



NATIONAL OPEN UNIVERSITY OF NIGERIA

FACULTY OF AGRICULTURAL SCIENCES

COURSE CODE: ANP201

**COURSE TITLE: INTRODUCTION TO ANIMAL
BIOTECHNOLOGY**

Course Material Instructor: Dr. Inuwa Shehu Usman,
Department of Plant Science,
Ahmadu Bello University, Zaria; Tel: 0803-451-1385, e-mail:
inuwasu@gmail.com

Course Objectives:

Course Aim: to introduce students into concepts and applied aspects of animal biotechnology that allow students to understand and discuss strategies for animal improvement through biotechnology.

This is a lecture, tutorial, practical based subject focussing on providing students with a theoretical and practical understanding of animal biotechnology. The subject covers animal molecular biology, recombinant DNA technology, production of transgenic animals, reproductive biotechnology, biotechnology in animal breeding and ethics

Specific Objectives:

1. To provide students with a scientific and technical understanding of animal biotechnology.
2. To introduce students to the commercial and ethical aspects of the biotechnology industry, and to challenge students with some of the moral and ethical issues that face biotechnologists, legislators and the general public.
3. To present biotechnology options for improving livestock production in developing countries and to encourage students to derive informed opinions on the potential benefit or danger of biotechnology and its impact on animal agriculture

Learning Outcomes:

After completion of this course the student will be able to:

1. Visualize the organization, structure and control of genes.
2. Understand the details of gene expression.
3. Describe posttranscriptional processing, initiation of translation

and posttranslational modifications, subcellular targeting, stability and degradation of RNA and proteins.

4. Understand the fundamental concepts and techniques for the use of recombinant DNA technology.
5. Describe the contribution biotechnology is making and is likely to make in animal animal science now and in the future.
6. Accomplishing the above course objectives will enable the students to apply these newly acquired skills in the critical evaluation of professional literature or scientific presentations in animal biotechnology. The students will be able to evaluate these skills during the course.

75 % of the course will be **lectures**

25 % of the course will be **laboratory demonstrations**

10 % of the course will be a **case study on ethical aspects** of transgenic crops

Case study on ethical aspects of transgenic crops:

Students will read and discuss conceptual papers reflecting the view on ethical aspects of transgenic plants of a specific group and students will present this view

Suggested Text:

Grading:

Exam objectives:

Continuous assessment 1: Gene expression and regulation

(genome organization, transcription, processing of transcripts, translation, protein folding and sorting, regulation of gene expression, degradation of RNA and proteins).

Continuous assessment 2: Concepts and techniques for the use of recombinant DNA technology, gene isolation, vector construction, genetic transformation, and applications of biotechnology to animal agriculture

Final exam (date to be announced): The final exam will evaluate the students' ability

to apply the acquired skills in the critical evaluation of professional literature. A scientific article in plant molecular genetics/biotechnology will be handed to students at least 1 week before the exam. During the exam questions will address molecular concepts, molecular techniques, results and conclusions associated with the article.

Expectations:

Students are expected to be prepared and participate in class discussions, ask questions and push for clarity. .

Course Outline:

1.0 History and concepts in biotechnology

1.1 History, foundations and Scope of animal biotechnology;

1.2 Scope of animal biotechnology

1.3 Mendelian Genetics

1.4 Nucleic acids

2.0 The Chemical Nature of Genes

2.1 Gene the basis of inheritance,

2.2 DNA replication; the genetic code; transcription; translation; RNA and proteins)

2.3 Regulation of gene expression

2.4 DNA Errors and repairs

2.5 Mutations

2.6 Laboratory Exercise on extraction of DNA and RNA from Animal Tissues

3.0 Principles of recombinant DNA technology

3.1 Recombinant DNA- Definition, How it functions and its importance

3.2 Recombination and Cloning

3.3 Transforming E.Coli to obtain products for human therapy

3.4 Molecular tools and techniques: Polymerase chain reaction (PCR), Gel

Electrophoresis, Restricted fragment length polymorphism (RFLP) and DNA Probes

4.0 Biotechnology Applications in Animal Agriculture

- 4.1 Reproductive physiology (Artificial insemination (AI) and Embryo transfer (ET)),
- 4.2 Transgenic animals,
- 4.3 Animal health (Disease diagnosis, Vaccines)
- 4.4 Physiology of lactation and growth
- 4.5 Animal nutrition (Increasing digestibility of low-quality forages, Improving nutritive value of cereals, Removing anti-nutritive factors from feeds, Improving rumen function)

5.0 Environmental and Biosafety Issues In Animal Biotechnology

- 5.1 Environmental concerns
- 5.2 Biosafety Considerations
- 5.3 Constraints
- 5.4 Building the capacity

Examinations will include multiple choice questions and short essays to determine the depth of the students understanding as well as their ability to communicate that understanding to others. Topics will be chosen from lectures and textbook(s) mainly, but occasionally from the media, popular press or scientific literature so as to cover the breadth of the topics.

Outcomes:

Students will better understand the several aspects of biotechnology. They will begin to make an association between animal and human health with development of technology. They will understand how to modify physiological processes to obtain biotechnological products to be applied to agricultural, social and medical areas.

Expectations

5

All students are encouraged to consult the course material and participate fully in interactive

practicals and discussions.

Prerequisites:

Course format:

This is a 3 day a week lecture course. The lectures will be supplemented with interactive and

virtual laboratories, videotapes, overheads and/or slides. Questions and discussion during lecture are encouraged.

MODULE ONE: **History, definitions and concepts in biotechnology**

1.1: History and the foundations of modern biotechnology

Biotechnology is defined as the application of scientific and engineering principles to the processing of material by biological agents to provide goods and services'. The Spinks Report (1980) defined biotechnology as the application of biological organisms, systems or processes to the manufacturing and service industries. United States Congress's Office of Technology Assessment defined biotechnology as 'any technique that used living organisms to make or modify a product, to improve plants or animals or to develop microorganisms for specific uses'. The document focuses, on the development and application of modern biotechnology based on new enabling techniques of recombinant-DNA technology, often referred to as genetic engineering.

The history of biotechnology begins with *zymotechnology*, which commenced with a focus on brewing techniques for beer. By *World War I*, however, zymotechnology would expand to tackle larger industrial issues, and the potential of industrial *fermentation* gave rise to biotechnology. The oldest biotechnological processes are found in microbial fermentations, as born out by a Babylonian tablet circa 6000 B.C., unearthed in 1881 and explaining the preparation of beer. In about 4000 B.C. leavened bread was produced with the aid of yeast. The Sumerians were able to brew as many as twenty types of beer in the third millennium B.C. In the 14th century, first vinegar manufacturing industry was established in France near Orleans. In 1680, Antony van Leeuwenhoek first observed yeast cells with his newly designed microscope. In 1857, Louis Pasteur highlighted the lactic acid fermentation by microbes. By the end of 19th century large number of industries and group

of scientists were involved in the field of biotechnology and developed large scale sewage purification system employing microbes were established in

Germany and France. In 1914 to 1916, Delbruck, Heyduck and Hennerberg discovered the large-scale use of yeast in food industry. In the same period, acetone, butanol and glycerin were obtained from bacteria. In 1920, Alexander Fleming discovered penicillin and large scale manufacturing of penicillin started in 1944.

Fermentation to Produce Foods Fermentation is perhaps the most ancient biotechnological discovery. Over 10,000 years ago mankind was producing wine, beer, vinegar and bread using microorganisms, primarily yeast. Yogurt was produced by lactic acid bacteria in milk and molds were used to produce cheese. These processes are still in use today for the production of modern foods. However, the cultures that are used have been purified and often genetically refined to maintain the most desirable traits and highest quality of the products.

In 1897, the discovery that enzymes from yeast can convert sugar to alcohol lead to industrial processes for chemicals such as butanol, acetone and glycerol. Fermentation processes are still in use today in many modern biotech organizations, often for the production of enzymes to be used in pharmaceutical processes, environmental remediation and other industrial processes. Drying, salting and freezing foods to prevent spoilage by microorganisms were practiced long before anyone really understood why they worked or even fully knew what caused the food to spoil in the first place.

1.2: Scope of Animal Biotechnology

The concept of animal tissue culture first emerged in 1903, when scientists discovered the technique of dividing cells in vitro (in a test tube). Ross

Harrison made the beginning of animal tissue culture technique in 1907 using frog tissue. This technique was initially confined to cold-blooded animals. However, subsequent studies brought even the warm-blooded animals into its sphere. Over the years various tissues have been used as explants, and tissue culture technique has indeed become the backbone of animal biotechnology. The advances in recombinant DNA technology have occurred in parallel with the development of genetic processes and biological variations. The development of new technologies have resulted into production of large amount of biochemically-defined proteins of medical significance and created an enormous potential for pharmaceutical industries.

The modern biotechnological tools have had a remarkable influence on animal biotechnology as well. Many innovative techniques are constantly being used around the world to improve upon livestock. The foundation of this approach lies in alteration at various biochemical and molecular levels. Biotechnology in itself is a vast subject and its scope is extended to various branches of biology. The details about these applications are treated in various modules of this course. These techniques are proving extremely useful in developing disease-resistant, healthy, and more productive animals. Figure 1 gives the summary of key areas of application of animal biotechnology.

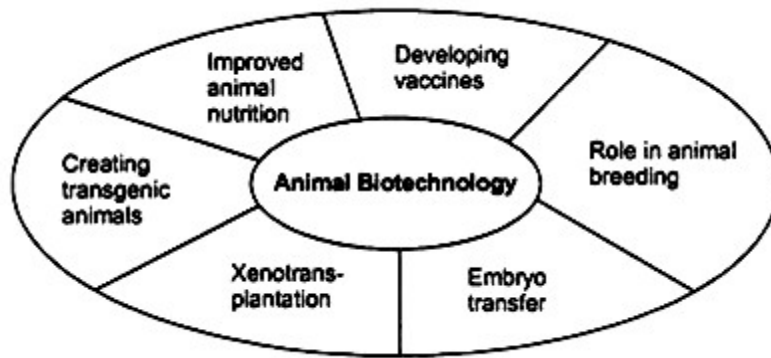


Figure 1: Schematics of Applications of Animal Biotechnology

1.3 Mendelian Genetics

For thousands of years farmers and herders have been selectively breeding their plants and animals to produce more useful [hybrids](#). It was somewhat of a hit or miss process since the actual mechanisms governing inheritance were unknown. Knowledge of these genetic mechanisms finally came as a result of careful laboratory breeding experiments carried out over the last century and a half.

By the 1890's, the invention of better microscopes allowed biologists to discover the basic facts of cell division and sexual reproduction. The focus of [genetics](#) research then shifted to understanding what really happens in the transmission of hereditary traits from parents to children. A number of hypotheses were suggested to explain heredity, but Gregor Mendel , a little known Central European monk, was the only one who got it more or less right. His ideas had been published in 1866 but largely went unrecognized until 1900, which was long after his death. His early adult life was spent in relative obscurity doing basic genetics research and teaching high school mathematics, physics, and

Greek in Brno (now in the Czech Republic). In his later years, he became the abbot of his monastery and put aside his scientific work.

While Mendel's research was with plants, the basic underlying principles of heredity that he discovered also apply to people and other animals because the mechanisms of heredity are essentially the same for all complex life forms.

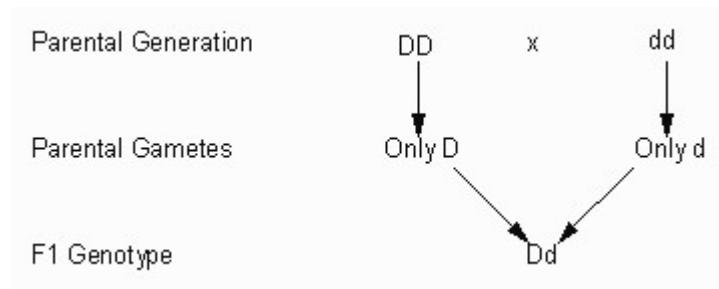
Through the selective cross-breeding of common pea plants (*Pisum sativum*) over many generations, Mendel discovered that certain traits show up in offspring without any blending of parent characteristics. For instance, the pea flowers are either purple or white--intermediate colors do not appear in the offspring of [cross-pollinated](#) pea plants.

Mendelian Genetics Definitions

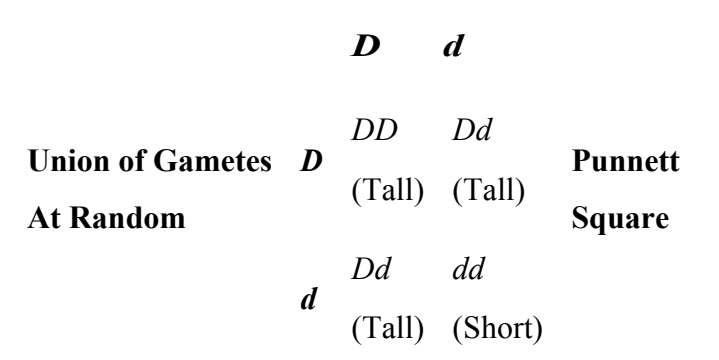
- **Allele** - one alternative form of a given allelic pair; tall and dwarf are the alleles for the height of a pea plant; more than two alleles can exist for any specific gene, but only two of them will be found within any individual
- **Allelic pair** - the combination of two alleles which comprise the gene pair
- **Homozygote** - an individual which contains only one allele at the allelic pair; for example DD is homozygous dominant and dd is homozygous recessive; pure lines are homozygous for the gene of interest
- **Heterozygote** - an individual which contains one of each member of the gene pair; for example the Dd heterozygote
- **Genotype** - the specific allelic combination for a certain gene or set of genes
- **Dominant** - the allele that expresses itself at the expense of an alternate allele; the phenotype that is expressed in the F₁ generation from the cross of two pure lines

- **Recessive** - an allele whose expression is suppressed in the presence of a dominant allele; the phenotype that disappears in the F₁ generation from the cross of two pure lines and reappears in the F₂ generation

Using symbols we can depict the cross of tall and short pea plants in the following manner:



The F₂ generation was created by selfing the F₁ plants. This can be depicted graphically in a Punnett square. From these results Mendel coined several other terms and formulated his first law. First the Punnett Square is shown.



The Punnett Square allows us to determine specific genetic ratios.

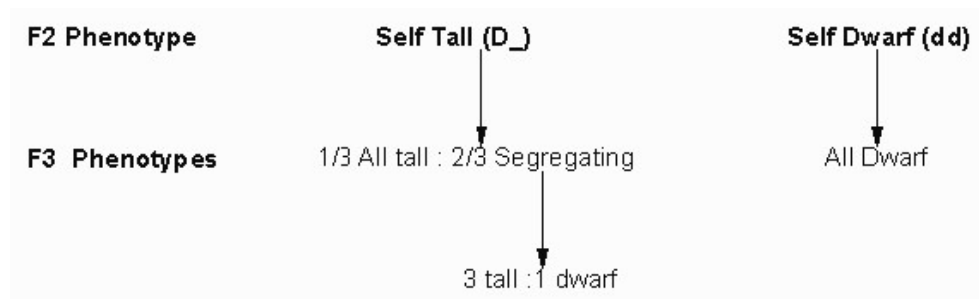
Genotypic ratio of F₂: 1 DD : 2 Dd : 1 dd (or 3 D₋ : 1 dd)

Phenotypic ratio of F₂: 3 tall : 1 dwarf

Mendel's First Law - the law of segregation; during gamete formation each member of the allelic pair separates from the other member to form the genetic constitution of the gamete

Confirmation of Mendel's First Law Hypothesis

With these observations, Mendel could form a hypothesis about segregation. To test this hypothesis, Mendel selfed the F₂ plants. If his law was correct he could predict what the results would be. And indeed, the results occurred as he expected.



From these results we can now confirm the genotype of the F₂ individuals.

Phenotypes	Genotypes	Genetic Description
F ₂ Tall Plants	1/3 <i>DD</i> 2/3 <i>Dd</i>	Pure line homozygote dominant Heterozygotes
F ₂ Dwarf Plants	all <i>dd</i>	Pure line homozygote recessive

Thus the F₂ is genotypically 1/4 *Dd* : 1/2 *Dd* : 1/4 *dd*

This data was also available from the Punnett Square using the gametes from the F₁ individual. So although the phenotypic ratio is 3:1 the genotypic ratio is 1:2:1

Mendel performed one other cross to confirm the hypothesis of segregation --- the backcross. Remember, the first cross is between two pure line parents to produce an F₁ heterozygote.



At this point instead of selfing the F₁, Mendel crossed it to a pure line, homozygote dwarf plant.

Backcross: $Dd \times dd$

	Male
	Gametes
	<i>D</i>
	<i>DD</i>
Female	<i>D</i> (Tall)
Gametes	<i>d</i> (Short)

Backcross One or (BC₁) Phenotypes: 1 Tall : 1 Dwarf

BC₁ Genotypes: 1 Dd : 1 dd

Backcross - the cross of an F₁ hybrid to one of the homozygous parents; for pea plant height the cross would be Dd x DD or Dd x dd; most often, though a backcross is a cross to a fully recessive parent

Testcross - the cross of any individual to a homozygous recessive parent; used to determine if the individual is homozygous dominant or heterozygous

So far, all the discussion has concentrated on monohybrid crosses.

Monohybrid cross - a cross between parents that differ at a single gene pair (usually $AA \times aa$)

Monohybrid - the offspring of two parents that are homozygous for alternate alleles of a gene pair

Remember --- a monohybrid cross is not the cross of two monohybrids.

Monohybrids are good for describing the relationship between alleles. When

Results from Mendel's Experiments

Parental Cross	F ₁ Phenotype	F ₂ Phenotypic Ratio	F ₂ Ratio
Round x Wrinkled Seed	Round	5474 Round:1850 Wrinkled	2.96:1
Yellow x Green Seeds	Yellow	6022 Yellow:2001 Green	3.01:1
Red x White Flowers	Red	705 Red:224 White	3.15:1
Tall x Dwarf Plants	Tall	1787 Tall:227 Dwarf	2.84:1

Terms and Results Found in the Table

Phenotype - literally means "the form that is shown"; it is the outward, physical appearance of a particular trait

Mendel's pea plants exhibited the following phenotypes:

- - round or wrinkled seed phenotype
- - yellow or green seed phenotype
- - red or white flower phenotype
- - tall or dwarf plant phenotype

Mendel observed seven traits that are easily recognized and apparently only occur in one of two forms:

1. flower color is purple or white
2. flower position is axil or terminal
3. stem length is long or short
4. seed shape is round or wrinkled
5. seed color is yellow or green
6. pod shape is inflated or constricted
7. pod color is yellow or green

Mendel picked common garden pea plants for the focus of his research because they can be grown easily in large numbers and their reproduction can be manipulated. Pea plants have both male and female reproductive organs. As a result, they can either self-pollinate themselves or cross-pollinate with another plant. In his experiments, Mendel was able to selectively cross-pollinate purebred plants with particular traits and observe the outcome over many generations. This was the basis for his conclusions about the nature of genetic inheritance. In cross-pollinating plants that either produce yellow or green pea seeds exclusively, Mendel found that the first offspring generation (f1) always has yellow seeds. However, the following generation (f2) consistently has a 3:1 ratio of yellow to green. This 3:1 ratio occurs in later generations as well. Mendel realized that this underlying regularity was the key to understanding the basic mechanisms of inheritance. He came to three important conclusions from these experimental results:

1. The hereditary determinants are of a particulate nature. These determinants are called genes.

2. Each parent has a gene pair in each cell for each trait studied. The F₁ from a cross of two pure lines contains one allele for the dominant phenotype and one for the recessive phenotype. These two alleles comprise the gene pair.
3. One member of the gene pair segregates into a gamete, thus each gamete only carries one member of the gene pair.
4. Gametes unite at random and irrespective of the other gene pairs involved

Mendel's observations from these experiments can be summarized in two principles:

1. the principle of segregation
2. the principle of independent assortment

According to the **principle of segregation**, for any particular trait, the pair of alleles of each parent separate and only one allele passes from each parent on to an offspring. Which allele in a parent's pair of alleles is inherited is a matter of chance. We now know that this segregation of alleles occurs during the process of sex cell formation (i.e., meiosis).

According to the **principle of independent assortment**, different pairs of alleles are passed to offspring independently of each other. The result is that new combinations of genes present in neither parent are possible. For example, a pea plant's inheritance of the ability to produce purple flowers instead of white ones does not make it more likely that it will also inherit the ability to produce yellow pea seeds in contrast to green ones. Likewise, the principle of independent assortment explains why the human inheritance of a particular eye color does not increase or decrease the likelihood of having 6 fingers on each hand. Today, we know this is due to the fact that the genes for independently assorted traits are located on different [chromosomes](#) .

Genetic analysis predates Gregor Mendel, but Mendel's laws form the theoretical basis of our understanding of the genetics of inheritance. These two principles of inheritance, along with the understanding of unit inheritance and dominance, were the beginnings of our modern science of genetics. However, Mendel did not realize that there are exceptions to these rules. What is seen in the F₁ generation? We always see only one of the two parental phenotypes in this generation. But the F₁ possesses the information needed to produce both parental phenotypes in the following generation. The F₂ generation always produced a 3:1 ratio where the dominant trait is present three times as often as the recessive trait. Mendel coined two terms to describe the relationship of the two phenotypes based on the F₁ and F₂ phenotypes. .

1.4: Nucleic Acids

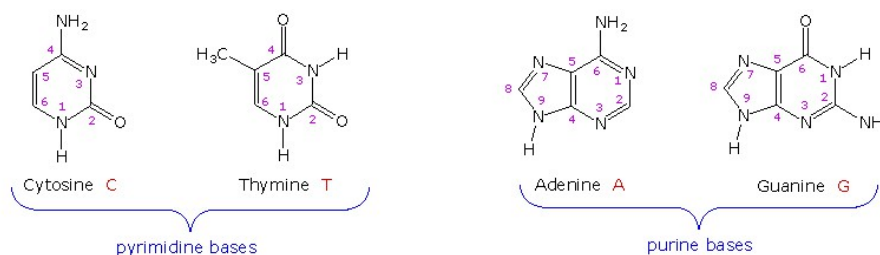
A nucleic acid is a long molecule made up of smaller molecules called nucleotides. Nucleic acids were discovered in 1868, when twenty-four-year-old Swiss physician Friedrich Miescher isolated a new compound from the nuclei of white blood cells. This compound was neither a protein nor lipid nor a carbohydrate; therefore, it was a novel type of biological molecule. Miescher named his discovery "nuclein," because he had isolated it from the nuclei of cells. Today, his discovery is known as deoxyribonucleic acid (DNA).

DNA is sometimes called "the blueprint of life" because it contains the code, or instructions for building and organism and ensuring that organism functions correctly. Just like a builder uses a blueprint to build a house, DNA is used as the blueprint, or plans, for the entire organism. DNA is material that governs inheritance of eye color, hair color, stature, bone density and many other human and animal traits. DNA is a long, but narrow string-like object. A one foot long string or strand of DNA is normally packed into a space roughly equal to a cube 1/millionth of an inch on a side. This is possible only because DNA is a very thin string.

Our body's cells each contain a complete sample of our DNA. One cell is roughly equal in size to the cube described in the previous paragraph. There are muscle cells, brain cells, liver cells, blood cells, sperm cells and others. Basically, every part of the body is made up of these tiny cells and each contains a sample or complement of DNA identical to that of every other cell within a given person. There are a few exceptions. For example, our red blood cells lack DNA. Blood itself can be typed because of the DNA contained in our white blood cells. Not only does the human body rely on DNA but so do most living things including plants, animals and bacteria.

Along with DNA, the other major type of nucleic acid in cells is ribonucleic acid (RNA). Both are shown in the above image, in single stranded form. DNA at on the top, and differs from RNA by its inclusion of thymine (T, in red), whereas RNA differs from DNA by its inclusion of uracil (U, yellow).

Nucleic acids are macromolecules, which means they are molecules composed of many smaller molecular units. These units are called nucleotides, and they are chemically linked to one another in a chain. In DNA, the nucleotides are referred to in shorthand as A, C, T, and G. In RNA, the nucleotides are A, C, U, and G. The order, or sequence, of the nucleotides in DNA allows nucleic acid to encode an organism's genetic blueprint.



The two monocyclic bases shown here are classified as **pyrimidines**, and the two bicyclic bases are **purines**. Each has at least one N-H site at which an organic substituent may be attached. They are all polyfunctional bases, and may exist in tautomeric forms.

Base-catalyzed hydrolysis of DNA gave four **nucleoside** products, which proved to be N-glycosides of 2'-deoxyribose combined with the heterocyclic amines. Structures and names for these nucleosides will be displayed above. The base components are colored green, and the sugar is black. As noted in the 2'-deoxycytidine structure on the left, the numbering of the sugar carbons makes use of primed numbers to distinguish them from the heterocyclic base sites. The corresponding N-glycosides of the common sugar ribose are the building blocks of RNA, and are named adenosine, cytidine, guanosine and uridine (a thymidine analog missing the methyl group). From this evidence, nucleic acids may be formulated as alternating copolymers of phosphoric acid (**P**) and nucleosides (**N**), as shown:



At first the four nucleosides, distinguished by prime marks in this crude formula, were assumed to be present in equal amounts, resulting in a uniform structure, such as that of starch. However, a compound of this kind, presumably common to all organisms, was considered too simple to hold the hereditary information known to reside in the chromosomes. This view was challenged in 1944, when Oswald Avery and colleagues demonstrated that bacterial DNA was likely the genetic agent that carried information from one organism to another in a process called "transformation". He concluded that *"nucleic acids must be regarded as possessing biological specificity, the chemical basis of which is as yet undetermined."* Despite this finding, many scientists continued to believe that chromosomal proteins, which differ across species, between individuals, and even within a given organism, were the locus of an organism's genetic information.

It should be noted that single celled organisms like bacteria do not have a well-defined nucleus. Instead, their single chromosome is associated with specific proteins in a region called a "nucleoid". Nevertheless, the DNA from bacteria

has the same composition and general structure as that from multicellular organisms, including human beings.

Views about the role of DNA in inheritance changed in the late 1940's and early 1950's. By conducting a careful analysis of DNA from many sources, Erwin Chargaff found its composition to be species specific. In addition, he found that the amount of adenine (A) always equaled the amount of thymine (T), and the amount of guanine (G) always equaled the amount of cytosine (C), regardless of the DNA source. As set forth in the following table, the ratio of (A+T) to (C+G) varied from 2.70 to 0.35. The last two organisms are bacteria.

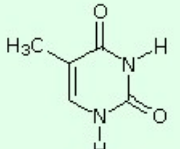
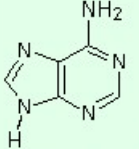
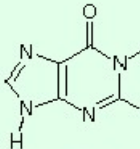
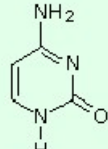
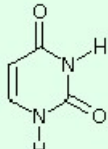
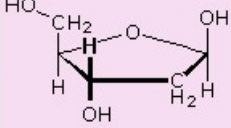
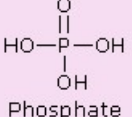
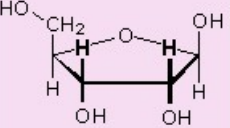
Nucleoside Base Distribution in DNA							
Organism	Base Composition (mole %)				Base Ratios		Ratio (A+T)/(G+C)
	A	G	T	C	A/T	G/C	
Human	30.9	19.9	29.4	19.8	1.05	1.00	1.52
Chicken	28.8	20.5	29.2	21.5	1.02	0.95	1.38
Yeast	31.3	18.7	32.9	17.1	0.95	1.09	1.79
Clostridium perfringens	36.9	14.0	36.3	12.8	1.01	1.09	2.70
Sarcina lutea	13.4	37.1	12.4	37.1	1.08	1.00	0.35

In a second critical study, Alfred Hershey and Martha Chase showed that when a bacterium is infected and genetically transformed by a virus, at least 80% of the viral DNA enters the bacterial cell and at least 80% of the viral protein remains outside. Together with the Chargaff findings this work established DNA as the repository of the unique genetic characteristics of an organism.

1.4.1 The Chemical Nature of DNA

The polymeric structure of DNA may be described in terms of monomeric units of increasing complexity. In the top shaded box of the following illustration, the three relatively simple components mentioned earlier are shown. Below that on the left, formulas for phosphoric acid and a nucleoside are drawn. [Condensation polymerization](#) of these leads to the DNA formulation outlined above. Finally, a 5'- monophosphate ester, called a **nucleotide** may be drawn as a single monomer unit, shown in the shaded box to the right. Since a monophosphate ester of this kind is a strong acid (pK_a of 1.0), it will be fully ionized at the usual physiological pH (ca.7.4). Names for these DNA components are given in the table to the right of the diagram. Isomeric 3'-monophosphate nucleotides are also known, and both isomers are found in cells. They may be obtained by selective hydrolysis of DNA through the action of nuclease enzymes. Anhydride-like di- and tri-phosphate nucleotides have been identified as important energy carriers in biochemical reactions, the most common being [ATP](#) (adenosine 5'-triphosphate).

Components of Nucleic Acids

	DNA only	DNA & RNA			RNA only
Nitrogen Bases	 <p>Thymine</p>	 <p>Adenine</p>	 <p>Guanine</p>	 <p>Cytosine</p>	 <p>Uracil</p>
Sugars & Phosphate	 <p>2-Deoxyribose</p>	 <p>Phosphate</p>			 <p>Ribose</p>

A complete structural representation of a segment of the DNA polymer formed from 5'-nucleotides may be viewed by [clicking on the above diagram](#). Several important characteristics of this formula should be noted.

- First, the remaining P-OH function is quite acidic and is completely ionized in biological systems.
- Second, the polymer chain is structurally directed. One end (5') is different from the other (3').
- Third, although this appears to be a relatively simple polymer, the possible permutations of the four nucleosides in the chain become very large as the chain lengthens.
- Fourth, the DNA polymer is much larger than originally believed. Molecular weights for the DNA from multicellular organisms are commonly 10^9 or greater.

Information is stored or encoded in the DNA polymer by the pattern in which the four nucleotides are arranged. To access this information the pattern must be "read" in a linear fashion, just as a bar code is read at a supermarket checkout. Because living organisms are extremely complex, a correspondingly large amount of information related to this complexity must be stored in the DNA. Consequently, the DNA itself must be very large, as noted above. Even the single DNA molecule from an *E. coli* bacterium is found to have roughly a million nucleotide units in a polymer strand, and would reach a millimeter in length if stretched out. The nuclei of multicellular organisms incorporate chromosomes, which are composed of DNA combined with nuclear proteins called histones. The fruit fly has 8 chromosomes, humans have 46 and dogs 78 (note that the amount of DNA in a cell's nucleus does not correlate with the number of chromosomes). The DNA from the smallest human chromosome is over ten times larger than *E. coli* DNA, and it has been estimated that the total DNA in a human cell would extend to 2 meters in length if unraveled. Since the nucleus is only about 5 μ m in diameter, the chromosomal DNA must be packed

tightly to fit in that small volume.

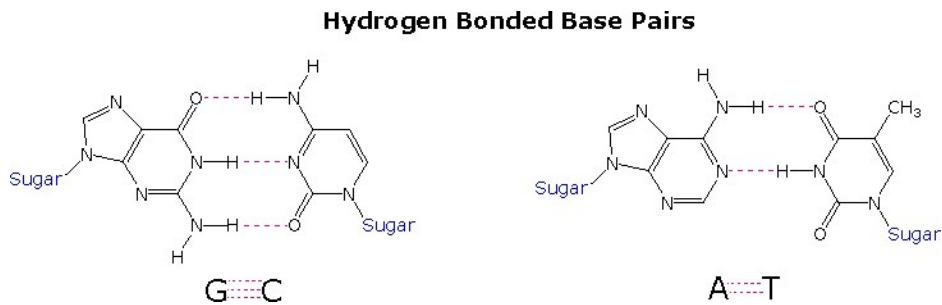
In addition to its role as a stable informational library, chromosomal DNA must be structured or organized in such a way that the chemical machinery of the cell will have easy access to that information, in order to make important molecules such as polypeptides. Furthermore, accurate copies of the DNA code must be created as cells divide, with the replicated DNA molecules passed on to subsequent cell generations, as well as to progeny of the organism. The nature of this DNA organization, or secondary structure, will be discussed in a later section.

1.4.1.1 The Secondary Structure of DNA

In the early 1950's the primary structure of DNA was well established, but a firm understanding of its secondary structure was lacking. Indeed, the situation was similar to that occupied by the proteins a decade earlier, before the alpha helix and pleated sheet structures were proposed by [Linus Pauling](#). Many researchers grappled with this problem, and it was generally conceded that the molar equivalences of base pairs (A & T and C & G) discovered by Chargaff would be an important factor. [Rosalind Franklin](#), working at King's College, London, obtained X-ray diffraction evidence that suggested a long helical structure of uniform thickness. Francis Crick and James Watson, at Cambridge University, considered hydrogen bonded base pairing interactions, and arrived at a double stranded helical model that satisfied most of the known facts, and has been confirmed by subsequent findings.

Once they had identified the favored base tautomers in the nucleosides, Watson and Crick were able to propose a complementary pairing, via hydrogen bonding, of guanosine (G) with cytidine (C) and adenosine (A) with thymidine (T). This pairing, which is shown in the following diagram, explained Chargaff's findings beautifully, and led them to suggest a double helix structure for DNA. Before viewing this double helix structure itself, it is instructive to examine the base

pairing interactions in greater detail. The G#C association involves three hydrogen bonds (colored pink), and is therefore stronger than the two-hydrogen bond association of A#T. These base pairings might appear to be arbitrary, but other possibilities suffer destabilizing steric or electronic interactions. The C#T pairing on the left suffers from carbonyl dipole repulsion, as well as steric crowding of the oxygens. The G#A pairing on the right is also destabilized by steric crowding (circled hydrogens).



A simple mnemonic device for remembering which bases are paired comes from the line construction of the capital letters used to identify the bases. A and T are made up of intersecting straight lines. In contrast, C and G are largely composed of curved lines. The RNA base uracil corresponds to thymine, since U follows T in the alphabet.

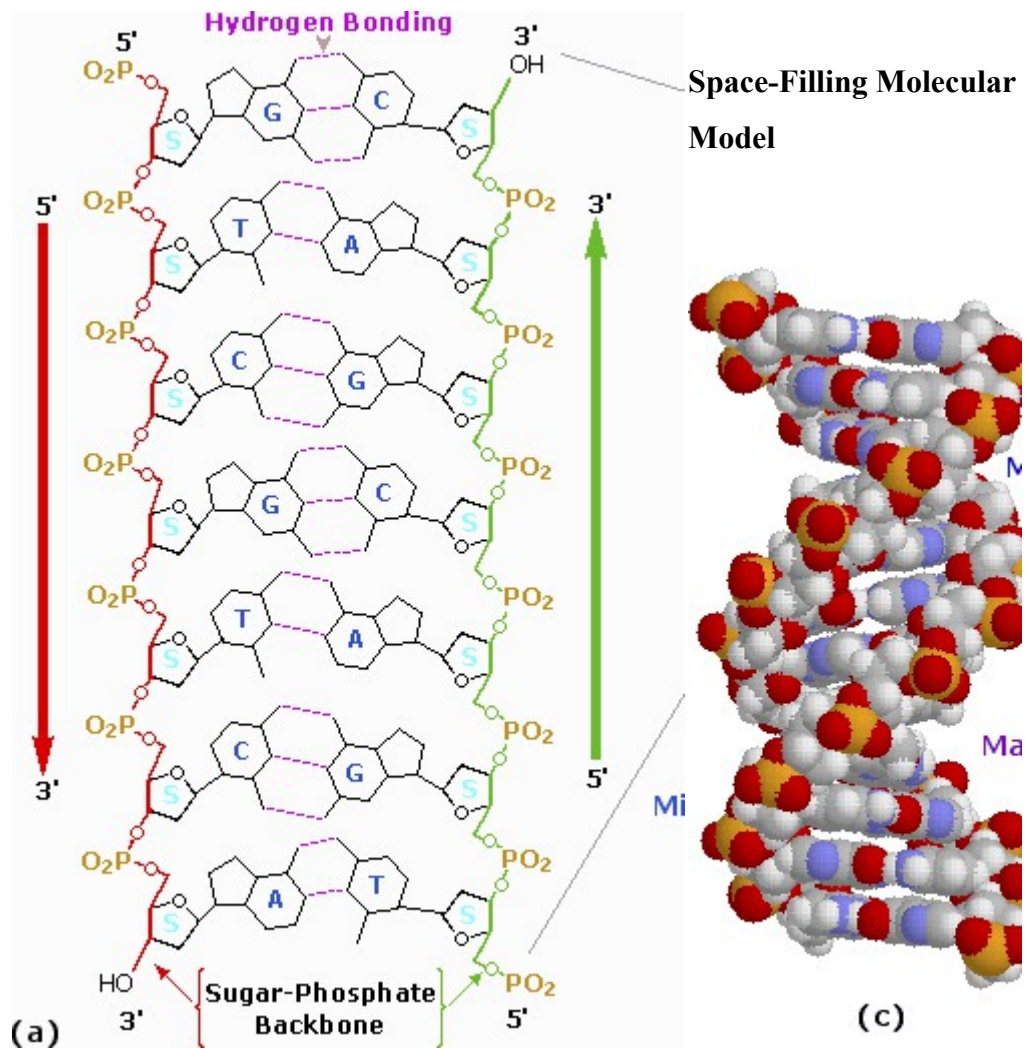
1.4.1.2 The Double Helix

After many trials and modifications, Watson and Crick conceived an ingenious double helix model for the secondary structure of DNA. Two strands of DNA were aligned anti-parallel to each other, i.e. with opposite 3' and 5' ends, as shown in part **a** of the following diagram. Complementary primary nucleotide structures for each strand allowed intra-strand hydrogen bonding between each pair of bases. These complementary strands are colored red and green in the diagram. Coiling these coupled strands then leads to a double helix structure,

shown as cross-linked ribbons in part **b** of the diagram. The double helix is further stabilized by hydrophobic attractions and pi-stacking of the bases. A space-filling molecular model of a short segment is displayed in part **c** on the right.

The helix shown here has ten base pairs per turn, and rises 3.4 Å in each turn. This right-handed helix is the favored conformation in aqueous systems, and has been termed the **B-helix**. As the DNA strands wind around each other, they leave gaps between each set of phosphate backbones. Two alternating grooves result, a wide and deep **major groove** (*ca.* 22Å wide), and a shallow and narrow **minor groove** (*ca.* 12Å wide). Other molecules, including polypeptides, may insert into these grooves, and in so doing perturb the chemistry of DNA. Other helical structures of DNA have also been observed, and are designated by letters (e.g. **A** and **Z**).

The Double Helix Structure for DNA



A model of a short DNA

A strand of DNA is made up of tiny building-blocks. There are only four, different basic building-blocks. Scientists usually refer to these using four letters, A, T, G, and C. These four letters are short nicknames for more complicated building-block chemical names, but actually the letters (A, T, G and C) are used much more commonly than the chemical names so the latter will not be mentioned here. Another term for DNA's building blocks is the term, "bases." A, T, G and C are bases.

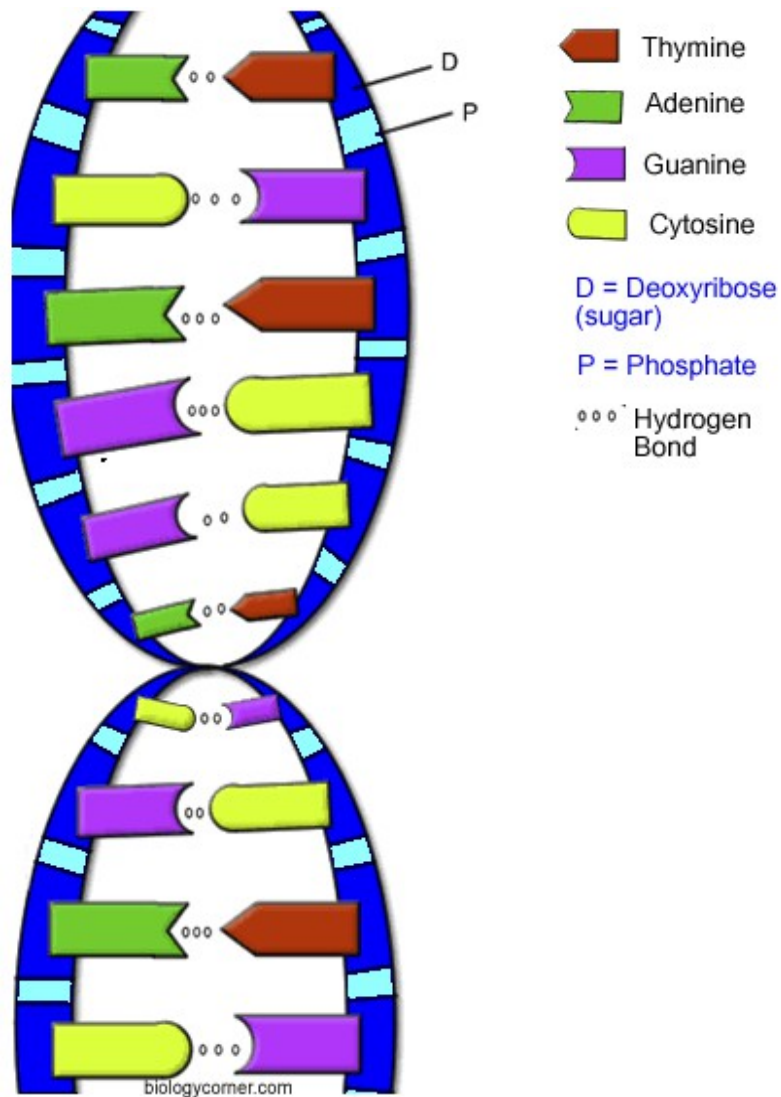
For example, to refer to a particular piece of DNA, we might write: AATTGCCTTTTAAAAA. This is a perfectly acceptable way of describing a piece of DNA. Someone with a machine called a DNA synthesizer could actually synthesize the same piece of DNA from the information AATTGCCTTTTAAAAA alone.

The sequence of bases (letters) can code for many properties of the body's cells. The cells can read this code. Some DNA sequences encode important information for the cell. Such DNA is called, not surprisingly, "coding DNA." Our cells also contain much DNA that doesn't encode anything that we know about. If the DNA doesn't encode anything, it is called non-coding DNA or sometimes, "junk DNA."^[1]

The DNA code, or genetic code as it is called, is passed through the sperm and egg to the offspring. A single sperm cell contains about three billion bases consisting of A, T, G and C that follow each other in a well defined sequence along the strand of DNA. Each egg cell also contains three billion bases arranged in a well-defined sequence very similar, but not identical to the sperm.

Both coding and non-coding DNAs may vary from one individual to another. These DNA variations can be used to identify people or at least distinguish one person from another.

The structure of DNA was established by James Watson and Francis Crick.



The shape of the DNA molecule is a double-helix (like a twisted ladder). The sides of the ladder are composed of alternating sugars (deoxyribose) and phosphates. The rungs of the ladder are composed of nucleotides.

Nucleotides (also called Bases)

Adenine, Thymine, , Guanine, Cytosine or A, T, G, C

Nucleotides pair in a specific way - called the Base-Pair Rule

Adenine pairs to Thymine

Guanine pairs to Cytosine

Memory helper - think "A T Granite City") - which is where you live. The rungs of the ladder can occur in any order (as long as the base-pair rule is followed)

For instance, a stretch of DNA could be AATGACCAT - which would code for a different gene than a stretch that read: GGGCCATAG. All in all, there are billions of bases (nucleotides) in cells, which code for all the things an organism needs to function. The DNA molecule is a double helix, and it is composed of three main parts:

- Five-carbon sugar (deoxyribose)
 - Phosphate molecule
 - One of four nitrogen-containing bases
 - § Adenine(A)
 - § Guanine(G)
 - § Cytosine(C)
 - § Thymine(T)
- The double helix is like a ladder. Two anti-parallel strands are comprised of sugars and phosphates, and the bases comprise the "rungs" of the "ladder."
 - The message encoded in DNA is made up of the four-letter alphabet A, G, C, and T, with each letter representing one of the bases.

1.4.2 RNA, a Different Nucleic Acid

The high molecular weight nucleic acid, DNA, is found chiefly in the nuclei of complex cells, known as **eucaryotic cells**, or in the nucleoid regions of **procaryotic cells**, such as bacteria. It is often associated with proteins that help

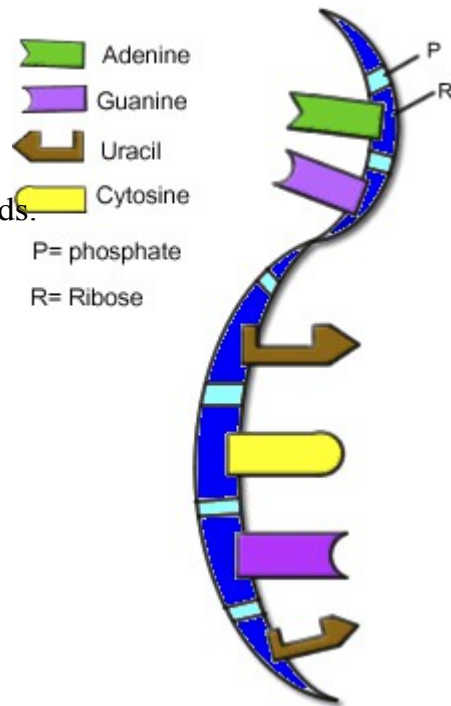
to pack it in a usable fashion. In contrast, a lower molecular weight, but much more abundant nucleic acid, **RNA**, is distributed throughout the cell, most commonly in small numerous organelles called **ribosomes**. Three kinds of RNA are identified, the largest subgroup (85 to 90%) being ribosomal RNA, **rRNA**, the major component of ribosomes, together with proteins. The size of rRNA molecules varies, but is generally less than a thousandth the size of DNA. The other forms of RNA are messenger RNA, **mRNA**, and transfer RNA, **tRNA**. Both have a more transient existence and are smaller than rRNA. All these RNA's have similar constitutions, and differ from DNA in two important respects. As shown in the following diagram, the sugar component of RNA is ribose, and the pyrimidine base uracil replaces the thymine base of DNA. The RNA's play a vital role in the transfer of information (transcription) from the DNA library to the protein factories called ribosomes, and in the interpretation of that information (translation) for the synthesis of specific polypeptides. These functions will be described later.

DNA remains in the nucleus, but in order for it to get its instructions translated into proteins, it must send its message to the ribosomes, where proteins are made. The chemical used to carry this message is **Messenger RNA**

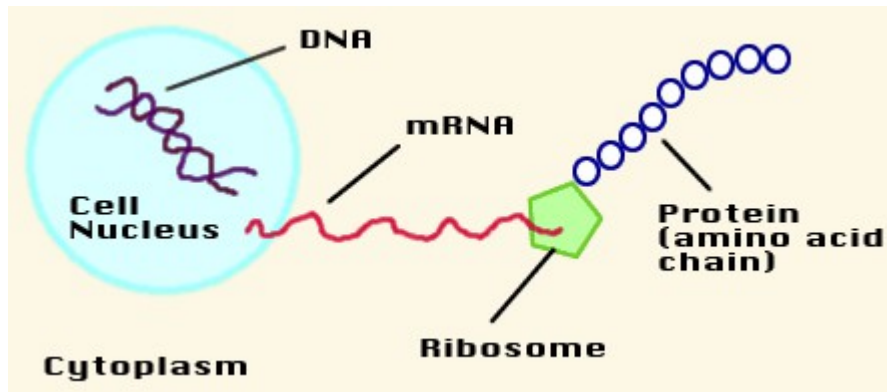
RNA = ribonucleic acid.

RNA is similar to DNA except:

1. has one strand instead of two strands.
2. has uracil instead of thymine
3. has ribose instead of deoxyribose



mRNA has the job of taking the message from the DNA to the nucleus to the ribosomes.



References:

1. Pujari, S. Animal Biotechnology: An Introduction to Animal Biotechnology. Copyright © 2015 YourArticleLibrary.com

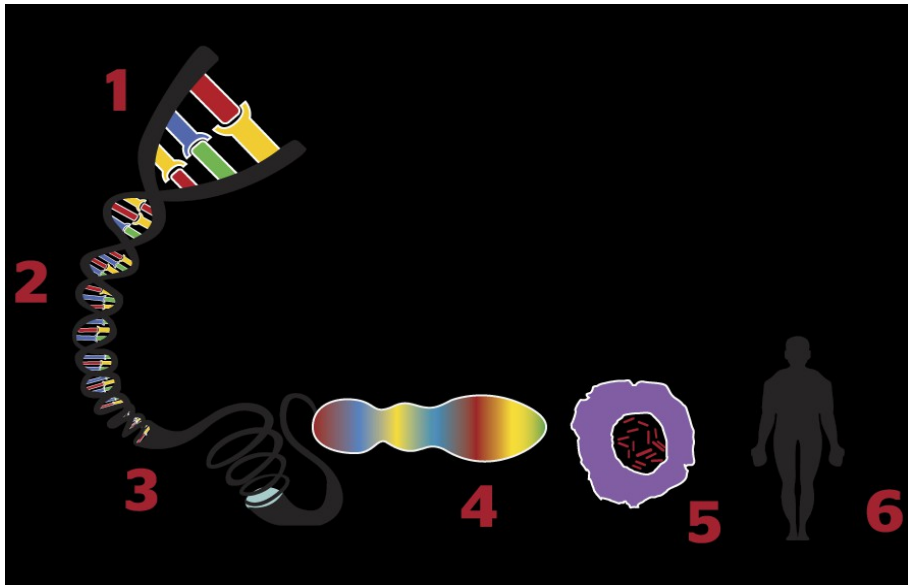
2. Mendelian Genetics. Copyright © 1997-2013 by Dennis O'Neil. All rights reserved
3. Nucleic Acids. This page is the property of William Reusch. Comments, questions and errors should be sent to whreusch@msu.edu

MODULE TWO

The Chemical Nature of Genes

2.1 Gene the basis of Inheritance

Genes are small segments of a long molecule called DNA that is composed of four different nucleotides: adenine (A), thiamine (T), cytosine (C) and guanidine (G). DNA is a double stranded molecule (A always pairs with T, and C with G). These number and order of the bases spell out the language known as the [genetic code](#), which is universal to all life forms. The nucleotides in a [gene](#) encode a recipe or template to produce a protein; the information in genes is sufficient to produce all the proteins the cells of an organism need. Genes instruct cells how to function, and what characteristic they should [express](#). For example, in flowers, the blue pigment expressed in petal cells. Gene activity is tightly [regulated](#).



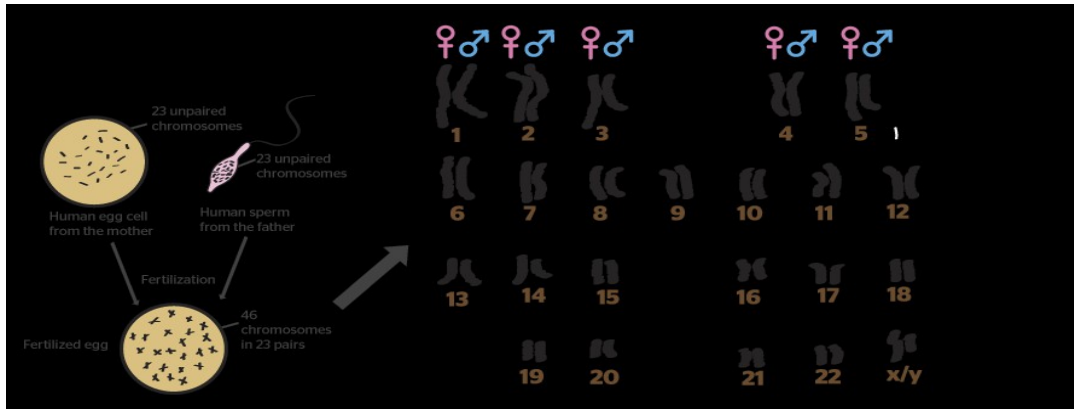
2.1.1 Genes and chromosomes

It is the chemical component of chromosomes, which are located in the nucleus of every cell. Stretches of DNA (or stretches of chromosomes) code for genes. **Gene** - a segment of DNA that codes for a protein, which in turn codes for a trait (skin tone, eye color..etc), a gene is a stretch of DNA. A locus (with a hard "c", LOW-KUS) is simply a location in the DNA. The plural of locus is, loci (with a soft "c", pronounced LOW-S-EYE). Again, the DNA is a long string like object as illustrated below. A locus is simply a location in the DNA. Such locations, or loci, reside at specific places on chromosomes. Extended DNA molecules would be very long (human DNA would be two meters long) but DNA is coiled and packaged in structures known as *chromosomes*. Chromosomes are contained in the cell's nucleus. Different species have different numbers of chromosomes (humans have 46 chromosomes, or 23 sets of chromosome pairs; peas have 14 chromosomes or 7 pairs; and tomatoes 24 chromosomes or 12 pairs).

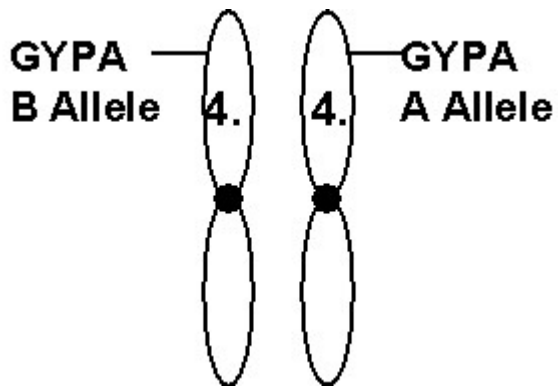
When a cell is getting ready to divide creating two daughter cells, it packs its DNA into bundles called chromosomes. Chromosomes are just bundles of DNA. For humans, there are consistently 23 pairs of chromosomes, each with a consistent size and shape. Chromosomes are numbered. Chromosome number 1 is the largest chromosome; chromosome number 2 a little smaller and so on. Among the 23 pairs of chromosomes there is a pair called the sex chromosomes. This is something of a misnomer, since there are many functions on the "sex" chromosomes that have nothing to do with sex. In females, the sex-chromosome pair consists of two similar size chromosomes called X chromosomes. Males have one X and one small Y chromosome.

Most adult cells contain two sets of chromosomes. However, sexual cells (sperm cells from the father and egg cells from the mother) contain a single chromosome

set. During reproduction, each parent contributes one set of chromosomes to their offspring.



Unless it has been purified, our DNA is actually not a loosely tangled string as illustrated but rather is well organized and packaged into what are called chromosomes. A chromosome is a tightly folded bundle of DNA. Chromosomes are most visible when cells divide. In a microscope, chromosomes look something like this without the numbers and letters:



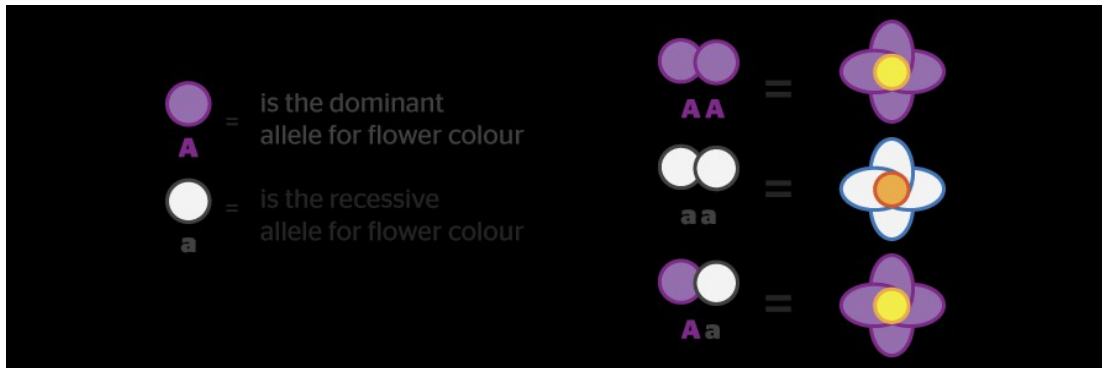
The illustration shows a pair of chromosomes named chromosome number 4, one pair among 23 pairs of chromosomes. The illustration also shows the position of a locus that happens to be called "GYPA." In this example, the

chromosome on the left has the variation called the B allele while the chromosome on the right has the variation called the A allele.

2.1.2 Alleles

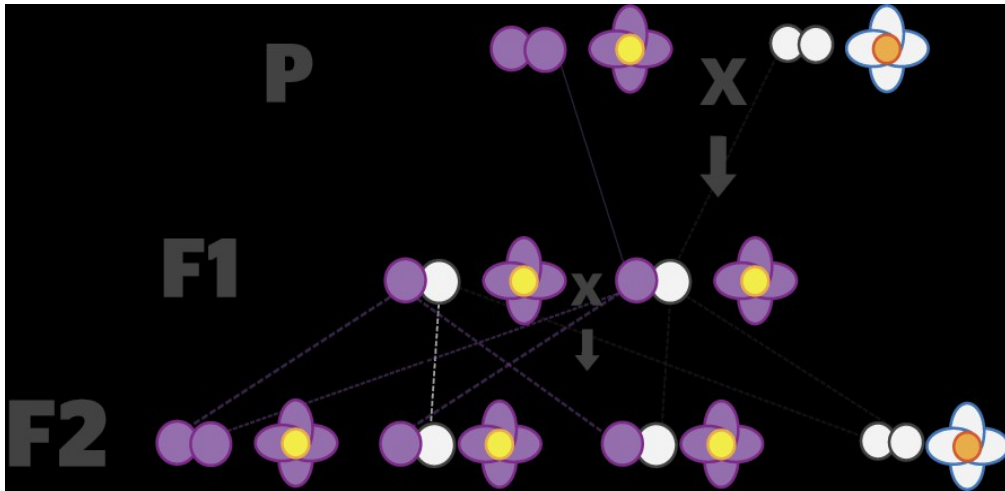
Alleles (ALL-EELS') are just variations at a particular site on a chromosome. Since each chromosome has a similar chromosome partner (except for males with their X and Y chromosomes) each locus is duplicated. Loci can vary a bit. If a person has two identical versions of the locus, they are said to be homozygous (HOMO-Z-EYE'-GUS). If there is a difference, they are said to be heterozygous (HETERO-Z-EYE'-GUS). Different forms of a gene are called alleles. Sometimes different alleles can result in variation in the trait expressed, such as flower colour. Each gene has two alleles (one in each paired chromosome). An individual with two copies of the same allele is called homozygous. An individual with two copies of different alleles is called heterozygous.

The interactions between the two alleles of a gene can be described as dominant or recessive. A dominant allele will mask the presence of a recessive allele.



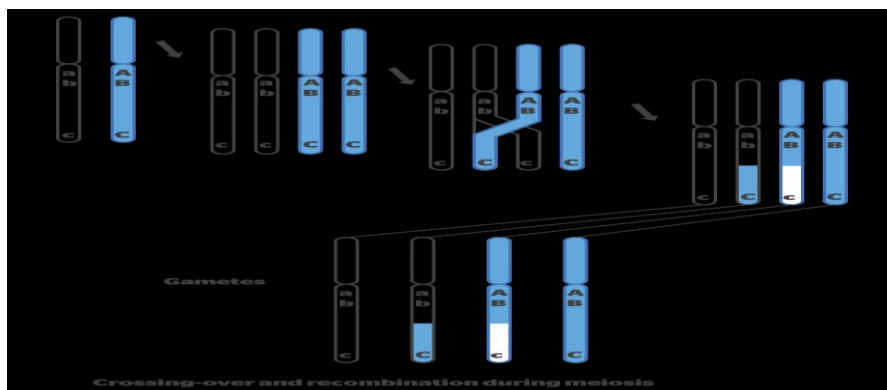
When two parents homozygous for different alleles are crossed, the progeny (also called F1, for filial generation 1) is an heterozygote. F1 generations of homozygous parents are very even; they are uniformly heterozygous. This is also true for commercial hybrid seed. In the following generation (F2) the alleles

segregate independently, forming the basis for genetic variability. Sexual reproduction also introduces genetic variation due to recombination.



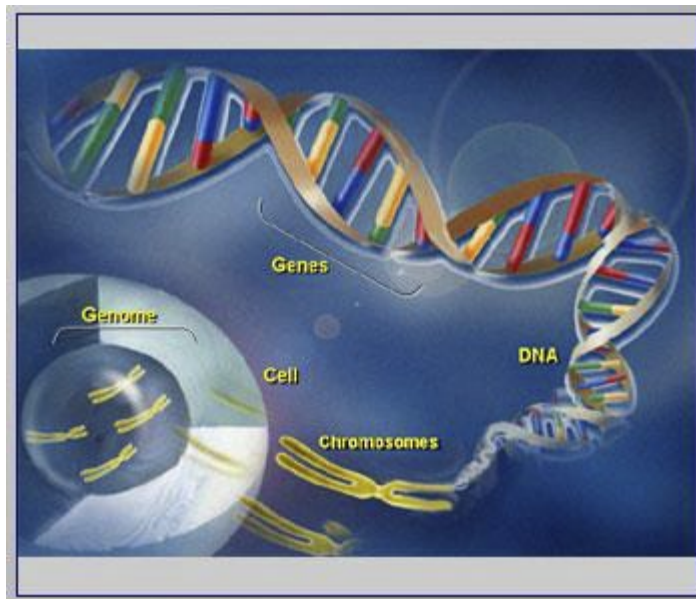
The process of sexual reproduction involves two parents, each contributing one gamete. Gametes are produced by a process called meiosis, which starts by the duplication of the chromosomes, followed by two rounds of cell divisions and halving of the chromosome number. Gametes have half the chromosome number of other adult cells of an organism.

A key feature of meiosis is the exchange of chromosome pieces which occurs in the first division of this process, called recombination or crossing over. Recombination is a very important source of genetic variation between individuals of sexually reproducing species, and the driving force for the process of natural selection.



- Genes are the units of heredity and are the instructions that make up the body's blueprint. They code for the proteins that determine virtually all of a person's characteristics.
- Humans have an estimated 35,000 genes.
- Most genes come in pairs and are made of strands of genetic material called *deoxyribonucleic acid*, or *DNA*.
- Genetic disorders are caused by one or more changes, or *mutations*, in the instruction code of a particular gene(s), preventing the gene(s) from functioning properly.
- The study of human genetics is the study of human variation that is carried in the genes.
- The physical location of a gene is its *locus*. Different versions of genes are called *alleles*. For example, an eye color gene may have a blue allele and a brown allele.

Genes are organized in structures called *chromosomes*

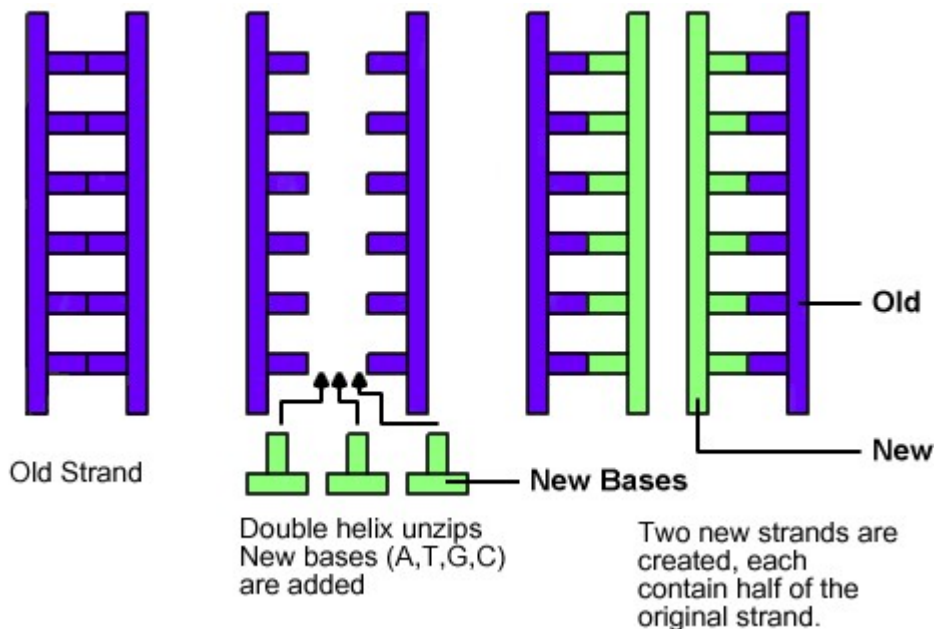


2.2 DNA replication

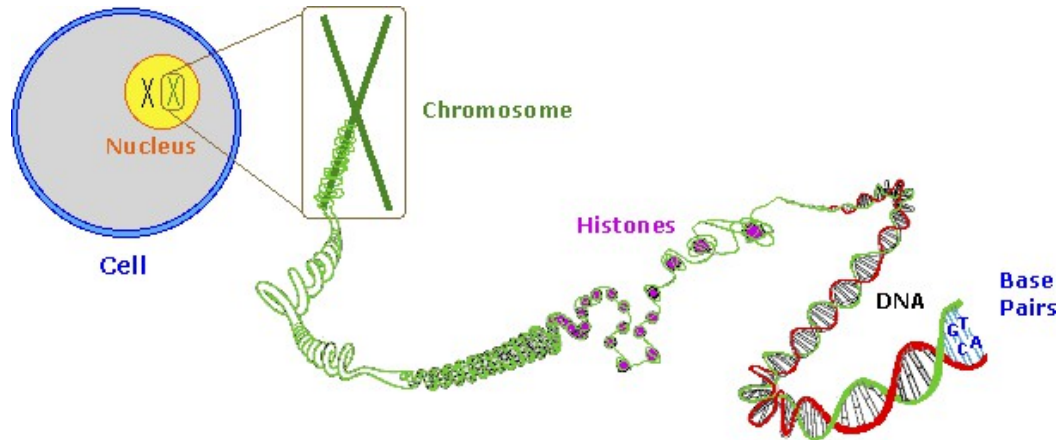
This unit explains the mechanisms of DNA replication . In their 1953 announcement of a double helix structure for DNA, Watson and Crick stated, *"It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material."* The essence of this suggestion is that, if separated, each strand of the molecule might act as a template on which a new complementary strand might be assembled, leading finally to two identical DNA molecules. Indeed, replication does take place in this fashion when cells divide, but the events leading up to the actual synthesis of complementary DNA strands are sufficiently complex that they will not be described in any detail.

Replication is the process where DNA makes a copy of itself. Why does DNA need to copy? Simple: Cells divide for an organism to grow or reproduce, every new cell needs a copy of the DNA or instructions to know how to be a cell. DNA replicates right before a cell divides.

DNA replication is **semi-conservative**. That means that when it makes a copy, one half of the old strand is always kept in the new strand. This helps reduce the number of copy errors.



As depicted in the following drawing, the DNA of a cell is tightly packed into chromosomes. First, the DNA is wrapped around small proteins called histones (colored pink below). These bead-like structures are then further organized and folded into chromatin aggregates that make up the chromosomes. An overall packing efficiency of 7,000 or more is thus achieved. Clearly a sequence of unfolding events must take place before the information encoded in the DNA can be used or replicated.

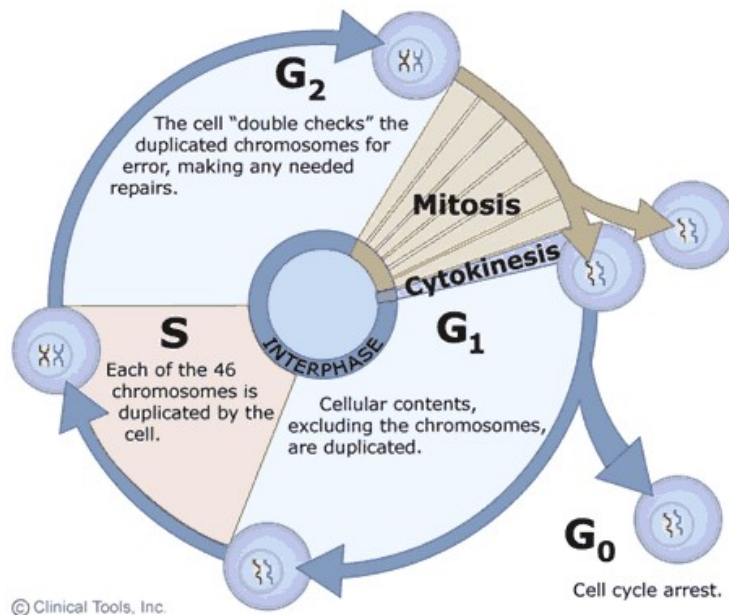


Once the double stranded DNA is exposed, a group of enzymes act to accomplish its replication. The DNA being replicated must be in a ready state for the start of replication, and there also has to be a clear start point from which replication proceeds. As each piece of DNA must only be copied one, there also has to be an end point to replication.

DNA replication must be carried out accurately, with an efficient proof reading and repair

mechanism in place for any mismatches or errors. And finally, the system of replication must also be able to distinguish between the original DNA template and then newly copied DNA. In order to be able to put these principles into context, it is helpful to look at the eukaryotic cell cycle to see where the main checkpoints are in the process.

Actively dividing eukaryote cells pass through a series of stages known collectively as the **cell cycle**: two gap phases (G1 and G2); an S (for synthesis) phase, in which the genetic material is duplicated; and an M phase, in which mitosis partitions the genetic material and the cell divides.



G1 phase. Metabolic changes prepare the cell for division. At a certain point - the restriction point - the cell is committed to division and moves into the S phase.

S phase. DNA synthesis replicates the genetic material. Each chromosome now consists of two sister chromatids.

G2 phase. Metabolic changes assemble the cytoplasmic materials necessary for mitosis and cytokinesis.

M phase. A nuclear division (mitosis) followed by a cell division (cytokinesis).

The period between mitotic divisions - that is, G1, S and G2 - is known as interphase. The main check points in DNA replication occur: between G1 phase and S phase: at the start of mitosis (M phase) and finally between M phase and G1 phase, when the decision is made whether to go quiescent or not. Replication in terms of the cell:

Must be ready: G1

Must all start at the same time: G1 - S

Must know where to start: Origin of Replication

Must all finish: Complete S

Must ensure that each piece of DNA is only replicated once, so need to know where to end: Replicon

■ Proof reading and repair: G2

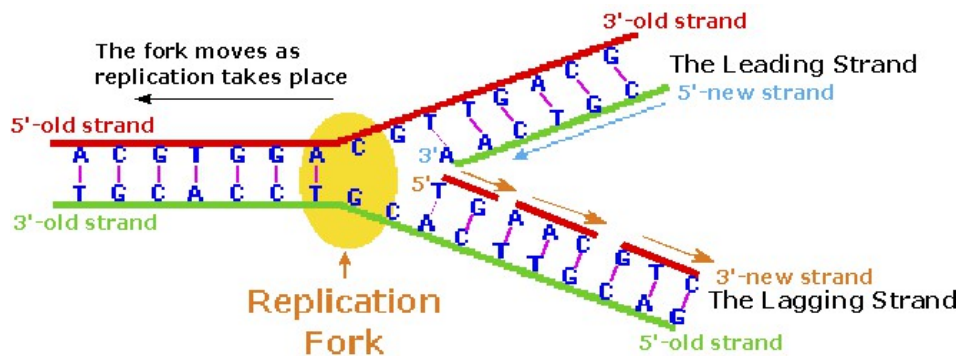
■

Able to distinguish between original and copy: Epigenetics

Because of the directional demand of the polymerization, one of the DNA strands is easily replicated in a continuous fashion, whereas the other strand can only be replicated in short segmental pieces. This is illustrated in the following diagram. Separation of a portion of the double helix takes place at a site called the **replication fork**. As replication of the separate strands occurs, the replication fork moves away (to the left in the diagram), unwinding additional lengths of DNA. Since the fork in the diagram is moving toward the 5'-end of the red-colored strand, replication of this strand may take place in a continuous fashion (building the new green strand in a 5' to 3' direction). This continuously formed new strand is called the **leading strand**. In contrast, the replication fork moves toward the 3'-end of the original green strand, preventing continuous polymerization of a complementary new red strand. Short segments of complementary DNA, called Okazaki fragments, are produced, and these are linked together later by the enzyme **ligase**. This new DNA strand is called the **lagging strand**.

Replication Summary

1. DNA replication occurs during the S Phase of the cell cycle
2. It is semi-conservative, i.e. it produces two copies that each contain one of the original *strands* and **one new strand**
3. It takes place in a 5' to 3' direction with a leading strand and a lagging strand (which is discontinuous), and the use of an RNA primer



4. In bacteria, there is only a single origin of replication
5. In eukaryotes, there are multiple origins of replication
6. Replication is bi-directional

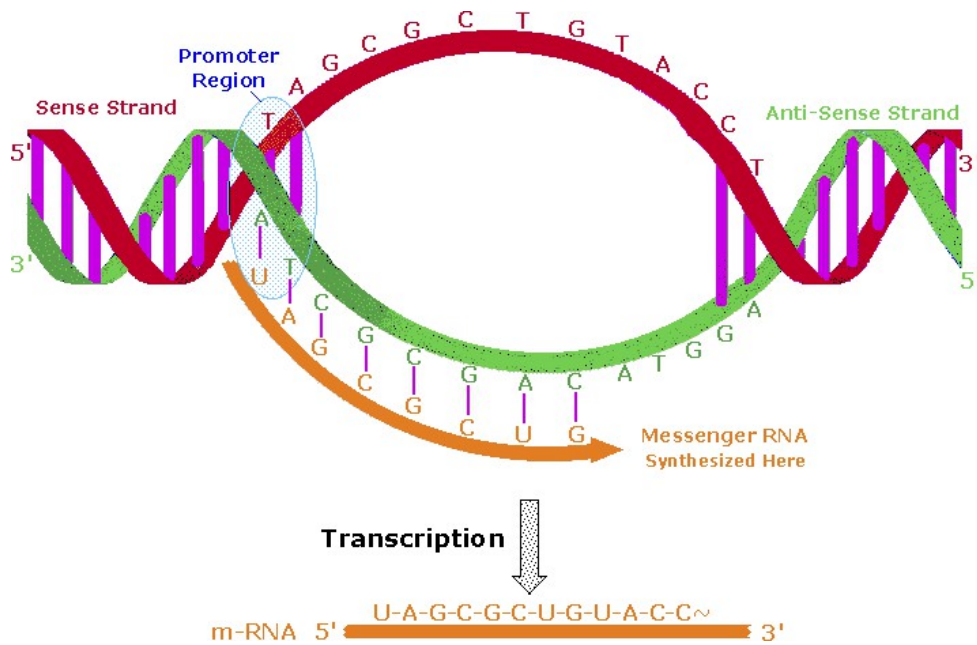
7. In yeast, the ARS (Autonomously Replicating Sequence) element is present. Yeast can be considered to be the eukaryotic equivalent of *E. coli*, and it has approximately 400 ARS elements in 12 chromosomes.

Replication Enzymes

- DNA Polymerase: Matches the correct nucleotide and then joins adjacent nucleotides together
- Primase: Provides an RNA primer to start polymerisation
- Ligase: Joins adjacent DNA strands together
- Helicase: Unwinds the DNA and melts it
- Single Strand Binding Proteins: Keep the DNA single stranded after it has been melted by helicase
- Gyrase: A topoisomerase that relieves torsional strain in the DNA molecule
- Telomerase: Finishes off the ends of the DNA strand

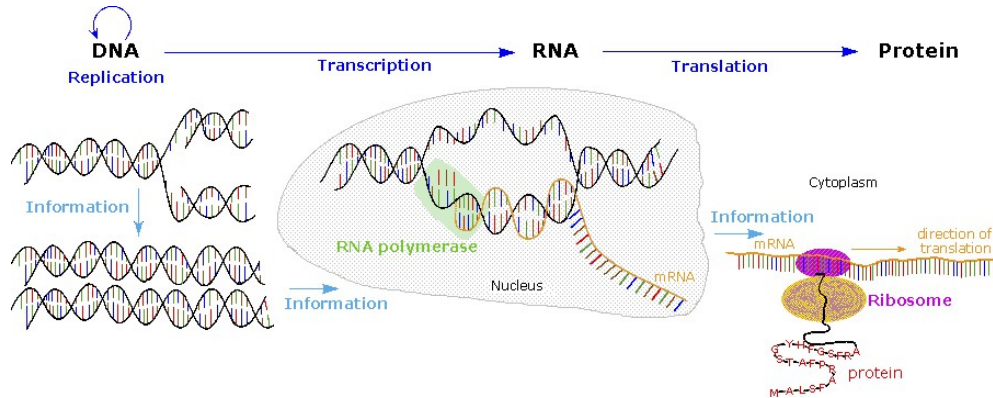
2. 2.3 Transcription

RNA is made from DNA. Francis Crick proposed that information flows from DNA to RNA in a process called **transcription**, and is then used to synthesize polypeptides by a process called **translation**. Transcription takes place in a manner similar to DNA replication. A characteristic sequence of nucleotides marks the beginning of a gene on the DNA strand, and this region binds to a promoter protein that initiates RNA synthesis. The double stranded structure unwinds at the promoter site., and one of the strands serves as a template for RNA formation, as depicted in the following diagram. The RNA molecule thus formed is single stranded, and serves to carry information from DNA to the protein synthesis machinery called ribosomes. These RNA molecules are therefore called **messenger-RNA (mRNA)**.



An important distinction must be made here. One of the DNA strands in the double helix holds the genetic information used for protein synthesis. This is called the **sense strand**, or information strand (colored red above). The complementary strand that binds to the sense strand is called the **anti-sense strand** (colored green), and it serves as a template for generating a mRNA molecule that delivers a copy of the sense strand information to a ribosome. The promoter protein binds to a specific nucleotide sequence that identifies the sense strand, relative to the anti-sense strand. RNA synthesis is then initiated in the 3' direction, as nucleotide triphosphates bind to complementary bases on the template strand, and are joined by phosphate diester linkages. An animation of this process for DNA replication was [presented earlier](#). A characteristic "stop sequence" of nucleotides terminates the RNA synthesis. The messenger molecule (colored orange above) is released into the cytoplasm to find a ribosome, and the DNA then rewinds to its double helix structure. In eucaryotic cells the initially transcribed m-RNA molecule is usually modified and shortened by an "editing" process that removes irrelevant material. The DNA of such organisms is often thousands of times larger and more complex than that composing the single chromosome of a procaryotic bacterial cell. This difference is due in part to repetitive nucleotide sequences (ca. 25% in the human genome). Furthermore, over 95% of human DNA is found in intervening sequences that separate genes and parts of genes. The informational DNA segments that make up genes are called **exons**, and the noncoding segments are called **introns**. Before the mRNA molecule leaves the nucleus, the nonsense bases that make up the introns are cut out, and the informationally useful exons are joined together in a step known as **RNA splicing**. In this fashion shorter mRNA molecules carrying the blueprint for a specific protein are sent on their way to the ribosome factories.

The **Central Dogma** of molecular biology, which at first was formulated as a simple linear progression of information from DNA to RNA to Protein, is summarized in the following illustration. The replication process on the left consists of passing information from a parent DNA molecule to daughter molecules. The middle transcription process copies this information to a mRNA molecule. Finally, this information is used by the chemical machinery of the ribosome to make polypeptides.



2.2.4 The genetic code

2.2.4.1 Codons and Building Proteins

- DNA is the code used to produce proteins.
 - First, the DNA is replicated, so that there are two identical copies of the DNA. Then, the DNA is transcribed into RNA, which is then translated, or read, by tRNA to make the proteins.
 - This process of replication, transcription, and translation, is called the “*central dogma*”.
 - Just like DNA, RNA is composed of a four-letter alphabet. However, the thymine (T) in DNA is replaced by a uracil (U) in RNA.
 - The message encoded in RNA is read in three-letter words called *codons*.
- Codons code for specific amino acids, which are the building blocks of proteins. Therefore, by knowing the sequence of bases in a gene, it is possible to predict the codons and, ultimately, the amino acid sequence of the protein the gene makes.
- There are twenty possible amino acids. However, most amino acids can be coded for by more than one codon.
 - The beginning of a coding sequence is signaled by a start codon, a unique sequence for this purpose; the start codon also codes for a methionine.

- There are three codons that indicate the end of the amino acid sequence. These are called stop codons.

		Second Letter											
		U			C			A			G		
First Letter	U	UUU	Phenylalanine	UCU	Serine	UAU	Tyrosine	UGU	Cysteine	U			
		UUC		UCC		UAC		UGC		C			
		UUA	Leucine	UCA		UAA	Stop codon	UGA	Stop codon	UAG	A		
	UUG		UCG	UAG	Stop codon	UGG	Tryptophan		G				
	C	CUU	Leucine	CCU	Proline	CAU	Histidine	CGU	Arginine	U			
		CUC		CCC		CAC		CGC		C			
		CUA		CCA		CAA	CGA	A					
		CUG		CCG		CAG	CGG	G					
	A	AUU	Isoleucine	ACU	Threonine	AUU	Asparagine	AGU	Serine	U			
		AUC		ACC		AAC		AGC		C			
		AUA	Methionine, initiation codon	ACA		AAA	Lysine	AGA	Arginine	A			
		AUG	ACG	AAG		AGG		G					
	G	GUU	Valine	GCU	Alanine	GAU	Aspartic acid	GGU	Glycine	U			
		GUC		GCC		GAC		GCC		C			
		GUA		GCA		GAA	GGA	A					
		GUG		GCG		GAG	GGG	G					

2.2.4.2 Codons and Building Proteins

- The RNA message (mRNA) is read in three-letter “words” called **codons**. According to the instructions in the RNA message, amino acids are added sequentially to create a protein chain as the message is read. · Each codon signals for a certain amino acid to be added to the protein chain. In the example shown the codon ‘**GUG**’ signals for the amino acid valine. Therefore, an RNA message reading ‘**GUGGAGUUU**’ would code for a protein chain of **valine, glutamic acid, and phenylalanine**. · In order to translate the RNA message (m-RNA), transfer RNA (tRNA) must bring the appropriate amino acid to the mRNA template. The tRNA’s contain the complementary RNA code (in the case of **valine, CAC**). One by one, amino acids are then added together to form a polypeptide chain.

2.2.5 Translation

Proteins are made from the message on the RNA. However, you should understand that translation is a more complex process than transcription. This would, of course, be expected. After all, the coded messages produced by the German Enigma machine could be copied easily, but required a considerable decoding effort before they could be read with understanding. In a similar sense, DNA replication is simply a complementary base pairing exercise, but the

translation of the four letter (bases) alphabet code of RNA to the twenty letter (amino acids) alphabet of protein literature is far from trivial. Clearly, there could not be a direct one-to-one correlation of bases to amino acids, so the nucleotide letters must form short words or **codons** that define specific amino acids. Many questions pertaining to this genetic code were posed in the late 1950's:

- **How many RNA nucleotide bases designate a specific amino acid?**

If separate groups of nucleotides, called codons, serve this purpose, at least three are needed. There are $4^3 = 64$ different nucleotide triplets, compared with $4^2 = 16$ possible pairs.

- **Are the codons linked separately or do they overlap?**

Sequentially joined triplet codons will result in a nucleotide chain three times longer than the protein it describes. If overlapping codons are used then fewer total nucleotides would be required.

- **If triplet segments of mRNA designate specific amino acids in the protein, how are the codons identified?**

For the sequence ~CUAGGU~ are the codons CUA & GGU or ~C, UAG & GU~ or ~CU, AGG & U~?

- **Are all the codon words the same size?**

In Morse code the most widely used letters are shorter than less common letters. Perhaps nature employs a similar scheme.

Physicists and mathematicians, as well as chemists and microbiologists all contributed to unravelling the genetic code. Although earlier proposals assumed efficient relationships that correlated the nucleotide codons uniquely with the twenty fundamental amino acids, it is now apparent that there is considerable redundancy in the code as it now operates. Furthermore, the code consists exclusively of non-overlapping triplet codons. Clever experiments provided some of the earliest breaks in deciphering the genetic code. Marshall Nirenberg found that RNA from many different organisms could initiate specific protein synthesis when combined with broken E.coli cells (the enzymes remain active). A synthetic polyuridine RNA induced synthesis of poly-phenylalanine, so the UUU codon designated phenylalanine. Likewise an alternating ~CACA~ RNA led to synthesis of a ~His-Thr-His-Thr~ polypeptide. The following table presents the present day interpretation of the genetic code. Note that this is the RNA alphabet, and an equivalent DNA codon table would have all the U nucleotides replaced by T. Methionine and tryptophan are uniquely represented by a single codon. At the other extreme, leucine is represented by eight codons. The average redundancy for the twenty amino acids is about three. Also, there are three **stop codons** that terminate polypeptide synthesis.

2.2.5 RNA and Protein Synthesis

The genetic information stored in DNA molecules is used as a blueprint for making proteins. Why proteins? Because these macromolecules have diverse primary, secondary and tertiary structures that equip them to carry out the numerous functions necessary to maintain a living organism. As noted in the [protein chapter](#), these functions include:

- Structural integrity (hair, horn, eye lenses etc.).
- Molecular recognition and signaling (antibodies and hormones).
- Catalysis of reactions (enzymes)..
- Molecular transport (hemoglobin transports oxygen).
- Movement (pumps and motors).

The critical importance of proteins in life processes is demonstrated by numerous genetic diseases, in which small modifications in primary structure produce debilitating and often disastrous consequences. Such genetic diseases include Tay-Sachs, phenylketonuria (PKU), sickle cell anemia, achondroplasia, and Parkinson disease. The unavoidable conclusion is that proteins are of central importance in living cells, and that proteins must therefore be continuously prepared with high structural fidelity by appropriate cellular chemistry.

Early geneticists identified **genes** as hereditary units that determined the appearance and / or function of an organism (i.e. its phenotype). We now define genes as sequences of DNA that occupy specific locations on a chromosome. The original proposal that each gene controlled the formation of a single enzyme has since been modified as: **one gene = one polypeptide**. The intriguing question of how the information encoded in DNA is converted to the actual construction of a specific polypeptide has been the subject of numerous studies, which have created the modern field of **Molecular Biology**.

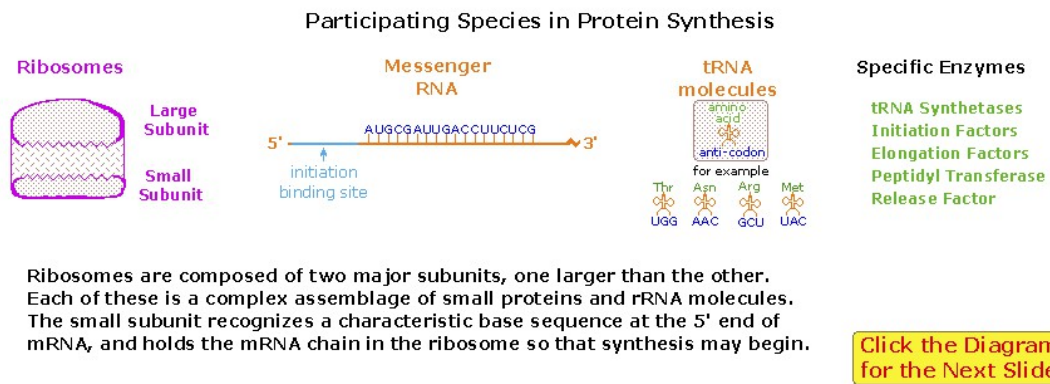


2.3 Regulation of gene expression

The translation process is fundamentally straightforward. The mRNA strand bearing the transcribed code for synthesis of a protein interacts with relatively

small RNA molecules (about 70-nucleotides) to which individual amino acids have been attached by an ester bond at the 3'-end. These **transfer RNA's** (tRNA) have distinctive three-dimensional structures consisting of loops of single-stranded RNA connected by double stranded segments. This cloverleaf secondary structure is further wrapped into an "L-shaped" assembly, having the amino acid at the end of one arm, and a characteristic **anti-codon** region at the other end. The anti-codon consists of a nucleotide triplet that is the complement of the amino acid's codon(s). Models of two such tRNA molecules are shown to the right. When read from the top to the bottom, the anti-codons depicted here should complement a codon in the previous table.

A cell's protein synthesis takes place in organelles called **ribosomes**. Ribosomes are complex structures made up of two distinct and separable subunits (one about twice the size of the other). Each subunit is composed of one or two RNA molecules (60-70%) associated with 20 to 40 small proteins (30-40%). The ribosome accepts a mRNA molecule, binding initially to a characteristic nucleotide sequence at the 5'-end (colored light blue in the following diagram). This unique binding assures that polypeptide synthesis starts at the right codon. A tRNA molecule with the appropriate anti-codon then attaches at the starting point and this is followed by a series of adjacent tRNA attachments, peptide bond formation and shifts of the ribosome along the mRNA chain to expose new codons to the ribosomal chemistry. The following diagram is designed as a slide show illustrating these steps. The outcome is synthesis of a polypeptide chain corresponding to the mRNA blueprint. A "stop codon" at a designated position on the mRNA terminates the synthesis by introduction of a "Release Factor".

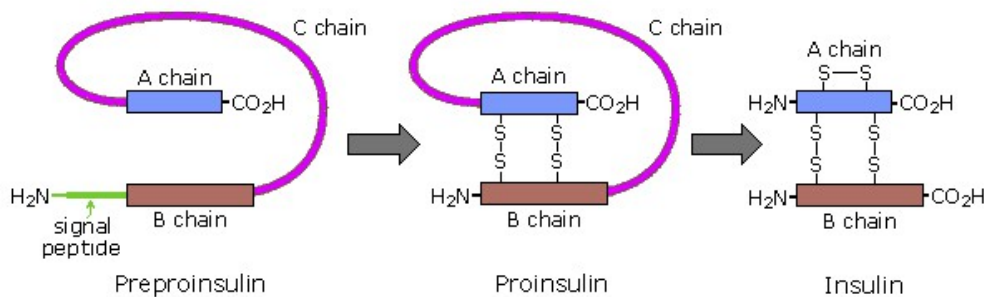


2.3.1. Post-translational Modification

Once a peptide or protein has been synthesized and released from the ribosome it often undergoes further chemical transformation. This **post-translational modification** may involve the attachment of other moieties such as acyl groups, alkyl groups, phosphates, sulfates, lipids and carbohydrates. Functional changes such as dehydration, amidation, hydrolysis and oxidation (e.g. disulfide bond

formation) are also common. In this manner the limited array of twenty amino acids designated by the codons may be expanded in a variety of ways to enable proper functioning of the resulting protein. Since these post-translational reactions are generally catalyzed by enzymes, it may be said: "*Virtually every molecule in a cell is made by the ribosome or by enzymes made by the ribosome.*"

Amino acids may be enzymatically removed from the amino end of the protein. Because the "start" codon on mRNA codes for the amino acid methionine, this amino acid is usually removed from the resulting protein during post-translational modification. Peptide chains may also be cut in the middle to form shorter strands. Thus, [insulin](#) is initially synthesized as a 105 residue preprotein. The 24-amino acid signal peptide is removed, yielding a proinsulin peptide. This folds and forms disulfide bonds between cysteines 7 and 67 and between 19 and 80. Such dimeric cysteines, joined by a disulfide bond, are named **cystine**. A protease then cleaves the peptide at arg31 and arg60, with loss of the 32-60 sequence (chain C). Removal of arg31 yields mature insulin, with the A and B chains held together by disulfide bonds and a third cystine moiety in chain A. The following cartoon illustrates this chain of events.



That every organism is the product of the interaction of its genetic material and its environment is one of the most fundamental of all biological principles. The genetic material (DNA) determines the potentialities for development. The environment determines the degree and kind of the genetic potential that is actually manifested. Or, to put it a slightly different way, the form and function of plants are dependent on their genetic makeup and the ability of the genes express themselves through regulation of protein synthesis. The environment influences the growth, differentiation, and metabolism of a plant by controlling, through effects on both protein synthesis and enzyme activity, expression of its genetic makeup.

Since the impressive, Nobel prize-winning research of G. W. Beadle and E. L. Tatum in the early 1950s on the mechanism of gene action in *Neurospora* that culminated in the one- gene-one enzyme theory, we have understood that the fundamental action of a gene is to control the synthesis of an enzyme or

structural protein. The structure of each polypeptide chain is governed by an independent gene. In the case of several enzymes, known to be made up of one or more polypeptide chains, it is apparent that the structure of each polypeptide chain is governed by a separate independent gene. Furthermore, we now know that there are functionally different types of genes, and that much of the DNA of a eukaryotic organism acts to control the ultimate expression of the structural genes, which contain the genetic code for specific proteins.

To summarize: a gene is a stretch of DNA that contains a pattern for the amino acid sequence of a protein. In order to actually make this protein, the relevant DNA segment is first copied into messenger-RNA. The cell then synthesizes the protein, using the mRNA as a template.

Sizes of Genomic DNA for various Species in kbp

E. Coli	4,640,000bp
Yeast	12,100,000bp
Fruit Fly	140,000,000bp
Human	3,000,000,000bp
Pea	4,800,000,000bp

2.4 DNA Errors and repairs

This unit explains the ways in which DNA can repair any replication errors. It also looks at some of the causes of DNA damage and what failure of the repair mechanism can lead to. One of the benefits of the double stranded DNA structure is that it lends itself to repair, when structural damage or replication errors occur. Several kinds of chemical change may cause damage to DNA: These include:

- Chemical mutagens
- Radiation
- Free radicals

UV-B light causes crosslinking between adjacent cytosine and thymine bases creating pyrimidine dimers. This is called direct DNA damage. UV-A light creates mostly free radicals. The damage caused by free radicals is called indirect DNA damage. Ionizing radiation such as that created by radioactive decay or in cosmic rays causes breaks in DNA strands. Thermal disruption at elevated temperature increases the rate of depurination (loss of purine bases from the DNA backbone) and single strand breaks. For example, hydrolytic depurination is seen in the thermophilic bacteria, which grow in hot springs at 85-250 °C.[6] The

rate of depurination (300) purine residues per genome per generation) is too high in these species to be repaired by normal repair machinery, hence a possibility of an adaptive response cannot be ruled out. Industrial chemicals such as vinyl chloride and hydrogen peroxide, and environmental chemicals such as polycyclic hydrocarbons found in smoke, soot and tar create a huge diversity of DNA adducts- ethenobases, oxidized bases, alkylated phosphotriesters and Crosslinking of DNA just to name a few. The natural ageing process and respiration also causes DNA damage at the rate of around 10000 lesions/cell/day. The main types of DNA damage that occurs are: base loss and base modification.

All these transformations disrupt base pairing at the site of the change, and this produces a structural deformation in the double helix.. Inspection-repair enzymes detect such deformations, and use the undamaged nucleotide at that site as a template for replacing the damaged unit. These repairs reduce errors in DNA structure from about one in ten million to one per trillion.

2.3.1 Strand Directed Mismatch Repair System

DNA mismatch repair is a system which recognises and repairs erroneous insertion, deletion and mis-incorporation of bases that can arise during DNA replication. It also repairs some forms of DNA damage. Mismatch repair is strand-specific. During DNA synthesis the newly synthesised (daughter) strand will often include errors. In order to carry out the repairs, the mismatch repair machinery distinguishes the newly synthesised strand from the template (parental).

- The mismatch repair system carries out the following functions:
- Removes replication errors which are not recognised by the replication machine
- Detects distortions in the DNA helix
 - Distinguishes the newly replicated strand from the parental strand by means of methylation of A residues in GATC in bacteria
- Methylation occurs shortly after replication occur
- Reduces error rate 100x
- 3 step process: recognition of mismatch; excision of segment of DNA containing mismatch;
resynthesis of excised fragment

In mammals the newly synthesised strand is preferentially nicked and can be distinguished in this manner from the parental strand. If there is a defective copy of the mismatch repair gene, then a predisposition to

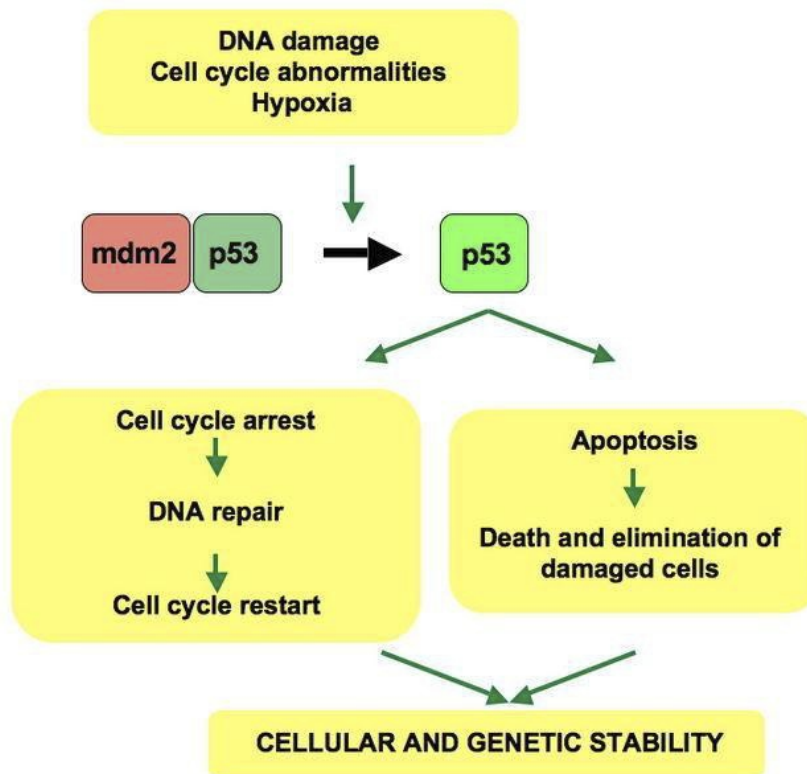
cancer is the end result.

Despite the 1000's of alterations that occur in our DNA each day, very few are actually retained as mutations and this is due to highly efficient DNA repair mechanisms. This is a very important mechanism, and this is highlighted by the high number of genes that are devoted to DNA repair. Also, if there is an inactivation or loss of function of the DNA repair genes, then this results in increased mutation rates. Defects in the DNA repair mechanisms are associated with several disease states as can be seen in the following table:

Disorder	Frequency	Defect	Hereditary/non
Fanconi's anaemia	1/22,000 in some populations	Deficient excision repair	Non - hereditary
Hereditary nonpolyposis colon cancer	1/200	Deficient mismatch repair	Hereditary
Werner's syndrome	3/1,000,000	Deficient helicase	Non-hereditary
Xeroderma pigmentosum	1/250,000	Deficient excision repair	Hereditary

2.4.1 Failure of DNA repair

When DNA repair fails, fewer mutations are corrected and this leads to an increase in the number of mutations in the genome. In most cases, the protein p53 monitors the repair of damaged DNA, however, if the damage is too severe, then p53 promotes programmed cell death (apoptosis). However, mutations in genes which encode the DNA repair proteins can be inherited and this leads to an overall increase in the number of mutations as errors or damage to the DNA is no longer repaired efficiently.



Source: http://en.wikipedia.org/wiki/File:P53_pathways.jpg

In a normal cell p53 is inactivated by its negative regulator, mdm2. Upon DNA damage or other stress, various pathways will lead to the dissociation of the p53 and mdm2 complex. Once activated, p53 will either induce a cell cycle arrest to allow repair and survival of the cell or apoptosis to discard the damage cell. How p53 makes this choice is currently unknown.

2.5 Mutation: Types, causes and effects

This section explains the mechanisms of DNA replication and also the ways in which DNA can repair any replication errors. It also looks at some of the causes of DNA damage and what failure of the repair mechanism can lead to. A gene mutation is a permanent alteration in the DNA sequence that makes up a gene, such that the sequence differs from what is found in most people. Mutations range in size; they can affect anywhere from a single DNA building block (base pair) to a large segment of a chromosome that includes multiple genes.

Gene mutations can be classified in two major ways:

- Hereditary mutations are inherited from a parent and are present throughout a person's life in virtually every cell in the body. These mutations are also called germline mutations because they are present in the parent's egg or sperm cells, which are also called germ cells. When an egg and a sperm cell unite, the resulting fertilized egg cell receives DNA from both parents. If this DNA has a mutation, the child that grows from the fertilized egg will have the mutation in each of his or her cells.
- Acquired (or somatic) mutations occur at some time during a person's life and are present only in certain cells, not in every cell in the body. These changes can be caused by environmental factors such as ultraviolet radiation from the sun, or can occur if a mistake is made as DNA copies itself during cell division. Acquired mutations in somatic cells (cells other than sperm and egg cells) cannot be passed on to the next generation.

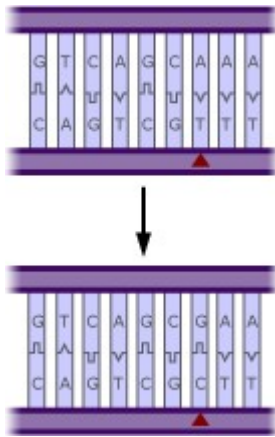
Genetic changes that are described as de novo (new) mutations can be either hereditary or somatic. In some cases, the mutation occurs in a person's egg or sperm cell but is not present in any of the person's other cells. In other cases, the mutation occurs in the fertilized egg shortly after the egg and sperm cells unite. (It is often impossible to tell exactly when a de novo mutation happened.) As the fertilized egg divides, each resulting cell in the growing embryo will have the mutation. De novo mutations may explain genetic disorders in which an affected child has a mutation in every cell in the body but the parents do not, and there is no family history of the disorder.

Somatic mutations that happen in a single cell early in embryonic development can lead to a situation called mosaicism. These genetic changes are not present in a parent's egg or sperm cells, or in the fertilized egg, but happen a bit later when the embryo includes several cells. As all the cells divide during growth and development, cells that arise from the cell with the altered gene will have the mutation, while other cells will not. Depending on the mutation and how many cells are affected, mosaicism may or may not cause health problems.

Most disease-causing gene mutations are uncommon in the general population. However, other genetic changes occur more frequently. Genetic alterations that occur in more than 1 percent of the population are called polymorphisms. They are common enough to be considered a normal variation in the DNA. Polymorphisms are responsible for many of the normal differences between people such as eye color, hair color, and blood type. Although many polymorphisms have no negative effects on a person's health, some of these variations may influence the risk of developing certain disorders.

A mutation is a change in [DNA](#), the hereditary material of life. An organism's DNA affects how it looks, how it behaves, and its physiology. So a change in an organism's DNA can cause changes in all aspects of its life. Mutations are

essential to evolution; they are the raw material of [genetic variation](#). Without mutation, evolution could not occur. In figure below allows you to visualize a typical mutation.



DNA: The molecular basis of mutations

Since mutations are simply changes in DNA, in order to understand how mutations work, you need to understand how DNA does its job. Your DNA contains a set of instructions for "building" a human. These instructions are inscribed in the structure of the DNA molecule through a genetic code. It works like this:

DNA is made of a long sequence of smaller units strung together. There are four basic types of unit: A, T, G, and C. These letters represent the type of [base](#) each unit carries: adenine, thymine, guanine, and cytosine.

The sequence of these bases encodes instructions. Some parts of your DNA are control centers for turning [genes](#) on and off, some parts have no function, and some parts have a function that we don't understand yet. Other parts of your DNA are genes that carry the instructions for making [proteins](#) — which are long chains of [amino acids](#). These proteins help build an organism.

Protein-coding DNA can be divided into [codons](#) — sets of three bases that specify an amino acid or signal the end of the protein. Codons are identified by the bases that make them up — in the example at right, GCA, for guanine, cytosine, and adenine. The cellular machinery uses these instructions to assemble a string of corresponding amino acids (one amino acid for each three bases) that form a protein. The amino acid that corresponds to "GCA" is called alanine; there are twenty different amino acids synthesized this way in humans. "Stop" codons signify the end of the newly built protein. After the protein is

built based on the sequence of bases in the gene, the completed protein is released to do its job in the cell.

2.5.1 Types of mutations

There are many different ways that DNA can be changed, resulting in different types of mutation. Here is a quick summary of a few of these:

CTGGAG
CTGGGG

Substitution

A substitution is a mutation that exchanges one base for another (i.e., a change in a single "chemical letter" such as switching an A to a G). Such a substitution could:

1. change a codon to one that encodes a different amino acid and cause a small change in the protein produced. For example, [sickle cell anemia](#) is caused by a substitution in the beta-hemoglobin gene, which alters a single amino acid in the protein produced.
2. change a codon to one that encodes the same amino acid and causes no change in the protein produced. These are called silent mutations.
3. change an amino-acid-coding codon to a single "stop" codon and cause an incomplete protein. This can have serious effects since the incomplete protein probably won't function.

CTGGAG
CTGGTGGAG

Insertion

Insertions are mutations in which extra base pairs are inserted into a new place in the DNA.

~~CTGGAG~~
CTAG

Deletion

Deletions are mutations in which a section of DNA is lost, or deleted.

~~The fat cat sat~~
hef atc ats at

Frameshift

Since protein-coding DNA is divided into codons three bases long, insertions and deletions can alter a gene so that its message is no longer correctly parsed. These changes are called frameshifts.

For example, consider the sentence, "The fat cat sat." Each word represents a codon. If we delete the first letter and parse the sentence in the same way, it doesn't make sense.

In frameshifts, a similar error occurs at the DNA level, causing the codons to be parsed incorrectly. This usually generates truncated proteins that are as useless as "hef atc ats at" is uninformative.

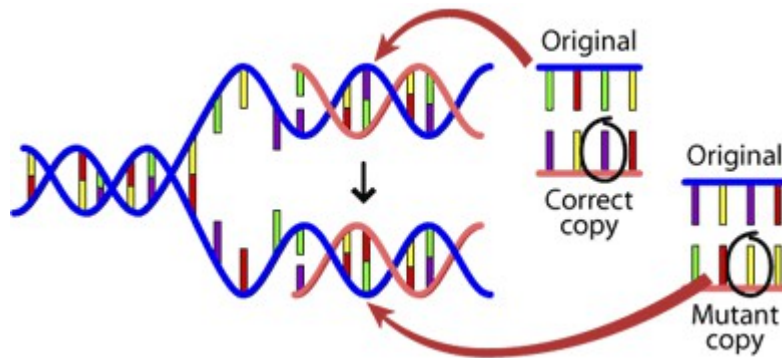
There are other types of mutations as well, but this short list should give you an idea of the possibilities.

2,5.2 The causes of mutations

Mutations happen for several reasons.

1. DNA fails to copy accurately

Most of the mutations that we think matter to evolution are "naturally occurring." For example, when a cell divides, it makes a copy of its DNA — and sometimes the copy is not quite perfect. That small difference from the original DNA sequence is a mutation.



2. External influences can create mutations

Mutations can also be caused by exposure to specific chemicals or radiation. These agents cause the DNA to break down. This is not necessarily unnatural — even in the most isolated and pristine environments, DNA breaks down. Nevertheless, when the cell repairs the DNA, it might not do a perfect job of the repair. So the cell would end up with DNA slightly different than the original DNA and hence, a mutation.

2.5.3 The effects of mutations

Since all cells in our body contain DNA, there are lots of places for mutations to occur; however, some mutations cannot be passed on to offspring and do not matter for evolution. [Somatic mutations](#) occur in non-reproductive cells and won't be passed onto offspring. For example, the golden color on half of this Red Delicious apple was caused by a somatic mutation. Its seeds will not carry the mutation.

The only mutations that matter to large-scale evolution are those that can be passed on to offspring. These occur in reproductive cells like eggs and sperm and are called [germ line mutations](#).

Effects of germ line mutations
A single germ line mutation can have a range of effects:

1. **No change occurs in phenotype.**
Some mutations don't have any noticeable effect on the phenotype of an organism. This can happen in many situations: perhaps the mutation occurs in a stretch of DNA with no function, or perhaps the mutation occurs in a protein-coding region, but ends up not affecting the [amino acid](#) sequence of the [protein](#).
2. **Small change occurs in phenotype.**
— A single mutation caused this cat's ears to curl backwards slightly.
3. **Big change occurs in phenotype.**
Some really important phenotypic changes, like DDT resistance in insects are sometimes caused by single mutations. A single mutation can also have strong negative effects for the organism. Mutations that cause the death of an organism are called lethals — and it doesn't get more negative than that.

Little mutations with big effects: Mutations to control genes. Mutations are often the victims of bad press — unfairly stereotyped as unimportant or as a cause of genetic disease. While many mutations do indeed have small or negative effects, another sort of mutation gets less airtime. Mutations to control genes can have major (and sometimes positive) effects.

Some regions of DNA control other genes, determining when and where other genes are turned "on". Mutations in these parts of the genome can substantially change the way the organism is built. The difference between a mutation to a

control gene and a mutation to a less powerful gene is a bit like the difference between whispering an instruction to the trumpet player in an orchestra versus whispering it to the orchestra's conductor. The impact of changing the conductor's behavior is much bigger and more coordinated than changing the behavior of an individual orchestra member. Similarly, a mutation in a gene "conductor" can cause a cascade of effects in the behavior of genes under its control.

2.6. Laboratory Exercise on extraction of DNA and RNA from Animal Tissues:

2.6.1 Introduction

The Purpose of this unit is to introduce the main steps in the DNA extraction protocol and review the chemistry involved in each step. In view of the role of DNA in inheritance of traits, it is important that you understand the steps involved in DNA Extraction. This will enable you to obtain DNA in a relatively purified form which can be used for further investigations, like PCR, sequencing, etc

The Basic Protocol consist of two parts: A technique to lyse the cells gently and solubilize the DNA, and then enzymatic or chemical methods to remove contaminating proteins, RNA, or macromolecules

- In plants, the nucleus is protected within a nuclear membrane which is surrounded by a cell membrane and a cell wall. Four steps are used to remove and purify the DNA from the rest of the cell.

1. Lysis
2. Precipitation
3. Wash
4. Resuspension

The key steps involved in DNA extraction can be summarized as follows:

1. **Lysis:** grind in mortar/pestle and use detergent
2. **Precipitation Part I:** addition of salts to interrupt hydrogen bonding between water and phosphates on the DNA In the **Part II:** addition of ethanol to pull DNA out of solution
3. **Wash and resuspend:** DNA is washed in ethanol, dried, and resuspended in H₂O or TE buffer

The above listed steps are given in more details in the following section;

1. LYSIS:

In DNA extraction this step commonly refers to the breaking of the cellular membranes most importantly, the plasma and nuclear membranes. The cell

membrane is disrupted by mechanical force (for example, grinding). This followed by the addition of a detergent which breaks down the cell membranes. Detergents are able to disrupt membranes due to the amphipathic (having both hydrophilic and hydrophobic regions) nature of both cellular membranes and detergent molecules. The detergent molecules are able to pull apart the membranes. The end result of LYSIS is that the contents of the cells are distributed in solution.

2. **PRECIPITATION:**

This a series of steps where DNA is separated from the rest of the cellular components. In a research lab, the first part of precipitation uses phenol/chloroform to remove the proteins from the DNA. The phenol denatures proteins and dissolves denatured proteins. Whereas chloroform is also a protein denaturant. The second part of research lab DNA precipitation is the addition of salts. The salts interrupt the hydrogen bonds between the water and DNA molecules. The DNA is then precipitated from the protein in a subsequent step with isopropanol or ethanol. In the presence of cations, ethanol induces a structural change in DNA molecules that causes them to aggregate and precipitate out of solution. The DNA is pelleted by spinning with a centrifuge and the supernatant removed

3. **Washing and Resuspension**

The precipitated DNA is laden with acetate salts. It is “washed” with a 70% ethanol solution to remove salts and other water soluble impurities but not resuspend the DNA. The clean DNA is now resuspended in a buffer to ensure stability and long term storage. The most commonly used buffer for resuspension is called 1xTE

An Overview of the procedure for DNA Extraction is presented in figure 2. 1.

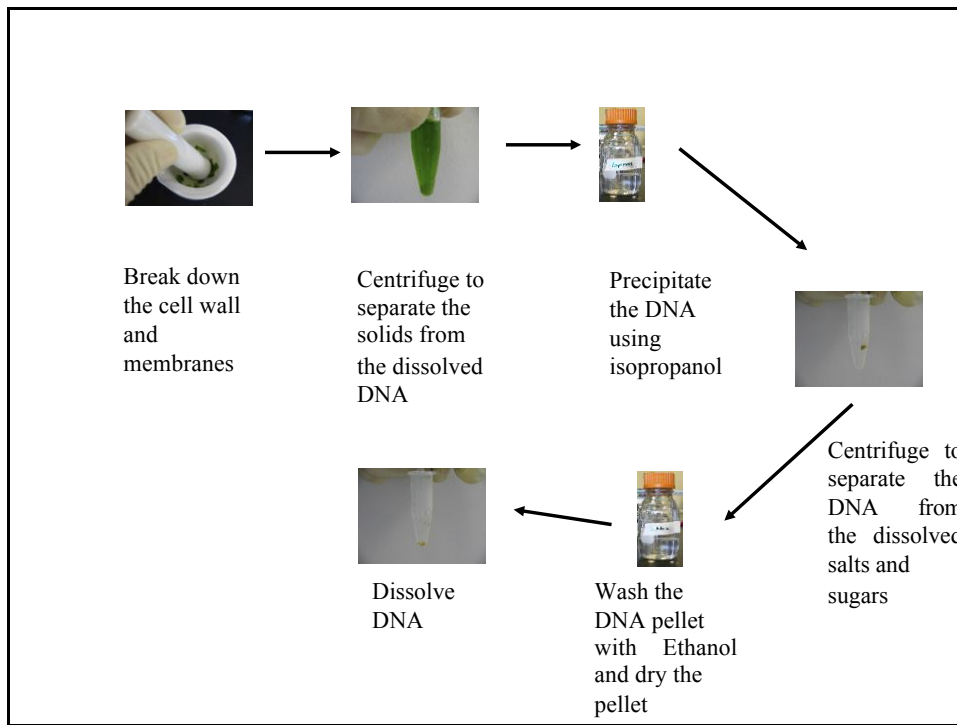
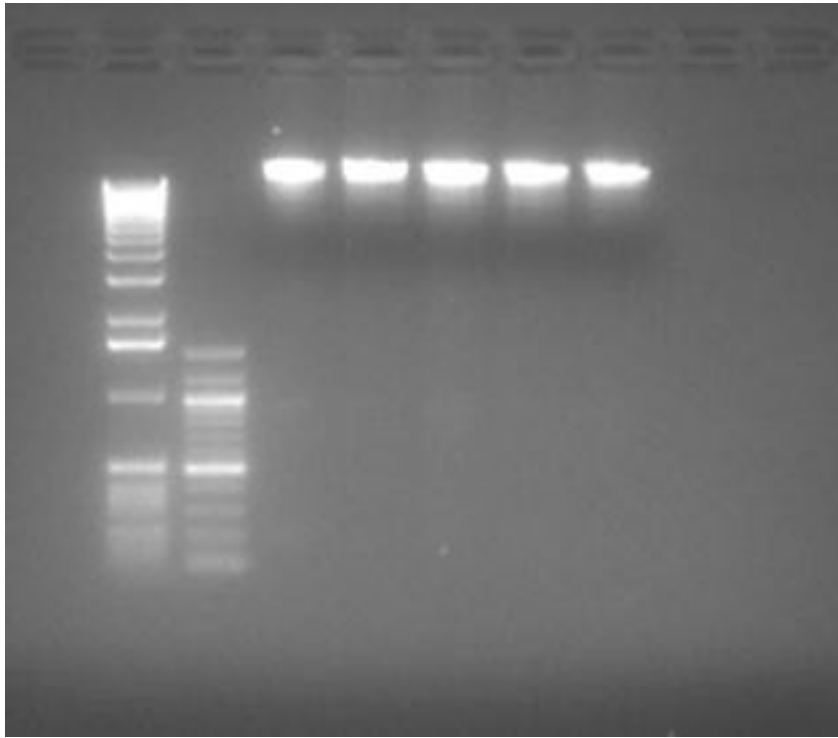


Figure 2.1: A flow chart showing the steps involved in DNA Extraction.

2.6.2 Checking the Quality of your DNA

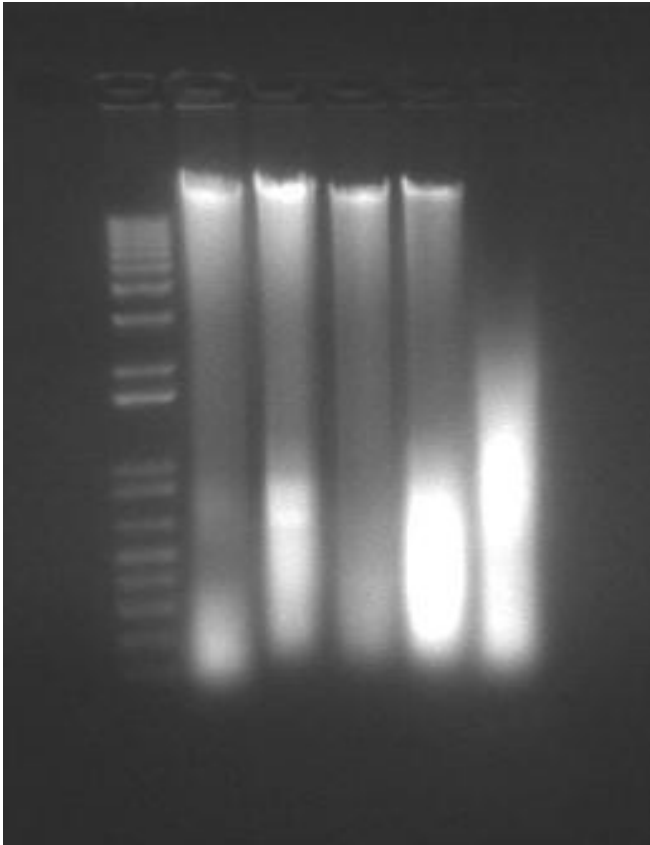
The product of your DNA extraction will be used in subsequent experiments. You could recall that in the opening of this unit it was indicated that the purpose of DNA extraction is for further testing, using PCR, sequencing, finger printing etc. You should note that poor quality DNA will not perform well in PCR. Therefore there is need to assess the quality of the DNA you extracted. This can be carried out using the following simple protocol: Mix 10 μL of DNA with 10 μL of loading buffer, Load this mixture into a 1% agarose gel and analyze results. Figure 2.2 is an agarose gel that has 5 genomic DNA samples from various animals. Note that the DNA runs at a very high molecular weight and as a collar, thick band. This is typical of DNA extracted in a research lab under optimal conditions. This DNA will be excellent for PCR and further experiments.



1 kbp and 100 bp ladders Genomic DNA of 5 species of animals

Figure 2.2: Showing the results of 1% Agrose gel electropherisis of Pure DNA extracted under ideal research conditions.

There are times when impure DNA can be obtained when the steps involved in the extraction are not followed optimally (Fig. 2.3). For instance using the above by missing phenol/chloroform step, the genomic DNA extracted will look different than the optimized DNA extraction on the previous Fig. 2.2). Please observe that the DNA has sheared (particularly for wheat) - broken up into numerous fragments and is not a clean single band at the top - these are the mid-ranged sized fragments (1000-10,000bp size range). The bright bands at the 100 - 1000 bp range are RNA, which also gets extracted using this protocol. This is expected. Even though this genomic DNA preparation is not perfect, it is suitable for use as a PCR template. Sample (A): This sample is fine; Sample (B): This sample is fine; Sample (C) : This sample is fine; Sample (D) : This sample is fine; Sample (E): This sample has severe degradation, can work for PCR but should be re-extracted



The genomic fragments run at ~12kbp because they are sheared during extraction

Figure 2.3: Showing the results of 1% agarose gel electrophoresis of impure DNA extracted under non ideal conditions

References:

- 1) Clancy,S. (2008). DNA damage & repair: mechanisms for maintaining DNA integrity.*Nature Education* **1(1)**:103
- 2) DNA Replication and Repair. BS 2009. University of Leicester 18th February 2010. 13p
- 3) **DNA & Mutations**
- 4) RILEY , D. E. 2005. DNA Testing: An Introduction For Non-Scientists: An Illustrated Explanation. Revised Edition Posted April 6, 2005. University of Washington (Copyright 1997-2005)

MODULE THREE

Principles of Recombinant DNA Technology

3.1.1 Recombinant DNA Defined

What Is rDNA? That's a very good question! rDNA stands for recombinant DNA. Before we get to the "r" part, you need to recall your understanding of DNA from the earlier modules. DNA is the keeper of the all the information needed to recreate an organism. All DNA is made up of a base consisting of sugar, phosphate and one nitrogen base. There are four nitrogen bases, adenine (A), thymine (T), guanine (G), and cytosine (C). The nitrogen bases are found in pairs, with A & T and G & C paired together. The sequence of the nitrogen bases can be arranged in an infinite ways, and their structure is known as the famous "double helix" which is shown in the image below. The sugar used in DNA is deoxyribose. The four nitrogen bases are the same for all organisms. The sequence and number of bases is what creates diversity. DNA does not actually make the organism, it only makes proteins. The DNA is transcribed into mRNA and mRNA is translated into protein, and the protein then forms the organism. By changing the DNA sequence, the way in which the protein is formed changes. This leads to either a different protein, or an inactive protein.

Now that we know what DNA is, this is where the recombinant comes in. Recombinant DNA is the general name for taking a piece of one DNA, and combining it with another strand of DNA. Thus, the name recombinant! Recombinant DNA is also sometimes referred to as "chimera." By combining two or more different strands of DNA, scientists are able to create a new strand of DNA. The most common recombinant process involves combining the DNA of two different organisms.

3.1.2 How rDNA Functions

Recombinant DNA works when the host cell expresses protein from the recombinant genes. A significant amount of recombinant protein will not be produced by the host unless expression factors are added. Protein expression depends upon the gene being surrounded by a collection of signals which provide instructions for the transcription and translation of the gene by the cell.

These signals include the promoter, the ribosome binding site, and the terminator. Expression vectors, in which the foreign DNA is inserted, contain these signals. Signals are species specific. In the case of E. Coli, these signals must be E. Coli signals as E. Coli is unlikely to understand the signals of human promoters and terminators. Problems are encountered if the gene contains introns or contains signals which act as terminators to a bacterial host. This results in premature termination, and the recombinant protein may not be processed correctly, be folded correctly, or may even be degraded.

Production of recombinant proteins in eukaryotic systems generally takes place in yeast and filamentous fungi. The use of animal cells is difficult due to the fact that many need a solid support surface, unlike bacteria, and have complex growth needs. However, some proteins are too complex to be produced in bacterium, so eukaryotic cells must be used.

3.1.3 Importance of rDNA

Recombinant DNA has been gaining in importance over the recent years, and recombinant DNA will only become more important in the 21st century as genetic diseases become more prevalent and agricultural area is reduced. Below are some of the areas where Recombinant DNA will have an impact.

1. Better Crops (drought & heat resistance)
2. Recombinant Vaccines (ie. Hepatitis B)
3. Prevention and cure of sickle cell anemia
4. Prevention and cure of cystic fibrosis
5. Production of clotting factors
6. Production of insulin
7. Production of recombinant pharmaceuticals
8. Plants that produce their own insecticides
9. Germ line and somatic gene therapy

3.2 Recombination and Cloning of Genes

How do we manipulate these natural processes for biotechnology; for instance, to make a bacterium that produces large quantities of insulin? One approach would be to cut the appropriate gene from human DNA and paste, or splice, it into a [vector](#) such as a plasmid or phage DNA. Our "scissors" are the class of enzymes called [restriction endonucleases](#)

3.2.1

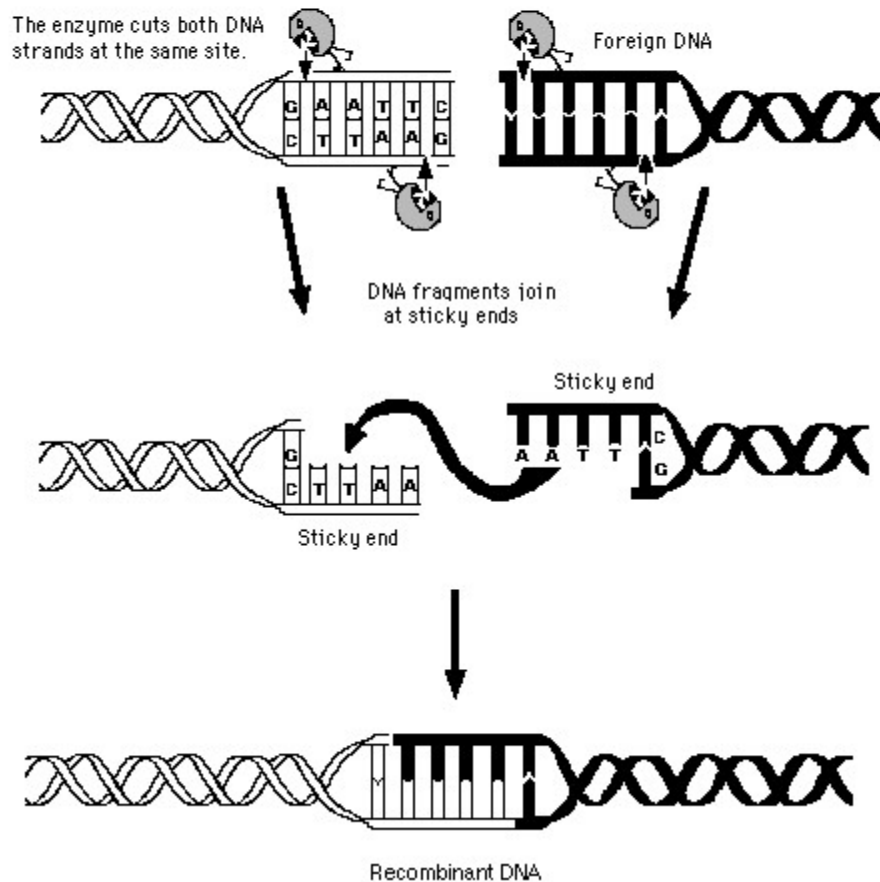
Restriction

Endonucleases

An "endonuclease" is an enzyme that cuts duplex DNA in the middle, not at an end (for exonuclease). Different species of bacteria have evolved different restriction endonucleases, each to cut foreign DNA that gets into their cells by mistake. To be cut, the DNA has to lack their own pattern of protective methylation. There are well over a hundred restriction enzymes, each cutting in a very precise way a specific base sequence of the DNA molecule.

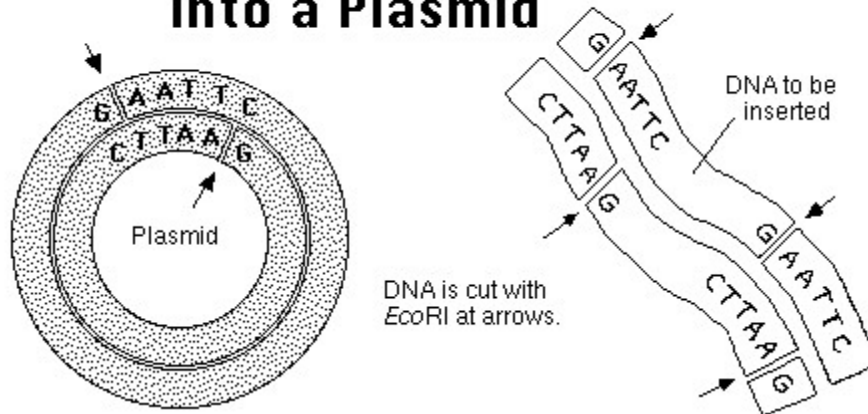
A restriction endonuclease cuts DNA only at a specific site, usually containing 4-6 base pairs. The enzyme has to cut the DNA backbone twice, recognizing the same type of site; therefore, the site "reads" the same way backwards as forwards--a [palindrome](#).

Restriction Enzyme Action of EcoRI

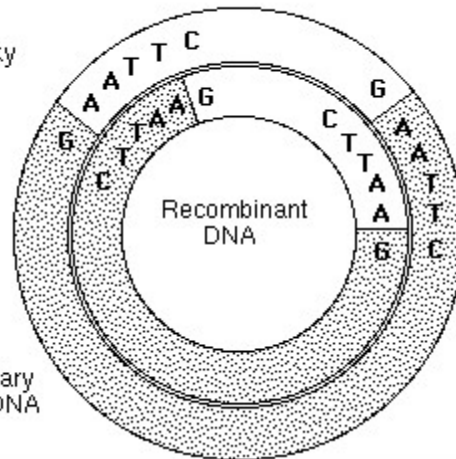


This "sticky ends" from two different DNA molecules can hybridize together; then the nicks are sealed using **ligase**. (**Where does ligase come from? What is its natural function?**) The result is **recombinant DNA**. When this recombinant vector is inserted into *E. coli*, the cell will be able to process the instructions to assemble the amino acids for insulin production. More importantly, the new instructions are passed along to the next generation of *E. coli* cells in the process known as **gene cloning**.

Inserting a DNA Sample into a Plasmid



Resulting DNAs have sticky (complementary) ends.



DNA is spliced by complementary base pairing and sealed with DNA ligase

Transfer and Cloning of the Insulin Gene

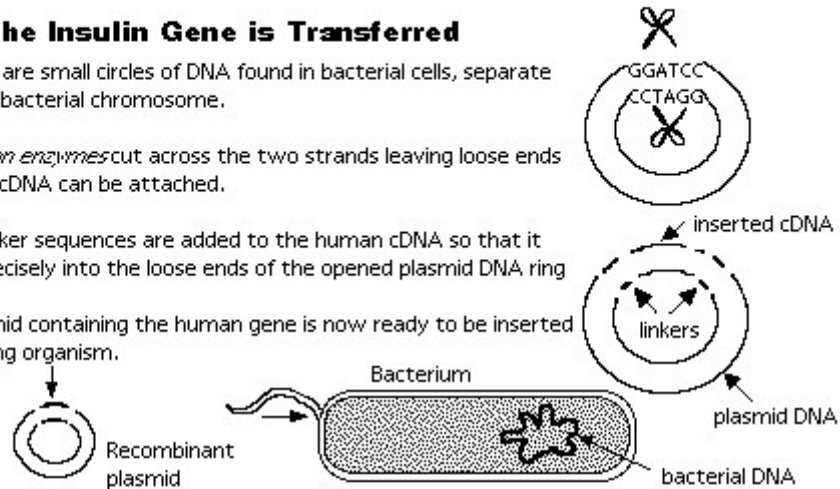
How the Insulin Gene is Transferred

Plasmids are small circles of DNA found in bacterial cells, separate from the bacterial chromosome.

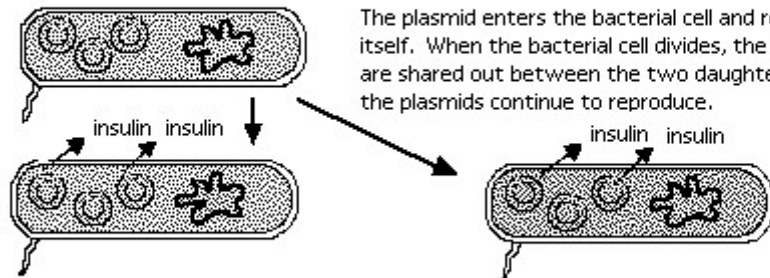
Restriction enzymes cut across the two strands leaving loose ends to which cDNA can be attached.

Special linker sequences are added to the human cDNA so that it will fit precisely into the loose ends of the opened plasmid DNA ring

The plasmid containing the human gene is now ready to be inserted into a living organism.



Cloning the Human Insulin Gene



The plasmid enters the bacterial cell and reproduces itself. When the bacterial cell divides, the plasmids are shared out between the two daughter cells and the plasmids continue to reproduce.

In this way a clone of identical cells is formed and if the human gene incorporated encodes for the hormone insulin, then such a clone can provide a reliable insulin source.

The above is a highly simplified description of recombinant DNA technology.

3.2.2 Cloning genes

In nature, DNA molecules recombine for various functions -- even DNA between different species. But twenty years ago, despite the work of Barbara McClintock and others, the extent of this recombination was not appreciated. DNA was still thought to be the "master molecule," not to be violated by "unnatural" manipulation. When scientists began to manipulate DNA in the test tube, many scientists feared that disastrous monsters would result, with unspecified dangers to people. In 1977 scientists at the Asilomar Conference proposed sweeping regulation on so-called "recombinant DNA," technologies which recombine DNA from different species in the test tube.

- These are called "sticky ends" because they are able to base pair with any DNA molecule containing the complementary sticky end.
- In this case, both DNA preparations have complementary sticky ends and thus can pair with each other when mixed.
- a **DNA ligase** covalently links the two into a molecule of **recombinant DNA**.

To be useful, the recombinant molecule must be replicated many times to provide material for analysis, [sequencing](#), etc. Producing many identical copies of the same recombinant molecule is called **cloning**. Cloning can be done [in vitro](#), by a process called the [polymerase chain reaction \(PCR\)](#). Here, however, we shall examine how cloning is done [in vivo](#).

Cloning in vivo can be done in

- unicellular microbes like [E. coli](#)
- unicellular [eukaryotes](#) like **yeast** and
- in mammalian cells grown in tissue culture.

In every case, the recombinant DNA must be taken up by the cell in a form in which it can be replicated and expressed. This is achieved by incorporating the DNA in a [vector](#). A number of viruses (both bacterial and of mammalian cells) can serve as vectors. But here let us examine an example of cloning using *E. coli* as the host and a **plasmid** as the vector.

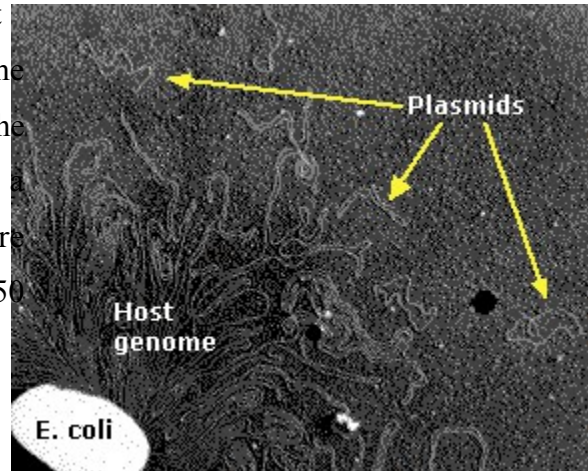
3.2.1 Plasmids

Plasmids are molecules of DNA that are found in bacteria separate from the bacterial chromosome.

They:

- are small (a few thousand base pairs)
- usually carry only one or a few genes
- are circular
- have a single [origin of replication](#)

Plasmids are replicated by the same machinery that replicates the bacterial chromosome. Some plasmids are copied at about the same rate as the chromosome, so a single cell is apt to have only a single copy of the plasmid. Other plasmids are copied at a high rate and a single cell may have 50 or more of them.



Genes on plasmids with high numbers of copies are

usually expressed at high levels. In nature, these genes often encode proteins (e.g., enzymes) that protect the bacterium from one or more [antibiotics](#).

Plasmids enter the bacterial cell with relative ease. This occurs in nature and may account for the rapid spread of antibiotic resistance in hospitals and elsewhere. Plasmids can be deliberately introduced into bacteria in the laboratory [transforming](#) the cell with the incoming genes.

Electron micrograph of an *E. coli* cell ruptured to release its DNA. The tangle is a portion of a single DNA molecule containing over 4.6 million base pairs encoding approximately 4,300 genes. The small circlets are plasmids. (Courtesy of Huntington Potter and David Dressler, Harvard Medical School.)

An Example

(courtesy of David Miklos and Greg Freyer of the Cold Spring Harbor Laboratory, who used these plasmids as the basis of a laboratory introduction to recombinant DNA technology that every serious biology student — high school or college — should experience!)

pAMP

- 4539 base pairs
- a single replication origin
- a gene (**amp^r**) conferring resistance to the antibiotic **ampicillin** (a relative of [penicillin](#))
- a **single** occurrence of the sequence

5' GGATCC 3'
3' CCTAGG 5'

that, as we saw above, is cut by the [restriction enzyme BamHI](#) .

a **single** occurrence of the sequence

5' AAGCTT 3'
3' TTCGAA 5'

that is cut by the restriction enzyme **HindIII**

Treatment of pAMP with a **mixture** of BamHI and HindIII produces:

- a fragment of **3755** base pairs carrying both the **amp^r** gene and the replication origin
- a fragment of **784** base pairs
- both fragments have sticky ends

pKAN

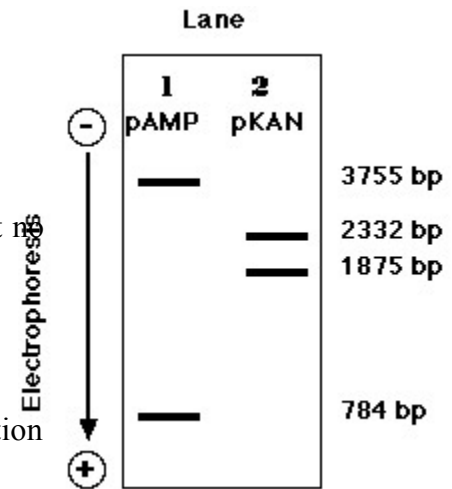
- 4207 base pairs
 - a single replication origin
 - a gene (**kan^r**) conferring resistance to the antibiotic [kanamycin](#) .
- a single site cut by **BamHI**
- a single site cut by **HindIII**

Treatment of pKAN with a **mixture** of BamHI and HindIII produces:

- a fragment of **2332** base pairs
- a fragment of **1875** base pairs with the **kan^r** gene (but not origin of replication)
- both fragments have sticky ends

These fragments can be visualized by subjecting the digestion mixtures to [electrophoresis](#) in an agarose gel. Because of its

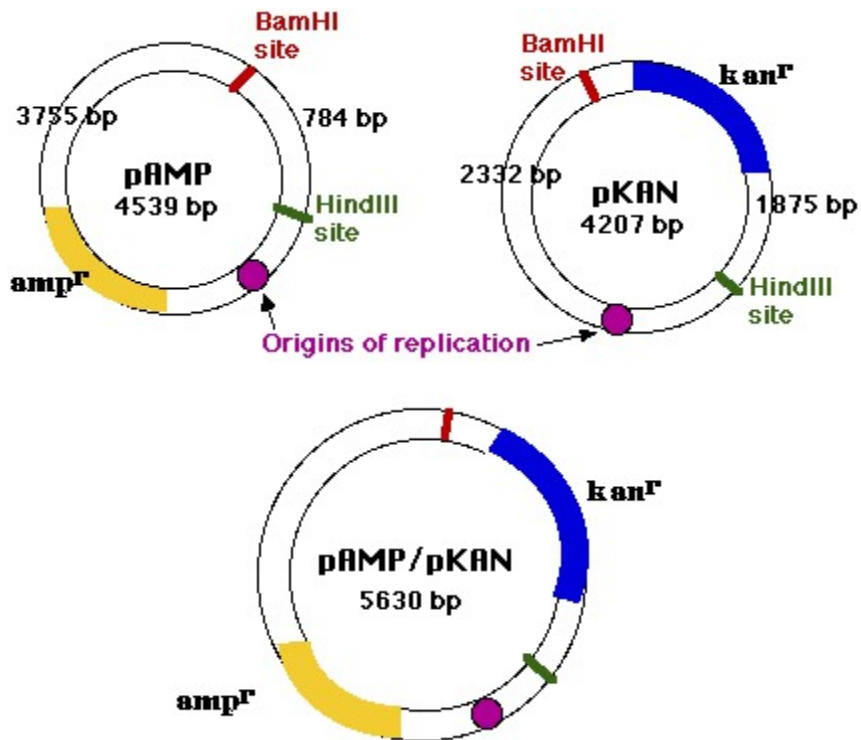
negatively-charged [phosphate groups](#), DNA migrates toward the positive electrode (anode) when a direct current is applied. The smaller the fragment, the farther it migrates in the gel.



3.3.2 Ligation Possibilities

If you remove the two restriction enzymes and provide the conditions for DNA ligase to do its work, the pieces of these plasmids can rejoin (thanks to the complementarity of their sticky ends).

Mixing the pKAN and pAMP fragments provides several (at least 10) possibilities of rejoined molecules. Some of these will not produce functional plasmids (molecules with two or with no replication origin cannot function).



One interesting possibility is the joining of

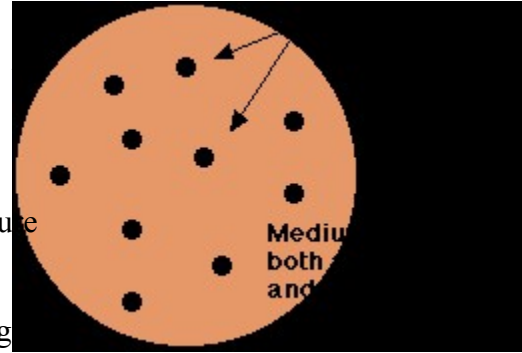
- the 3755-bp pAMP fragment (with **amp^r** and a replication origin) with the
- 1875-bp pKAN fragment (with **kan^r**)

Sealed with **DNA ligase**, these molecules are functioning plasmids that are capable of conferring resistance to **both** ampicillin and kanamycin. They are molecules of **recombinant DNA**.

Because the replication origin, which enables the molecule to function as a plasmid, was contributed by pAMP, pAMP is called the **vector**.

3.3 Transforming *E. coli*

Treatment of *E. coli* with the mixture of religated molecules will produce some colonies that are able to grow in the presence of both ampicillin and kanamycin.

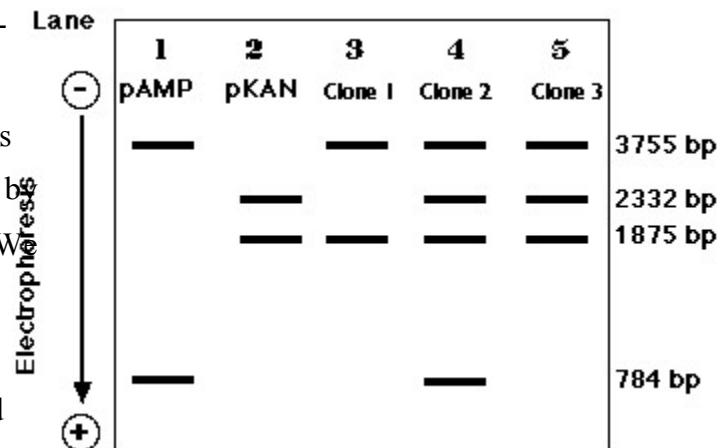


- A suspension of *E. coli* is treated with the mixture of religated DNA molecules.
- The suspension is spread on the surface of agar containing both ampicillin and kanamycin.
- The next day, a few cells — resistant to both antibiotics — will have grown into visible colonies containing billions of transformed cells.
- Each colony represents a **clone** of transformed cells.

However, *E. coli* can be simultaneously transformed by more than one plasmid, so we must demonstrate that the transformed cells have acquired the recombinant plasmid.

Electrophoresis of the DNA from doubly-resistant colonies (clones) tells the story.

- Plasmid DNA from cells that acquired their resistance from a **recombinant plasmid** only show only the **3755-bp** and **1875-bp** bands (**Clone 1**, lane 3).
- **Clone 2** (Lane 4) was simultaneously transformed by religated pAMP and pKAN. (We cannot tell if it took up the recombinant molecule as well.)
- **Clone 3** (Lane 5) was transformed by the recombinant molecule as well as by an intact pKAN.



3.3. 1Cloning other Genes

The recombinant vector described above could itself be a useful tool for cloning other genes. Let us assume that within its **kanamycin resistance gene** (**kan^r**) there is a single occurrence of the sequence

5'

GAATTC

3'

3' CTTAAG 5'

This is cut by the restriction enzyme [EcoRI](#), producing sticky ends.

If we treat any other sample of DNA, e.g., from human cells, with EcoRI, fragments with the same sticky ends will be formed. Mixed with EcoRI-treated plasmid and DNA ligase, a small number of the human molecules will become incorporated into the plasmid which can then be used to transform *E. coli*.

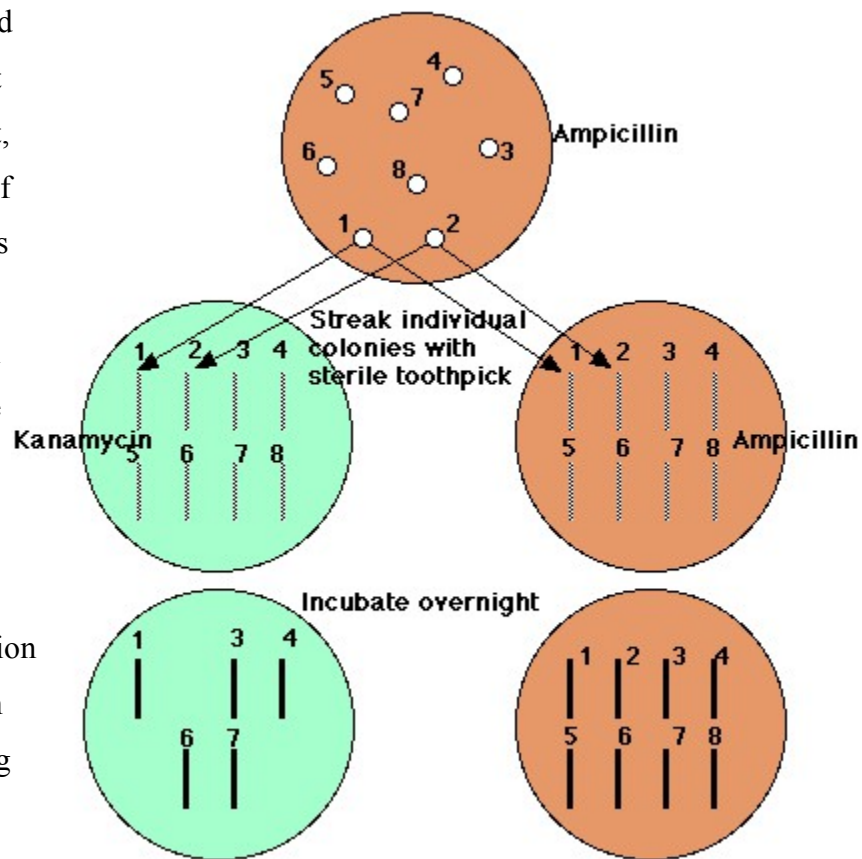
But how to detect those clones of *E. coli* that have been transformed by a plasmid carrying a piece of human DNA?

The key is that the EcoRI site is **within** the **kan^r** gene, so when a piece of human DNA is inserted there, the gene's function is destroyed.

All *E. coli* cells transformed by the vector, whether it carries human DNA or not, can grow in the presence of ampicillin. But *E. coli* cells transformed by a plasmid carrying human DNA will be unable to grow in the presence of kanamycin.

So,

- Spread a suspension of treated *E. coli* on agar containing ampicillin only
- grow overnight



- with a sterile toothpick transfer a small amount of each colony to an identified spot on agar containing kanamycin
- (do the same with another ampicillin plate)
- Incubate overnight

All those clones that continue to grow on ampicillin but fail to grow on kanamycin (here, **clones 2, 5, and 8**) have been transformed with a piece of human DNA.

3.3.2 Some recombinant DNA products being used in human therapy

Using procedures like this, many human genes have been cloned in *E. coli* or in yeast. This has made it possible — for the first time — to produce unlimited amounts of human proteins in vitro. Cultured cells (*E. coli*, yeast, mammalian cells) transformed with a human gene are being used to manufacture more than 100 products for human therapy. Some examples:

1. [insulin](#) for diabetics
2. [factor VIII](#) for males suffering from hemophilia A
3. [factor IX](#) for hemophilia B
4. [human growth hormone \(HGH\)](#)
5. [erythropoietin \(EPO\)](#) for treating anemia
6. several types of [interferons](#)
7. several [interleukins](#)
8. [granulocyte-macrophage colony-stimulating factor \(GM-CSF\)](#) for stimulating the bone marrow after a bone marrow transplant
9. [granulocyte colony-stimulating factor \(G-CSF\)](#) for stimulating [neutrophil](#) production (e.g., after [chemotherapy](#)) and for [mobilizing hematopoietic stem cells](#) from the bone marrow into the blood.
10. [tissue plasminogen activator \(TPA\)](#) for dissolving blood clots
11. [adenosine deaminase \(ADA\)](#) for treating some forms of [severe combined immunodeficiency \(SCID\)](#)
12. [parathyroid hormone](#)
13. many [monoclonal antibodies](#)

14. [hepatitis B surface antigen \(HBsAg\)](#) to vaccinate against the [hepatitis B virus](#)
15. [C1 inhibitor \(C1INH\)](#) used to treat [hereditary angioedema](#)

3.4. Molecular tools and techniques:

3.4.1 Polymerase chain reaction (PCR)

PCR is an abbreviation for "polymerase chain reaction." (POLL'-IM-ER-ACE). This term applies to a wide variety of different DNA tests that differ in reliability and effectiveness. Reliabilities of each kind of PCR test need independent verification. PCR itself doesn't accomplish DNA typing, it only increases the amount of DNA available for typing.

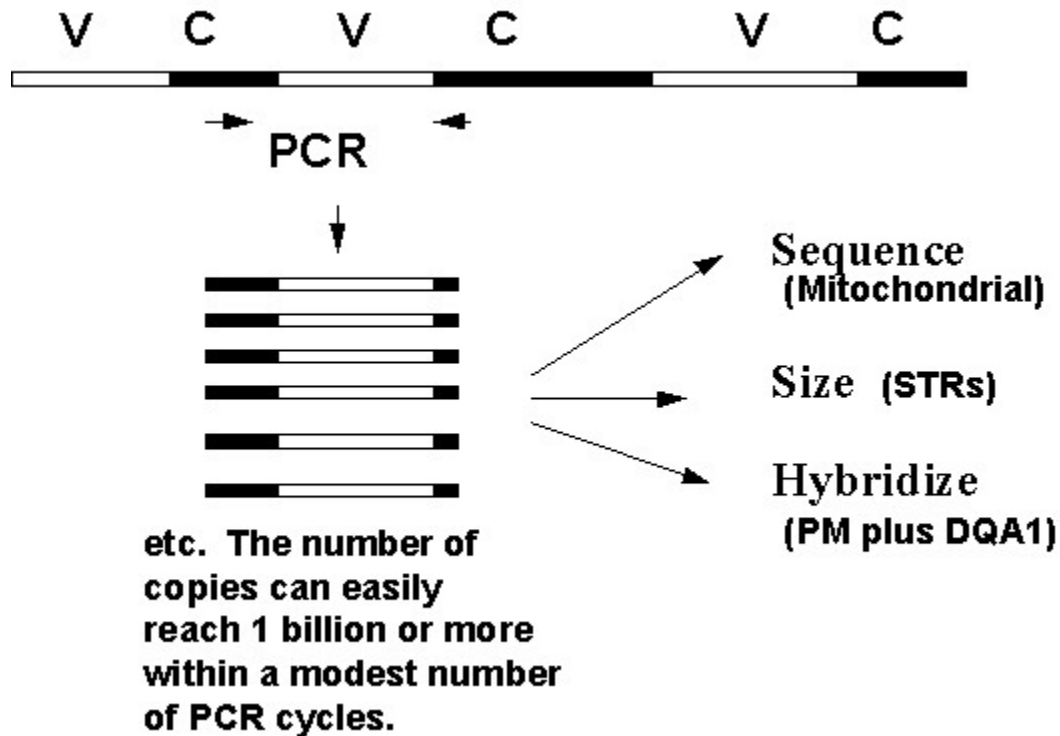
PCR uses constant regions of DNA sequence to prime the copying of variable regions of DNA sequence.

DNA HAS CONSERVED (the same for everyone) AND VARIABLE REGIONS (differing among people-polymorphic).

The variable regions (V) are usually interspersed among the conserved or constant (C) regions.



PCR typically uses two short pieces of known DNA called primers (small arrows below). These serve as starting points for the copying of a region of DNA.



Many forensic laboratories use commercial supplied DNA testing kits that contain key components for certain PCR-based tests. PM plus DQA1TM, Profiler PlusTM and CofilerTM and IdentifilerTM are all test kits commercially supplied by *PE Applied Biosystems*. PowerPlexTM is another test kit with variations supplied by *Promega*. PowerPlex kits have published primers, an advantage if the precise DNA targeted is to be recorded for posterity or studied for research. As of 2005, Profiler PlusTM and CofilerTM and PowerPlexTM are probably the most commonly used test kits in US forensic laboratories.

PCR-based testing often requires less DNA than RFLP testing and the DNA may be partially degraded, more so than is the case with RFLP. However, PCR still has sample size and degradation limitations that sometimes may be under-appreciated. PCR-based tests are also extremely sensitive to contaminating DNA at the crime scene and within the test laboratory. During PCR, contaminants may be amplified up to a billion times their original concentration. Contamination can influence PCR results, particularly in the absence of proper handling techniques and proper controls for contamination. PCR is less direct and somewhat more prone to error than RFLP.

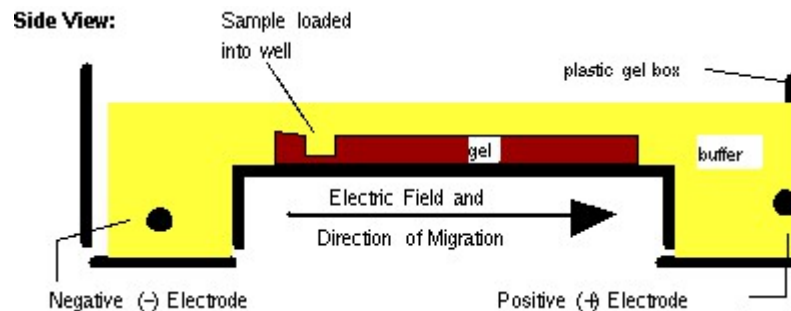
However, PCR has tended to replace RFLP in forensic testing primarily because PCR based tests are faster and more sensitive.

3.4.2 Gel Electrophoresis

DNA fragments of different sizes can be separated by an electrical field applied to a “gel”. The negatively charged DNA migrates away from the negative electrode and to the positive electrode. The smaller the fragment the faster it migrates. This technique separates molecules on the basis of their size.

- Cast slab of gel material, usually agarose or polyacrylamide. The gel is a matrix of polymers forming sub-microscopic pores.
- The size of the pores can be controlled by varying the chemical composition of the gel.

The gel is set up for electrophoresis in a tank holding pH buffer. Electrodes apply an electric field:



The molecules to separate (DNA RNA) carry a net **negative charge (why?)** so they move along the electric field toward the **positive cathode**. (To separate proteins, a detergent would be included which coats the protein with negative charge.)

The larger molecules are held up as they try to pass through the pores of the gel, while the smaller molecules are impeded less and move faster. This results in **separation by size**, with the larger molecules nearer the well and the smaller molecules farther away.

Note that this separates on the basis of **size (volume in solution)**, which is **not necessarily molecular weight**. For example:

- Two DNA molecules of the same molecular weight will run differently if one is supercoiled, because the supercoils constrain the shape to be smaller.
- Two RNA molecules of the same molecular weight will run differently if one has much intramolecular base pairing, making it "smaller."

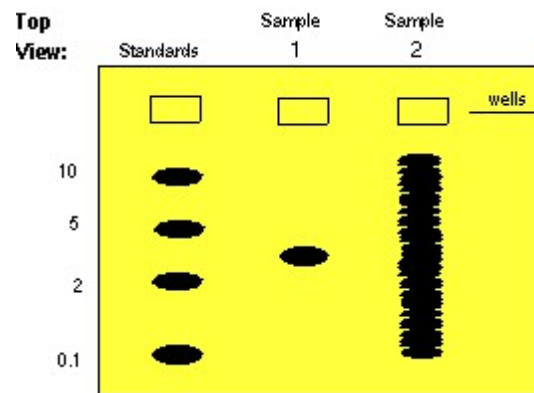
Aside from the above exceptions, the distance migrated is **roughly proportional to the log of the inverse of the molecular weight** (the log of $1/MW$). Gels are normally depicted as running vertically, with the wells at the top and the direction of migration downwards. This leaves the large molecules at the top and the smaller molecules at the bottom. Molecular weights are measured with different units for DNA, RNA, and protein:

- **DNA: Molecular weight is measured in base-pairs, or bp, and commonly in kilobase-pairs (1000bp), or kbp.**
- **RNA: Molecular weight is measured in nucleotides, or nt, and commonly in kilonucleotides (1000nt), or knt. [Sometimes, bases, or b and kb are used.]**
- **Protein: Molecular weight is measured in Daltons (grams per mole), or Da, and commonly in kiloDaltons (1000Da), or kDa.**

Molecular weight standards run in one well of the gel are used to calibrate the molecular weights of sample molecules. Below is a gel stained with a dye: a colored molecule which binds to a specific class of macromolecules in a sequence-independent manner (probes bind in a sequence-dependent manner).

Sample 1 contains only one size class of macromolecule - it could be a plasmid, a pure mRNA transcript, or a purified protein. In this case, you would not have to use a probe to detect the molecule of interest since there is only one type of molecule present. Blotting is usually necessary for samples that are not complex mixtures. By interpolation, its molecular weight is roughly 3.

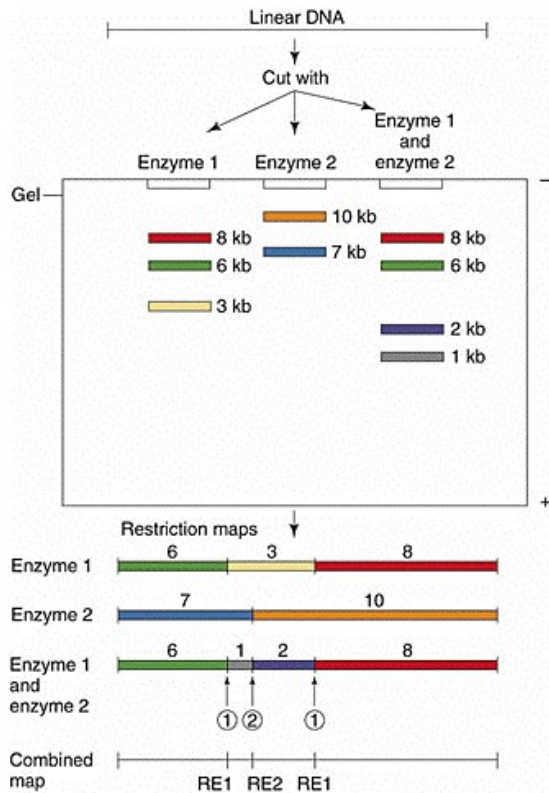
Sample 2 is what a sample of total DNA cut with a restriction enzyme, total cellular RNA, or total cellular protein would look like in a gel stained with a sequence-independent stain. There are so many bands that it is impossible to find the one we are interested in. Without a probe (which acts like a sequence-dependent stain) we cannot get very much information from a sample like this.



Different **stains** are used for different classes of macromolecules. DNA and RNA are generally stained with **ethidium bromide (EtBr)**, an **intercalating agent**. The DNA-EtBr complex fluoresces under UV light. Protein is stained with Coomassie Blue or Silver Stain.

3.4.3 Restricted Fragment Length Polymorphism (RFLP)

Frequently it is important to have a restriction enzyme site map of a cloned gene for further manipulations of the gene. This is accomplished by digestion of the gene singly with several enzymes and then in combinations. The fragments are subjected to gel electrophoresis to separate the fragments by size and the sites are deduced based on the sizes of the fragments.



In this example, digestion with Enzyme 1 shows that there are two restriction sites for this enzyme, but does not reveal whether the 3 kb segment is in the middle or on the end of the digested sequence, which is 17 kb long. Combined digestion by both enzyme 1 and enzyme 2 leaves the 6 and 8 kb segments intact but cleaves the 3 kb segment, showing that enzyme 2 cuts within this enzyme 1 fragment. If the 3 kb section were on the outside of the fragment being studied, digestion by enzyme 2 alone would yield a 1 or 2 kb fragment. Since this is not the case, of the three restriction fragments produced by enzyme 1, the 3 kb fragment must lie in the middle. That the RE2 site lies closer to the 6 kb section can be inferred from the 7 and 10 kb lengths of the enzyme 2 digestion.

RFLP DNA testing has four basic steps:

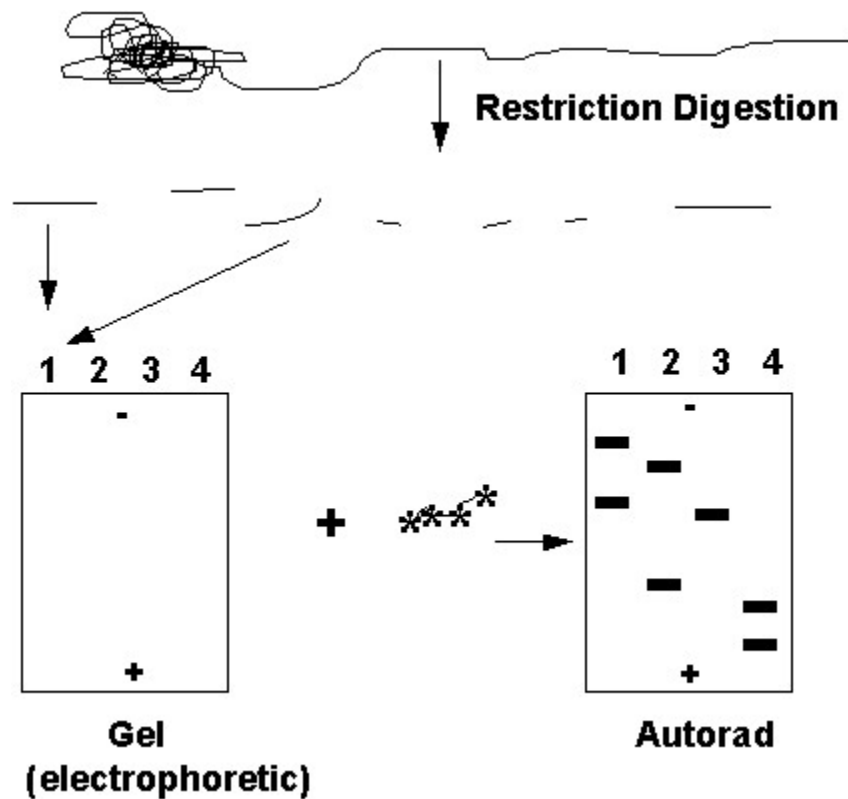
1. The DNA from crime-scene evidence or from a reference sample is cut with something called a restriction enzyme. The restriction enzyme recognizes a particular short sequence such as AATT that occurs many times in a given cell's DNA. One enzyme commonly used is called *Hae* III (pronounced: Hay Three) but the choice of enzyme varies. For RFLP to work, the analyst needs

thousands of cells. If thousands of cells are present from a single individual, they will all be cut in same place along their DNA by the enzyme because each cells DNA is identical to every other cell of that person.

2. The cut DNA pieces are now sorted according to size by a device called a gel. The DNA is placed at one end of a slab of gelatin and it is drawn through the gel by an electric current. The gel acts like a sieve allowing small DNA fragments to move more rapidly than larger ones.

3. After the gel has separated the DNA pieces according to size, a blot or replica of the gel is made to trap the DNA in the positions that they end up in, with small DNA fragments near one end of the blot and large ones near the other end. The blot is now treated with a piece of DNA called a probe. The probe is simply a piece of DNA that binds to the DNA on the blot in the position were a similar sequence (the target sequence) is located.

4. The size or sizes of the target DNA fragments recognized by the probe are measured. Using the same probe and enzyme, the test lab will perform these same steps for many people. These sizes and how they distribute among large groups of people form a database. From the database a rough idea of how common a given DNA size measured by a given probe is found. The commonness of a given size of DNA fragment is called a population frequency.



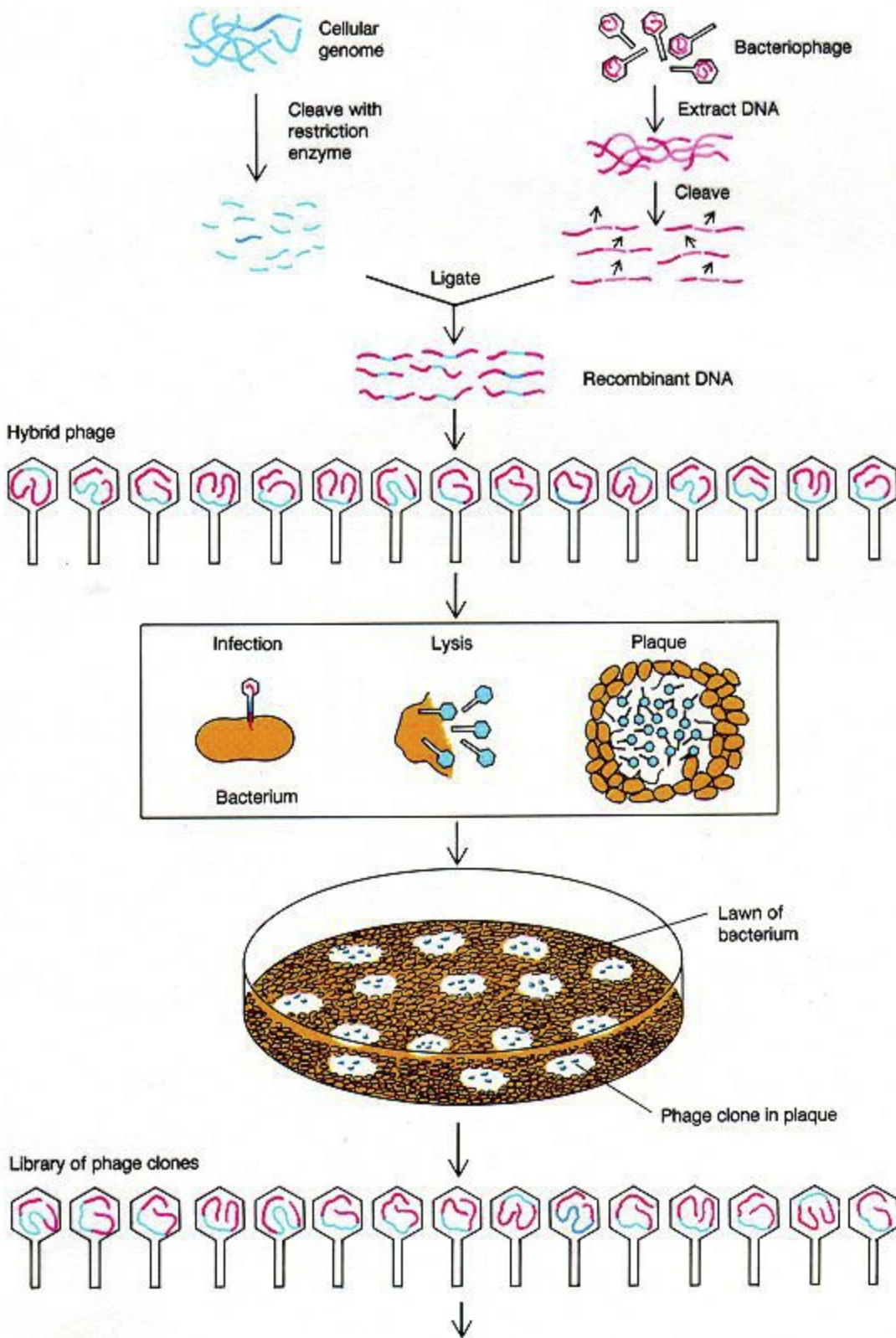
The restriction enzyme cuts the DNA into thousands of fragments of nearly all possible sizes. The sample is then electrophoretically separated. The DNA at this point is invisible in the gel unless the DNA is stained with a dye. A replica of the gel's DNA is made on something called a blot (also called a Southern blot) or membrane. The blot is then probed (mixed with) a special preparation of DNA that recognizes a specific DNA sequence or locus. Often, the probe is a radioactively labeled DNA sequence (represented by * labeled object in the figure above). Excess probe is washed off the blot, then the blot is laid onto X-ray film. Development reveals bands indicating the sizes of the alleles for the locus within each sample. The film is now called an "autorad." The band sizes are measured by comparing them with a "ladder" of known DNA sizes that is run next to the sample. A match may be declared if two samples have RFLP band sizes that are all within 5% of one another in size.

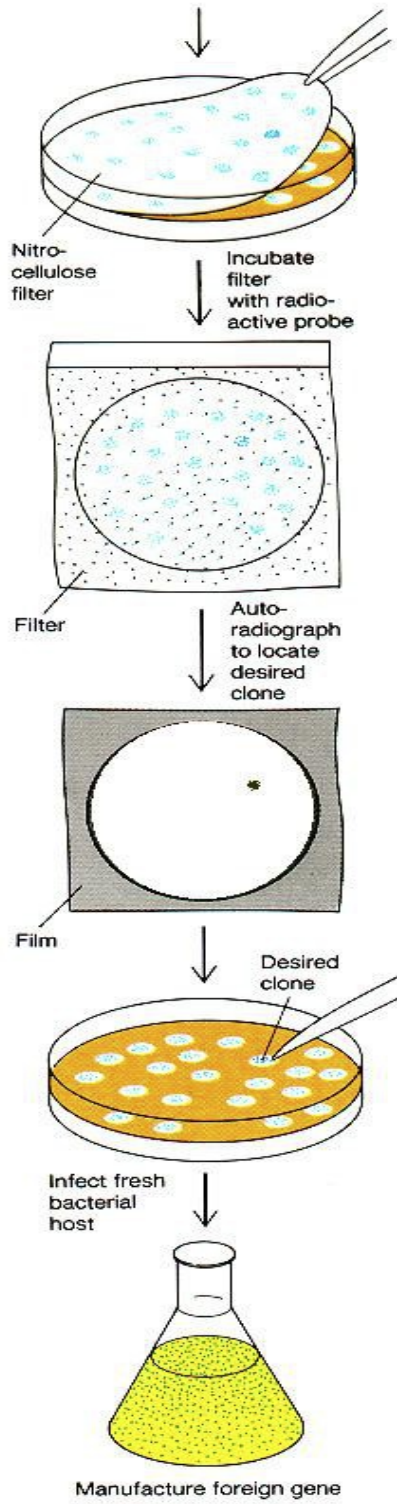
For RFLP analysis to be reliable, all complex steps of the analysis must be carefully controlled. Databases must be large meaning they include many people; they must be representative of the potential test subjects. Because of the complexities of populations, databases must be interpreted with extreme care. For example, DNA fragment sizes rare in one population may be very common

in other populations. Further, sub-populations or populations within populations must be considered.

3.4.4 DNA Probes

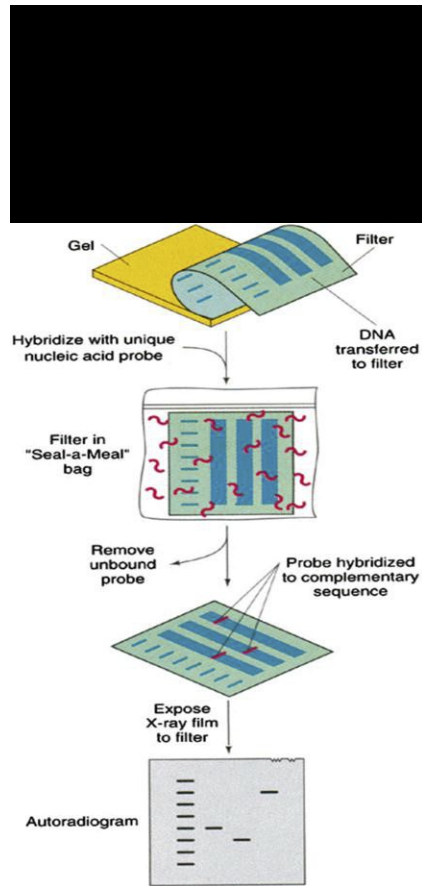
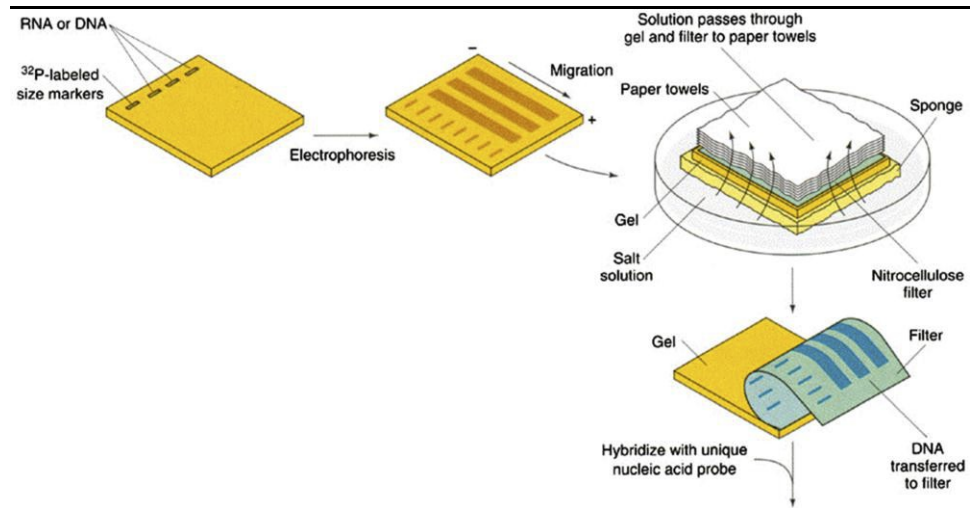
How would we actually locate the appropriately cloned gene? There are many different ways, depending on the specific case. Here is one example, in which a partial sequence of the protein enables us to [reverse the code](#) and determine an approximate DNA sequence to use for a [radiolabeled probe](#). The DNA probe will [hybridize](#) to clones containing the correct DNA, even if it is just one piece cut out of an entire genome.





3.4.5 Southern Blot

A Southern blot allows the detection of a gene of interest by probing DNA fragments that have been separated by electrophoresis with a “labeled” probe.



(From: AN INTRODUCTION TO GENETIC ANALYSIS 6/E BY Griffiths, Miller, Suzuki,

Leontin, Gelbart 1996 by W. H. Freeman and Company. (to seek for permission)

References:

1. The Basics of Recombinant DNA
2. Griffiths et al. 1996. **An introduction to genetic analysis** 6th Ed, Miller, W. H. Freeman and Company
3. Basic Laboratory Techniques in Molecular Biology.
<http://www.youtube.com/watch?v=Gy21KhSF3PM>
4. [MIT hypertextbook](#)

MODULE FOUR

Biotechnology Applications in Animal Agriculture

Introduction:

The modern biotechnological tools have had a remarkable influence on animal biotechnology as well. Many innovative techniques are constantly being used around the world to improve upon livestock. The foundation of this approach lies in alteration at various biochemical and molecular levels. There are a large number of technologies that have been developed for or adapted to the livestock of both developed and developing countries. However, the major technologies that are used effectively in livestock production in the developing world include conserving animal genetic resources, augmenting reproduction, embryo transfer (ET) and related technologies, diagnosing disease and controlling and improving nutrient availability. These techniques are proving extremely useful in developing disease-resistant, healthy, and more productive animals. Some of the areas where biotechnology can prove useful are presented in more details this module.

4.1 Reproductive physiology

One of the challenges for genetic improvement is to increase reproduction rates. Several reproduction techniques are available. The commonest of these are artificial insemination (AI), embryo transfer and associated technologies. Measurement of progesterone in milk or blood which is a widely used technique for monitoring ovarian function and for pregnancy tests is also an important technology for managing the reproductive function of the animal.

4.1.1 Artificial insemination

No other technology in agriculture, except hybrid seed and fertiliser use, has been so widely adopted globally as AI. Progress in semen collection and dilution, and cryopreservation techniques now enables a single bull to be used simultaneously in several countries for up to 100,000 inseminations a year (Gibson and Smith 1989). This implies that a very small number of top bulls can be used to serve a large cattle population. In addition, each bull is able to produce a large

number of daughters in a given time thus enhancing the efficiency of progeny testing of bulls. The high intensity and accuracy of selection arising from AI can lead to a four-fold increase in the rate of genetic improvement in dairy cattle relative to that from natural mating.

A wider and rapid use of selected males through AI will accelerate the rate of genetic improvement. Also, use of AI can reduce transmission of venereal diseases in a population and the need for farmers to maintain their own breeding males, facilitate more accurate recording of pedigree and minimise the cost of introducing improved stock. However, success of AI technology depends on accurate heat detection and timely insemination. The former requires a certain level of experience among farmers while the latter is dependent on good infrastructure, including transport network, and availability of reliable means of transport.

Though AI is widely available in developing countries it is used far less, particularly in Africa, than in developed countries. Its use has been limited largely to "exploratory" purposes mainly by research institutions. A few countries including Botswana, Ethiopia, Ghana, Malawi, Mali, Nigeria, Senegal and Sudan have taken the technology to the field, mostly for programmes of "upgrading" indigenous stock and as a service to a limited number of commercial farmers keeping exotic dairy cattle breeds. A few others have used the technology more widely. Kenya and Zimbabwe, for example, have elaborate AI systems which include national insemination services incorporating progeny testing schemes. However, even these have gone through periods of collapse or serious degeneration and have had to go through "rehabilitation" phases.

AI technology use is still more generally associated with dairy cattle than other domestic livestock species. The limitations of AI use in beef cattle include the difficulty in detecting heat in large beef herds kept on ranches and the less frequent handling of individual cows. In sheep and goats the failure to develop a simple, non-surgical insemination procedure has prevented extensive exploitation of the technology in sheep. AI is credited for providing the impetus for many other developments which have had a profound impact on reproductive biotechnology, such as embryo-transfer technology.

4.1.2 Embryo transfer (ET)

Although not economically feasible for commercial use on small farms at present, embryo technology can greatly contribute to research and genetic improvement in local breeds. There are two procedures presently available for production of embryos from donor females. One consists of superovulation, followed by AI and then flushing of the uterus to gather the embryos. The other, called *in vitro* fertilization (IVF), consists of recovery of eggs from the ovaries of the female then maturing and fertilizing them outside the body until they are ready for implantation into foster females. IVF facilitates recovery of a large number of embryos from a single female at a reduced cost thus making ET techniques economically feasible on a larger scale. Additionally, IVF makes available embryos suitable for cloning.

The principal benefit of embryo transfer is the possibility to produce several progeny from a female, just as AI can produce many offspring from one male. For example the average lifetime production of a cow can be increased from 4 to 25 calves. Increasing the reproductive rate of selected females has the following benefits: genetically outstanding animals can contribute more to the breeding programme, particularly if their sons are being selected for use in AI; the rate of genetic change can be enhanced with specially designed breeding schemes which take advantage of increased intensity of female selection combined with increased generation turnover; transport of embryos is much cheaper than that of live animals; risk of importing diseases is avoided; facilitates rapid expansion of rare but economically important genetic stocks; and the stress to exotic genotypes can be avoided by having them born to dams of local breeds rather than importing them as live animals.

Embryo transfer is still not widely used despite its potential benefits. In developing countries this is mainly due to absence of the necessary facilities and infrastructure. Even in developed countries, cost considerations still limit the use of commercial embryo transfer in specialised niches or for a small proportion of best cows in the best herds. Bovine embryo transfer is another technique of genetic manipulation. The principal advantage of embryo transfer is that it increases the reproductive capacity of useful cattle like cows and buffaloes. Such transfer can also decrease the generation interval between selection steps by having a large percentage of progeny of young donors. In some cases, embryo transfer even allows cows and buffaloes that have been rendered

infertile due to a disease, injury or ageing, to have progeny. Embryo transfer (ET) techniques have also been developed for camels and calves. Commercial embryo transfer is more popular with cattle than other species. This is mainly because ET is relatively easier in cattle than the other species and also because it is more economical in cattle (i.e. cattle are worth more). Additionally, the low reproductive rate and the long generation interval of cattle make ET much more advantageous in cattle.

Production of several closely related, and hence genetically similar, individuals through ET techniques can make critical contributions to research. For example a project at the International Laboratory for Research on Animal Diseases (ILRAD) to locate the genes responsible for tolerance of some cattle populations to trypanosomiasis required large numbers of closely related crosses of trypanotolerant and trypanosusceptible cattle. Use of ET has made it possible to generate such families thereby facilitating the search for genetic markers of trypanotolerance. Additionally, ET could be useful in studying the extent to which a trait is influenced by the embryo (direct component) or the reproductive tract (maternal component).

4.2. Transgenic animals

A transgenic animal is an animal whose hereditary DNA has been augmented by addition of DNA from a source other than parental germplasm through recombinant DNA techniques. Transfer of genes or gene constructs allows for the manipulation of individual genes rather than entire genomes. The technique has now become routine in the mouse and resulting transgenic mice are able to transmit their transgenes to their offspring thereby allowing a large number of transgenic animals to be produced. Successful production of transgenic livestock has been reported for pigs, sheep, rabbits and cattle (Niemann et al 1994).

Transgenesis offers considerable opportunity for advances in medicine and agriculture. In livestock, the ability to insert new genes for such economically important characteristics as fecundity, resistance to or tolerance of other environmental stresses would represent a major breakthrough in the breeding of commercially superior stock. Another opportunity that transgenic technology could provide is in the production of medically important proteins such as insulin and

clotting factors in the milk of domestic livestock. The genes coding for these proteins have been identified and the human factor IX construct has been successfully introduced into sheep and expression achieved in sheep milk. Moreover, the founder animal has been shown to be able to transmit the trait to its offspring.

The first reports of the production of transgenic animals created a lot of excitement among biological scientists. In the field of animal breeding, there were diverse opinions on how the technology might affect livestock genetic improvement programmes. Some believed that it would result in total reorganisation of conventional animal breeding theory while others considered the technology as an extension of current animal breeding procedures which, by broadening the gene pool, would make new and novel genotypes available for selection.

Genetic engineering has been used to introduce foreign genes into the animal genome or, alternatively, to knock out selected genes. Genes controlling growth were introduced into pigs to increase growth and improve carcass quality. Currently, research is underway to engineer resistance to diseases that affect the animals or that pose an indirect risk to human health, such as Marek's disease and salmonellosis in poultry, scrapie in sheep and mastitis in cattle. Other studies have tried to increase the casein content of milk or to engineer animals that produce pharmaceutical or industrial chemicals in their milk or semen. No agricultural applications have yet proved commercially successful. Nuclear transfer (NT) technology now provides an alternative route for cell-based transgenesis in domestic species, offering new opportunities for genetic modification. Livestock that produce human therapeutic proteins in their milk, that have organs suitable for xenotransplantation and that are resistant to diseases such as spongiform encephalopathies have been produced by NT from engineered cultured somatic cells. Although gene-based technologies have the potential to improve the efficiency of livestock production, thereby ensuring better returns for the farmers, the economic impact of transgenics in the livestock sector will be much less than in the crop sector.

4.2.1 Transgenic Sheep:

Dolly, the sheep was created in Scotland in 1997 by the nuclear transfer technique. Here the nucleus of a 'donor' mammary cell was injected into a recipient cell (egg) (the nucleus of which had been removed). This cell was then implanted into a receptive surrogate mother, and it eventually developed into Dolly - the 'clone of the donor'. This was followed by the birth of Polly - the transgenic lamb containing a human gene

The development of Dolly and Polly, the first cloned animals created waves all over the world. This feat is indeed significant, as it not only marks a great scientific achievement, but also paves the way for the generation of many other cloned animals, which carry valuable human proteins.

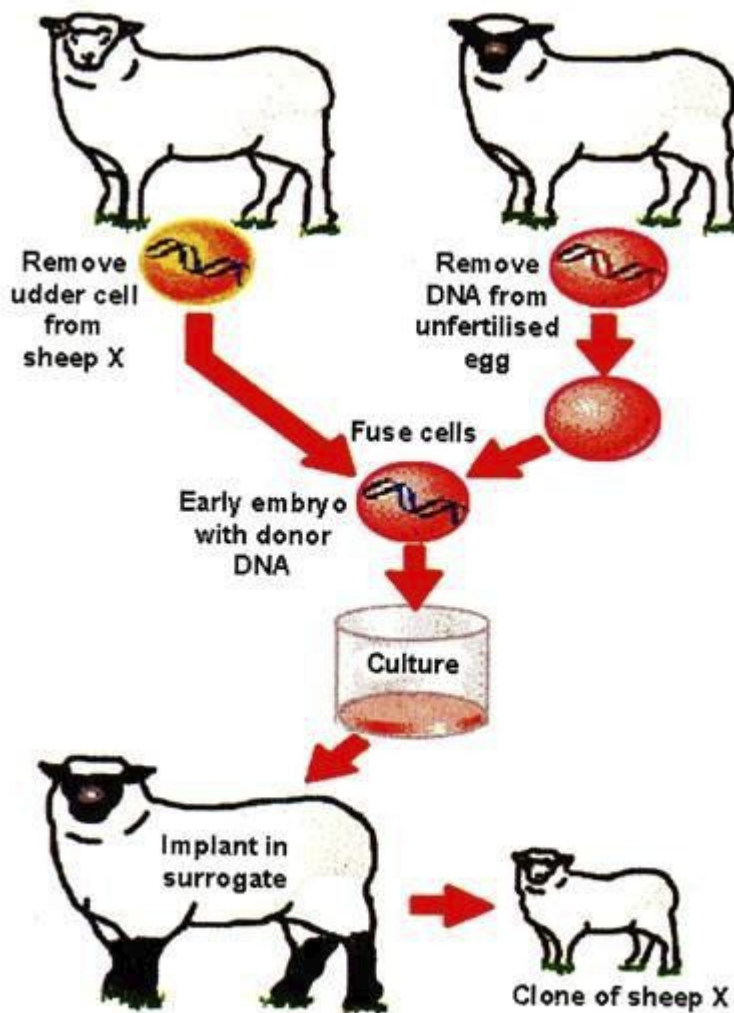


Figure 3: Cloning of Dolly (Sheep)

4.2.2 Transgenic Goat:

In this case the foetal cells were obtained from a thirty day-old female goat foetus. The AT III gene, a human gene encoding anti-clotting protein was hooked to the promoter, and injected into the nucleus of the newly fertilised egg.

After removing the nucleus of the recipient egg cell (enucleated condition), the donor egg cell was fused with foetal fibroblast cells possessing the human gene. Subsequently, the cloned embryo was transferred into a recipient female goat mother.

The female offspring thus developed is capable of producing milk containing human protein. This protein can be easily extracted from the milk and used for numerous pharmaceutical purposes. The development of these goats with human genes is one of the first applications of the nuclear transfer process.

PPL Therapeutics, a UK- based company has already developed five transgenic lambs. The company's director, Dr. Alan Colman says that these lambs are the realisation 'of the vision to produce instant flocks or herds that produce high concentrations of valuable therapeutic proteins very quickly. Of late, pigs have also been cloned using more innovative cloning techniques. These pigs could be very useful for the food industry.

4.3 Animal health

4.3.1 Disease diagnosis

Successful control of a disease requires accurate diagnosis. This has been greatly improved in recent years through developments in biotechnology. The most recent major development, the finding that it is possible to immortalise individual antibody-producing cells by hybridisation to produce antibodies of a given class, specificity and affinity (i.e. monoclonal antibodies) has provided a tool that permits the analysis of virtually any antigenic molecule. The use of monoclonal antibodies has revealed that the failure of vaccines (e.g. of rabies) to provide protection in all parts of the world was due to the diversity in the antigenic composition of the causative virus. The

(monoclonal antibody) technology is relatively simple and can readily be applied in developing countries. Monoclonal antibodies are currently supplied to developing countries directly or in the form of kits and simple reagents for completion of the tests. For example, kits for rinderpest virus diagnosis used in African countries come in this form.

The ability to generate highly specific antigens by recombinant DNA techniques has made it possible for an increasing number of enzyme-linked immunosorbent assays (ELISA) to have the capacity to differentiate between immune responses generated by vaccination from those due to infection. This has made it possible to overcome one of the major drawbacks of antibody detection tests: the fact that, because antibodies can persist in animals for long periods, their presence may not indicate current infection. ILRAD has developed a technique to overcome this problem in diagnosis of trypanosomiasis. The parasite antigen detection test uses monoclonal antibodies raised in laboratory mice to capture the parasite antigens which are then revealed by their reaction with a second layer antibody to which is conjugated an easily detectable enzyme. This test reveals current infections and facilitates differentiation between the major trypanosome species. This has important implications for disease control, especially because of the association of different parasite species with different epidemiological and disease circumstances.

The advent of PCR has enhanced the sensitivity of DNA detection tests considerably. For example, PCR used in combination with hybridisation analysis, has been shown (Brandon et al 1991) to provide a sensitive diagnostic assay to detect bovine leukosis virus. Other diagnostic techniques include nucleic acid hybridisation (NAD) and restriction endonuclease mapping (REM). As has been indicated above, one of the most valuable features of these molecular techniques is their specificity and sensitivity. A good example of the specificity of NAD is its application in distinguishing infections caused by *peste des petite ruminants* (PPR) virus from rinderpest diseases whose symptoms are clinically identical and which cannot be distinguished antigenically with available serological reagents. This technique also allows comparison of virus isolates from different geographical locations.

Other available diagnostic techniques which may have application in small ruminants and/or cattle include:

1. Nucleic acid probes (NAP) for heartwater (*Cowdria ruminantium*), *chlamydia psitacci*, Paratuberculosis and Bluetongue (Blancou 1990; Knowles and Gorham 1990);
2. Restriction endonuclease reaction (RENr) for diagnosis of *Corynebacterium pseudotuberculosis* (Knowles and Gorham 1990);
3. PCR for characterising subtypes of Bluetongue from different geographical regions (Osburn 1991);
4. Monoclonal antibodies (MAB) for differentiating false positive anti-Brucella titres caused by *Yersinia enterocolitica* and true positive anti-Brucella titres in latent infected animals (Haas et al 1990); and MAB for diagnosis of Toxoplasmosis, Pasteurellosis, *Mycoplasma* spp, PPR and Boder Disease (Blancou 1990; Lefevre and Diallo 1990).

Examples of tropical diseases for which diagnostic tests that are currently available or which could be developed in the near future are presented in Table 4.1.

Table 4.1. Tropical diseases for which probes and monoclonal antibodies (MAB) are available.

Viral diseases	Bacterial diseases
Rinderpest	Cowdriosis
Peste des petite ruminants	Contagious bovine pleuropneumonia
Bluetongue	Contagious agalactia
African horse sickness	Contagious caprine pleuropneumonia
Foot-and-mouth disease	Anaplasmosis (<i>A. centrale</i> and <i>A. marginale</i>)
Vesicular stomatitis	Haemorrhagic septicemia (only MAB available)

Source: Lefevre (1992).

4.3.2 Vaccines

Conventional means of controlling major livestock diseases include chemotherapy, vector control, vaccination, slaughter of infected stock, and other management practices (including grazing management and controlled stock movements). Vector control requires continuous application of pesticides. These are often unaffordable to farmers in developing country. Moreover, where these

drugs or pesticides are used, resistance by parasites is often encountered and reinfection following administration of drugs against parasitic diseases usually occurs. Additionally, in many cases drugs are not readily available locally. In some cases where they are available, they are ineffective, either because they have been partly preserved or they are not genuine.

Immunisation remains one of the most economical means of preventing specific diseases. An effective vaccine can produce long-lasting immunity. In some cases, vaccination can provide lifetime immunity. Moreover a small number of doses is usually required for protection. Level of infrastructure and logistical support required for a large-scale vaccination programme is such that a successful vaccination campaign can be implemented in remote rural areas. In general, vaccines offer a substantial benefit for comparatively low cost, a primary consideration for developing countries.

Vaccines have conventionally been produced by several methods some of which have become rather static with regard to efficacy, safety, stability and cost. Very effective vaccines against animal diseases such as rinderpest and pig cholera have been in use for more than 20 years and have helped to significantly reduce the incidence of these diseases world-wide. However, vaccines of questionable efficacy also exist. Impotency, instability, adverse side effects, and reversion of attenuated organisms to wild (disease-causing) forms represent some of the problems. However, research strategies for the development of better, cheaper and safer vaccines are constantly being sought. Through the use of monoclonal antibodies and recombinant DNA technologies, it is now possible to define and produce immunogenic components much more rapidly. These technologies are increasingly being used to clarify the pathogenetic mechanisms and immune response to microbial diseases. This should lead to the production of more effective vaccines in the future. To date many candidate vaccines have been produced by these techniques. However, only few of these are being produced commercially.

Billions of dollars are spent each year to improve upon farm animals and their health care. Scientists are now trying to use recombinant DNA technology to produce vaccines for animal stocks. An extremely effective vaccine has already been developed for swine pseudo rabies

(herpes-virus). This vaccine was conventionally manufactured by killing disease causing microbes.

Modern recombinant vaccines are not injected with these germs. They are thus safe to use, and involve no such risk. Conventional vaccine production is a high-cost, low volume affair. But modern recombinant production systems have opened up new vistas in the huge market for efficient vaccines. Recombinant vaccines also score on their fast pace of development. The conventional vaccines can take as long as twenty to thirty years of research and experimenting before they are ready to use. This has been causing a deficiency of important vaccines. The modern vaccines are made ready in a much shorter time span. What's more, these vaccines are active even at room temperature. Their movement and storage thus becomes much easier. Table 4. 2 summarises some vaccines developed by recombinant DNA technology.

Table 4. 2. Example of some novel animal vaccines.

Species	Disease	Producer
Avian	Coccidiosis,	Genex and A.H. Robins
	Newcastle virus	Codon and Salsbury labs
Bovine	Papilloma virus	Molecular genetics
	Viral diarrhoea	California Biotechnology
	Brucellosis	Ribi ImmunoChem
	Rinderpest	USDA and University of California, Davis
Swine	Parvovirus	Applied Biotechnology
	Dysentery	Codon
Equine	Influenza, Herpes	California Biotechnology
		Applied Biotechnology
Companion Canine	parvovirus	Applied Biotechnology

4.4 Physiology of lactation and growth

Recombinant bovine somatotropin (BST) is a genetically engineered synthetic analog of the natural growth hormone. Since the 1970s there have been a number of studies on the effects of BST on milk yield, reproductive performance and health as well as its likely effect on humans who consume such milk. Under good management and feeding, regular BST administration to lactating dairy cows increases milk yield by 15-30% and also increases efficiency of milk production.

BST is now readily available and is already in commercial use in the United States. However, the appropriateness of BST use to increase milk production in the USA is doubtful as the country already has a milk surplus and the public is also concerned about its effects on health. Indeed, the benefit of BST in the USA is perceived to be towards reducing the production costs of large dairy farms, but this could push smaller farmers out of the market. The appropriateness of BST use for developing countries is still a matter of debate. Those supporting its introduction argue that its use in commercial dairy farms could increase the national milk output. Those opposed point out the fact that BST does not improve milk yield in indigenous non-dairy breeds and that its use on crossbred and exotic dairy cattle will require more feeding, and that provision of adequate nutrition is already a problem for most dairy operations in developing countries.

Additionally, the effect of BST on reproductive performance is likely to be more adverse in the presence of higher biotic and abiotic stresses, including nutritional stresses. There is also need to examine the economics of BST application in view of the known association of its use with mastitis and other infections. BST use is thus bound to be associated with increased use of antibiotics and other veterinary drugs. Thus, in evaluating the potential role of BST in developing countries, one needs to consider not only the possible response levels of the cattle in these countries to BST treatment, but also the cost of BST, the amount and cost of other incremental inputs required for effective use of BST, and the milk prices. Ultimately, the main technical constraint to BST use in developing countries will not only be its cost, but the absence of an efficient delivery system; current use of the technology requires regular injections.

Porcine somatotropin (PST) and recombinant growth hormone stimulatory peptides (e.g. growth hormone releasing factor, GRF) along with BST have been shown to increase growth rates by 8-

38% in cattle, sheep and pigs. In almost all cases, administration of exogenous growth hormones have been associated with increased carcass protein and reduced carcass fat. Other growthpromoting agents (e.g. anabolic steroids and beta agonists) have been shown to have even larger effects but public concerns over the possible residual effects in meats have led to their being banned in most developing countries. The use of anabolic implants is, however, permitted in some countries such as the USA.

4.5 Animal nutrition

Nutrition represents one of the most serious limitations to livestock production in developing countries, especially in the tropics. Feed resources are inadequate in both quality and quantity, particularly during the dry seasons. Biotechnological options are available for improving rumen fermentation and enhancing the nutritive value and utilisation of agro-industrial by-products and other forages. Animal nutrition is another major concern that can be addressed through biotechnological tools. We have seen how certain bacteria have been efficiently used to over-express proteins for medicinal applications. Similarly, animal proteins like somatotropins can be over-expressed in bacteria and generated in larger quantities for commercial purposes.

Fibrous feeds, including crop residues, of low digestibility constitute the major proportion of feeds available to most ruminants under smallholder situations in developing countries. The associated low productivity can be overcome to some extent by several means, among which are: balancing of nutrients for the growth of rumen microflora thereby facilitating efficient fermentative digestion and providing small quantities of by-pass nutrients to balance the products of fermentative digestion, enhancing digestibility of fibrous feeds through treatment with alkali or by manipulating the balance of organisms in the rumen and genetic manipulation of rumen micro-organisms, currently acknowledged as potentially the most powerful tool for enhancing the rate and extent of digestion of low quality feeds. Rumen micro-organisms can also be manipulated by adding antibiotics as feed additives, fats to eliminate or reduce rumen ciliate protozoa (defaunation), protein degradation protectors, methane inhibitors, buffer substances, bacteria or rumen content and/or branched chain volatile fatty acids.

4.5.1 Increasing digestibility of low-quality forages

Low-quality forages are a major component of ruminant diets in the tropics. Thus, much progress can be made by improving the forage component of the ration. The characteristic feature of tropical forages is their slow rate of microbial breakdown in the rumen with the result that much of the nutrients of the feed are voided in the faeces. The slow rate of breakdown also results in reduced outflow rate of feed residues from the rumen which consequently depresses feed intake. At present, the main treatment methods for forages such as cereal straws are either mechanical (e.g. grinding), physical (e.g. temperature and pressure treatment) or a range of chemical treatments of which sodium hydroxide or ammonia are among the more successful.

The lignification of the cell walls prevents degradation by cellulase or hemicellulase enzymes. Fortunately, it is possible to use lignase enzyme produced by the soft-rot fungus (*Phanerochaete chrysosporium*) which causes a high degree of depolymerisation of lignin. The enzyme acts like a peroxidase and causes cleavage of carbon-carbon bonds. At present the levels of the lignase enzyme produced by the basidiomycete fungi are insufficient for the treatment of straw on a commercial scale. However, it is conceivable that the use of recombinant DNA engineering techniques will allow the modification of the lignase genes and associate proteins to increase their efficiency and stability. The lignin gene has to date been cloned and sequenced from *P. chrysosporium*.

4.5.2 Improving nutritive value of cereals

Moderate protein content and low amounts of specific amino acids limit the nutritive value of cereals and cereal by-products (e.g. barley is low in lysine and threonine). This is a major limitation in the ration formulation for non-ruminant livestock which necessitates addition of expensive protein supplements. There are on-going studies to enhance the low level of lysine in barley by genetically engineering the grain genome. Genetic modification through insertion of genes into rice protoplasts and generation of transformed plants has already been achieved.

4.5.3 Removing anti-nutritive factors from feeds

Anti-nutritive factors in plant tissues include protease inhibitors, tannins, phytohaemagglutinins and cyanogens in legumes, and glucosinolates, tannins and sinigrin in oilseed rape (*Brassica napus*) and other compounds in feeds belonging to the *Brassica* group. As with amino acid deficiencies, the adverse effects of these compounds are more marked in non-ruminants than in ruminants. Conventional plant breeding has been used to reduce and, in some cases, eliminate such anti-nutritive factors. An example is the introduction of cultivars of oilseed rape which are low in, or free from erucic acid and glucosinolates. A combination of genetic engineering and conventional plant breeding should lead to substantial reduction or removal of the major antinutritive factors in plant species of importance as animal feeds.

Transgenic rumen microbes (see below) could also play a role in the detoxification of plant poisons (Gregg 1989) or inactivation of antinutritional factors. Successful introduction of a caprine rumen inoculum obtained in Hawaii into the bovine rumen in Australia to detoxify 3-hydroxy 4(IH) pyridine (3,4 DHP), a breakdown product of the non-protein amino acid mimosine found in *Leucaena* forage (Jones and Megarritty 1986) demonstrates the possibilities.

4.5.4 Improving nutritive value of conserved feed

The conservation of plant material as silage depends upon anaerobic fermentation of sugars in the material which in turn is influenced by the ability of naturally occurring lactic acid bacteria to grow rapidly on the available nutrients under the existing physical environment. Unless the ensiled material is sterilised, lactic acid bacteria are always present. However, the ensiling conditions may not always be ideal for their development. In addition to the number and type of bacteria, other interrelated factors may affect quality of silage, including availability of water-soluble carbohydrates, the dry-matter content, the pH and extent of air exclusion. For example, lack of water-soluble carbohydrates may be overcome by wilting the material to raise the dry matter to a level at which less acid is required to stabilise the fermentation. The availability of sugars in the material and the rate at which the different micro-organisms multiply also influences the ensilage process.

Throughout this century, research workers have investigated ways through which the fermentation process in silage making can be controlled in order to improve the feeding quality of the resulting silage. Use of additives, to restrict the activity of the microorganisms, to stimulate the fermentation by the lactic acid bacteria or simply as nutrients has been one of the approaches. Additives used in the early studies included chloroform, toluene and cresol (to inhibit bacterial growth) and sulphuric acid and hydrochloric acid (to reduce the pH). Indeed, over the last 40 to 50 years, corrosive, acid-containing additives have been widely used in silage making. Other fermentation inhibitors which have been studied include organic acids, salts of acids, formaldehyde and other aldehydes, sodium hydroxide, and antibiotics. Of these, formic acid is probably the most widely studied and has been reported to have a beneficial effect on the fermentation process and on the nutritive value of silage. Sulphuric acid is cheaper than formic acid and is popular in some countries. However, acids are a hazard on the farm and can be particularly dangerous if recommended to uninformed farmers. Salts of acids are safer to handle but are less effective than the acids from which they are derived.

The hazardous nature of some of the chemical additives has necessitated a search for alternative compounds for improving the ensilage process. A group of compounds classified as fermentation stimulants have been widely studied. These include sugar sources (e.g. molasses and whey), enzymes and inocula of lactic acid bacteria. Molasses is of particular relevance to smallholder farmers in developing countries in the tropics where sugar-cane is produced and processed. Enzymes are essential for the breakdown of cell-wall carbohydrates to release the sugars necessary for the growth of the lactic acid bacteria. Although resident plant-enzymes and acid hydrolysis produce simple sugars from these carbohydrates, addition of enzymes derived from certain bacteria, e.g. *Aspergillus niger* or *Trichoderma viridi* increases the amount of available sugars. Commercial hemicellulase and cellulase enzyme cocktails are now available and improve the fermentation process considerably. However, prices of these products preclude their viability for farm level application, especially in developing countries.

There are two forms of indigenous lactic acid bacteria: the homofermentative type which converts hexose sugars to lactic acid with no loss of dry matter and the heterofermentative type which produces a range of compounds accompanied by loss of dry matter as carbon dioxide. Thus, the native bacteria are not the most efficient. Considerable research in the USA and Europe has been directed towards the development of microbial silage additives (inoculants). Commercial bacterial

inoculants designed to add sufficient homofermentative lactic acid bacteria to dominate the fermentation are now available. The objective of using such additives is to ensure the rapid production of the required amount of lactic acid from the carbohydrates present to preserve the ensiled material. Most such inoculants contain *Lactobacillus plantarum* with or without other bacteria such as *L. acidophilus*, *Pediococcus acidilactis* and *Streptococcus thermophilus*. In general, the results with bacterial inoculants have been quite variable. However, with an effective product, it is possible to improve the fermentation of low dry-matter silages and to enhance the efficiency of their utilization.

In order to improve the effectiveness of microbial inoculants in breaking down structural carbohydrates to glucose, detailed knowledge of the lactobacilli bacteria is essential. Work already undertaken on the molecular biology of *Lactobacillus plantarum* and other species (Armstrong and Gilbert 1991) suggest that the rapid progress in this area will make it possible to construct novel genes encoding highly active fibre-degrading enzymes. Such genes could then be inserted into strains of *L. plantarum*.

Successful silage making incorporating these technologies can only be achieved with strict adherence to recommended application procedures, including rates of additives, inoculants etc. This technology is available in most developing regions including Africa. However, it is not fully exploited. Indeed, in Africa silage making is still generally restricted to large-scale commercial farms.

4.5.4 Improving rumen function

Major areas of rumen function which might benefit from transgenic technology include: development of transgenic bacteria with enhanced cellulolytic activity, capability to cleave lignohemicellulose complexes, reduced methane production capability decreased proteolytic and/or deaminase activities, increased capability for nitrogen "fixation" and increased ability for microbial production of specific amino acids.

Genes encoding plant structural carbohydrate-degrading enzymes have been isolated from rumen bacteria. In contrast to conditions in which single species of organisms are grown in controlled environments and where the energy supply is usually in excess of demand, the rumen environment

is very complex, competition between different microbial species is intense and energy is usually the limiting growth factor. This is probably the main reason why reintroduction of genetically modified rumen bacteria into their natural habitat has met with variable success. Advances being made in transformation methods for obligate anaerobic bacteria will certainly result in successful genetic engineering of a range of rumen bacteria. However, it is not possible to predict if any of these bacteria will be capable of colonizing the rumen.

It can be seen that there are several potential opportunities for improving the efficiency of ruminant digestion and possibilities for utilizing a wider range of feeds than is currently possible. Modification of rumen microbial population is one such opportunity.

References

1. Denning C & Priddle H. (2003). - New frontiers in gene targeting and cloning: success, application and challenges in domestic animals and human embryonic stem cells. *Reproduction*, **126**, 1-11.
2. Gibson J.P. and Smith C. 1989. The incorporation of biotechnologies into animal breeding strategies. In: Babiuk L.A., Phillips J.P. and Moo-Young M. (eds), *Animal Biotechnology. Comprehensive Biotechnology First Supplement*. Pergamon Press, Oxford, UK. pp. 203-231.
3. Montagud A., E Navarro P Fernández de Córdoba and JF Urchueguía,
4. Madan M.L. 2005. Animal biotechnology: applications and economic implications in developing countries. *Rev. sci. tech. Off. int. Epiz.*, 2005, 24 (1). 127-139
5. Niemann H., Franzen H. and Smidt D. (eds), *Potential and Limitations of Biotechnology in Livestock Production in Developing Countries. Proceedings of a Symposium, Mariensee, Germany, May 1992*. Part I. ATSAF (Council for Tropical and Subtropical Agricultural Research), Bonn, Germany. pp. 68-80.
6. Niemann H., Halter R. and Paul D. 1994. Gene transfer in cattle and sheep: A summary perspective. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, Guelph, Canada, 7-12 August 1994. Volume 21*. International Committee for World Congress on Genetics Applied to Livestock Production, Guelph, Ontario, Canada. pp. 339-346.
7. Paterson A.H., Lander E.S., Hewitt J.D., Peterson S., Lincoln S.E. and Tanksley S.D. 1988. Resolution of quantitative traits into Mendelian factors, using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335:721-726.
8. Persley G.J. 1990. *Beyond Mendel's Garden: Biotechnology in the Service of World Agriculture*. CAB (Commonwealth Agricultural Bureaux) International, Wallingford, UK. 155 pp.

9. Rege , J.E.O. Biotechnology options for improving livestock production in developing countries, with special reference to sub-Saharan Africa. *International Livestock Centre for Africa* (ILCA)
P. O. Box 5689, Addis Ababa, Ethiopia.33p
10. Robinson J.J. and McEvoy T.G. 1993. Biotechnology - the possibilities. *Animal Production* 57:335-352.
11. Smith C. 1988. Genetic improvement of livestock in developing countries using nucleus breeding units. *World Animal Review* 6:2-10.
12. Smith C. 1988. Applications of embryo transfer in animal breeding. *Theriogenology* 29:203-212.
13. Smith C. 1989. Cloning and genetic improvement of beef cattle. *Animal Production* 49: 49-62.
14. Smith C., Meuwissen T.H.E. and Gibson J.P. 1987. On the use of transgenes in livestock improvement. *Animal Breeding Abstracts* 55:1-10.
15. Teale A.J. 1991. Biotechnologies applied to the diagnosis of disease, with special reference to trypanosomiasis and theileriosis of livestock. In: Hunter A.G. (ed), *Biotechnology in Livestock in Developing Countries*. Centre for Tropical Veterinary Medicine, University of Edinburgh, Edinburgh, UK. pp. 272-285.
16. Wray C. and Woodward M.J. 1990. Biotechnology and veterinary science: Production of veterinary vaccines. *Revue Scientifique et Technique de l'Office International des Epizooties* 9:779-794.

MODULE FIVE

Environmental and Biosafety Issues In Animal Biotechnology

5.1 Environmental concerns about biotechnology

There is much euphoria about developments in biotechnology and potential benefits, but little is said about the risks associated with biotechnology. For example genetically modified organisms could create ecological disaster if released into the environment.

Biosafety is, therefore, an issue of great concern for many developing countries. In a recent (June 1994) meeting of the Intergovernmental Committee on the Convention on Biological Diversity (CBD 1992), representatives of developing countries pointed out that biotechnology was evolving more rapidly than the capacity of their countries to install effective safety procedures for the handling and use of living modified organisms and that there was need for adequate and transparent safety procedures to manage and control the risks associated with the use and release of such organisms. To deal with the basic ethical questions and the risks associated with genetic engineering, regulatory mechanisms should be created and internationally acceptable guidelines or regulations put in place. The political and regulatory processes affecting biotechnology and its products must draw upon professional competence of the highest standard. In general, however, developed countries are lukewarm to the idea of a legally binding international protocol on biosafety, possibly because it is a heavy responsibility with potentially massive cost implications for the technology-rich countries. However, biosafety is an issue which must be addressed sooner than later.

5.2 SAFETY CONCERNS

Rapid advances in biology made since the structure of DNA was clarified provide techniques that have enhanced food production and improved human health. Advances are expected to continue and are likely to have an even greater impact in the future. However, the benefits of advanced technology rarely come without attendant hazards. The focus here is to identify the science based concerns related to modern, genetically-based animal biotechnology.

Biotechnology being a set of techniques by which living creatures are altered for the benefit of humans and other animals. Animal biotechnology has a long history, beginning as far back as 8,000 years ago with the domestication and artificial selection of animals. Rapid changes in animal production had been made in previous decades through procedures such as artificial selection, vaccination to enhance health, and artificial insemination to enhance reproduction. However, modern, genetically-based, biotechnology only began in the 1960s, following the discovery of the genetic code. In this report the committee moves beyond the scientific advances to examine new genetically based technologies.

New procedures involving direct gene insertion and manipulation allow for much more rapid selection of desirable traits than older procedures. The practice of biotechnology does not occur in the absence of the social, policy, and regulatory environments. In view of foregoing, five overarching concerns emerged.

1. The first was whether anything theoretically could go wrong with any of the technologies. For example, is it theoretically possible that a DNA sequence from a vector used for gene transfer could escape and unintentionally become integrated into the DNA of another organism and thereby create a hazard?
2. The second was whether the food and other products of animal biotechnology, whether genetically engineered, or from clones, are substantially different from those derived by more traditional, extant technologies.
3. A third major concern was whether the technologies result in novel environmental hazards.
4. The fourth concern was whether the technologies raise animal health and welfare issues.
5. Finally, there was concern as to whether ethical and policy aspects of this emerging technology have been adequately addressed. Are the statutory tools of the various government departments and agencies involved sufficiently defined? Are the technologic expertise and capacity within agencies sufficient to cope with the new technologies should they be deemed to pose a hazard?

Among these issues, the effects on the environment were considered to have the greatest potential

for long-term impact. The taxonomic groups that present the greatest environmental concerns are aquatic organisms and insects, because their mobility poses serious containment problems, and because unlike domestic farm birds and mammals, they easily can become feral and compete with indigenous populations.

There are a number of safety issues relating to biotechnology products that differ from those raised by low molecular weight products and need to be taken into account when designing the safety evaluation programme for a biotechnology derived pharmaceutical product (Table 5.1). The quality and consistency of the product requires careful control in terms of product identity, potency and purity because of concerns about microbiological safety, impurities arising from the manufacturing process (e.g. host-cell contaminants, endotoxin, residual DNA levels and process chemicals), and the fidelity of the protein sequence and post-translational modifications during process improvements and scale-up. The immunogenic nature of heterologous proteins, vectors, cells, tissues and process contaminants must also be considered in the design of the safety evaluation programme and appropriate monitoring for anti-product antibodies, particularly neutralizing antibodies included in toxicity studies to aid interpretation of the findings. For gene transfer products, there are concerns about the distribution and persistence of vector sequences, the potential for expression of vector sequences in non-target cells, tissues and, in particular, the potential for inadvertent gonadal distribution and germ-line integration. In 1997, the Food and Drug Administration (FDA) became aware that preclinical studies from multiple clinical trial applications indicated evidence of vector DNA in animal gonadal tissues following extragonadal administration. These positive polymerase chain reaction (pcr) signals were for DNA extracts from whole gonads subsequent to vector administration. The observations involved multiple classes of vectors, formulations and routes of administration. The following table summarizes key safety issues of biotechnology.

Table 5.1. Safety issues related to biotechnology

<i>Microbiological safety</i>
Adventitious agents bacterial, mycoplasma, fungal, viral, TESs, etc.
Potential for reconstitution of replication competent viral vectors for gene transfer products

Cells harbour homotropic and xenotropic viruses, e.g. retroviruses, Epstein-Barr virus, cytomegalovirus Potential for environmental spread, e.g., vector dissemination and viral shedding
<i>Immunological.</i>
Anti-product antibodies Host cell proteins or other process impurities Viral vectors for gene transfer products Cell impurities in cell therapy products
'Foreign' epitopes on cell/tissue/organ-based products Development of immune tolerance to DNA vaccines <i>Pharmacological</i>
Excessive pharmacology or unintentional receptor binding Distribution to and exposure of non-target tissues
Altered cell phenotype, cell products, function and localisation for cell therapies
<i>Biodistribution</i>
Distribution and persistence of gene and cell therapies Genetic transfer to germ cells for gene transfer products Insertional mutagenesis for gene transfer products Vector dissemination and viral shedding

5.3 Constraints on applying the technology

The application of new molecular biotechnologies and new breeding strategies to the livestock breeds used in smallholder production systems in developing countries is constrained by a number of factors. In the developing world, poverty, malnutrition, disease, poor hygiene and unemployment are widespread, and biotechnologies must be able to be applied in this context. Over the last few decades, the green revolution has brought comparative prosperity to farmers with land, but the majority of farmers, who are landless

or marginal farmers and subsist only on livestock, have been neglected and remain poor. The critical issues affecting livestock productivity have recently been re-examined. Research that aims to enhance productivity and sustainability should focus on improving. The major constraints on applying biotechnologies have been enumerated by Madan (39) and include:

- a)* the absence of an accurate and complete database on livestock and animal owners so that programmes can be implemented
- b)* the biodiversity present within species and breeds in agro-ecological systems
- c)* the fact that models of biotechnological intervention differ distinctly between developed and developing economies
- d)* the fact that many animal species and breeds are unique to the developing world; each has its own distinct developmental, production, disease resistance and nutrient utilisation characteristics
- e)* the lack of trained scientists, technicians and fieldworkers to develop and apply the technologies, both in the government and in the private sectors
- f)* the absence of an interface between industry, universities and institutions, which is necessary to translate technologies into products
- g)* the inability to access technologies from the developed world at an affordable price in order to make a rightful, positive and sustainable contribution to livestock production and the economic welfare of farmers
- h)* the high cost of technological inputs such as materials, biologicals and equipment
- i)* the failure to address issues of biosafety and to conduct risk analyses of new biologicals, gene products, transgenics and modified food items, and, above all
- j)* the negligible investment in animal biotechnology

5.4 Building capacity

Owing to the constraints outlined above, the economic benefits of animal biotechnology cannot be realised without a conscious, sustained, holistic, multi-stakeholder, participatory approach. There is a great need to ensure that capacity is not just created but also is retained and enhanced. Capacity-building activities must be carried out at all levels: the awareness of policy and decision makers must be raised, the necessary legal and regulatory frameworks must be initiated, the technical and regulatory capacities must be enhanced and institutions may need to be overhauled. More importantly, it is necessary to assess and deploy competent operators and institutional capacity continuously so that, as biotechnology advances, the procedures required for its safe use can be constantly evaluated, upgraded and applied. This is a daunting task, but it can be achieved through firm commitment and partnerships. Developing and commercializing improved technologies in most developing countries has been the responsibility of the public sector, and technology has been disseminated freely (51). This situation will have to continue if superior genetics, diagnostics and vaccines are to be delivered. However, research and almost all commercial development of biotechnology in the developed world are being driven largely by the private sector (52).

The global trends in funding for research and development and production do not address the concerns, needs and opportunities of the developing world. Developing countries are finding it increasingly expensive to access and use new technologies. There is limited private- and public-sector investment in animal health and production, particularly in relation to modern biotechnologies that are 'resource hungry'. Although several discoveries have been made in laboratories in the developing world, in most cases these have not been converted into useful technologies or products. The key potential users - resource-poor often illiterate farmers with a limited knowledge base - do not feel that applying these technologies is worth the effort, cost and risk involved. This is mainly because there is no agency or industry that can scale up and package the technology. Also, in the developed world, there is an economic incentive to market biotechnological services and products;

this is lacking in the developing world because of the limited purchasing power of resource-poor stakeholders. Research in biotechnology in recent years has also been motivated by economic considerations, and little research is conducted in the developing world because of the probable lack of returns on the investment. For understandable reasons, current funding policies in developing countries focus on areas that will yield practical benefit in the short term. In determining future policy, policy-makers and funding bodies must not lose sight of the substantial benefits that can be gained in the longer term by investing in strategic research into vaccine development.

Adequate multi-institutional (national and international) support through an international donor consortium is needed to develop cost-effective, cheap and easily adaptable biotechnological products. The amount spent by international agencies on animal biotechnology in developing countries is currently very low and constitutes only a small percentage of the total spending on agriculture. The World Bank, the Food and Agriculture Organization, the Consultative Group on International Agricultural Research, the United Nations Development Programme, the United States Agency for International Development, the Swedish International Development Cooperation Agency, the International Development Research Centre, the Asian Development Bank and other donor and funding agencies have to designate a higher percentage of funds to the livestock sector (39). It has been convincingly shown that investing in livestock has a dramatic and far-reaching impact on the human development index. This is a strong argument in control of disease) and the augmentation of feed resources. Adopting biotechnology has resulted in distinct benefits in terms of animal improvement and economic returns to the farmers. Over the past decade, the ILRI has focused on biotechnological applications, especially in Africa, and several developing countries now have multi-institutional programmes to develop and apply biotechnology. The developing world will have to respond to the many gene based technologies now being developed with a sense of commitment, trained manpower, infrastructure and funding.

References:

1. Animal Biotechnology: Science Based Concerns. ISBN: 0-309-50218-7,201 pages,6x9,

(2002) Copyright © National Academy of Sciences.

<http://www.nap.edu/catalog/10418.html>

2. Madan M.L. (2003). - Opportunities and constraints for using gene-based technologies in animal agriculture in developing countries and possible role of international donor agencies in promoting R&D in this field. In *FAO/IAEA* international symposium on applications of gene-based technologies for improving animal production and health in developing countries, Vienna, Austria, 6-10 October 2003. Food and Agriculture Organization/International Atomic Energy Agency, Vienna, 103-104
3. Pingali P.L. & Traxler G. (2002). - Changing locus of agricultural research: will the poor benefit biotechnology and privatization from trends? *Food Policy* 27, 223-238