6.3 Genomes

Human Genome Project (HGP) a collaborative worldwide project to sequence the nitrogenous bases in human DNA

In 1990 an ambitious project called the **Human Genome Project (HGP)** was launched. The goal of the HGP was to determine the complete 3 billion nitrogenous base sequence of human DNA. More than 2000 scientists from around the world collaborated to crack the code of the human DNA (**Figure 1**). Information about their research was exchanged among scientists by email, by articles in scientific journals, and by sharing when they met at conferences or visited each other's laboratories. In 1998, a private company, Celera Genomics, joined the race and, using a more computerized approach, began their own efforts to read the human genome.

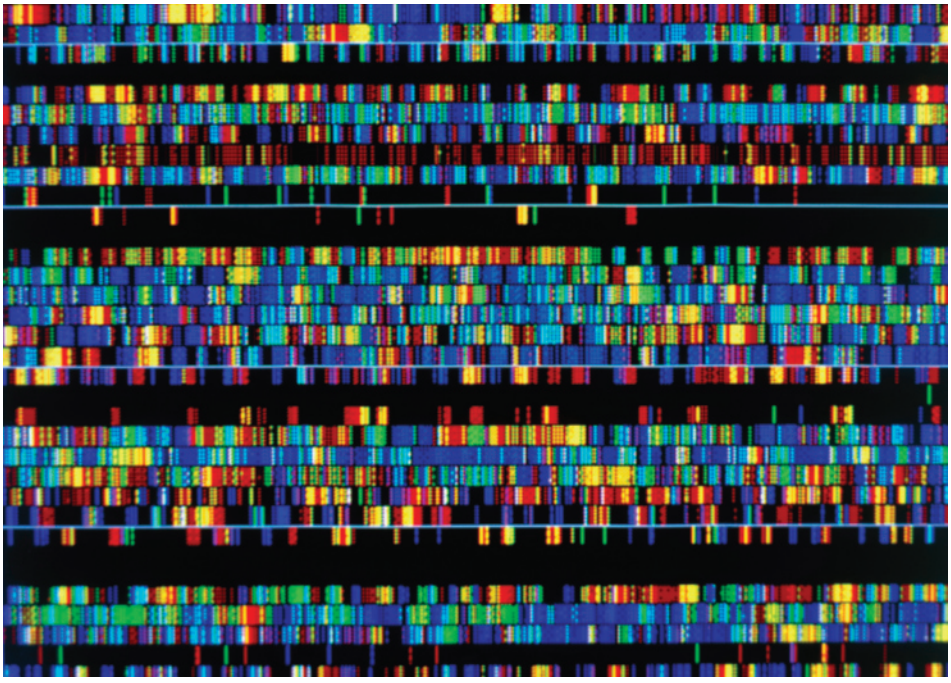

Figure 1 This computer screen displays a human DNA sequence as series of coloured bands. Each colour represents a specific base.

The Human Genome

In 2000, Celera Genomics and the HGP announced simultaneously that they had deciphered the human genome sequence, and each institution published a working draft in the prestigious journals *Science* and *Nature*. Three years later a more complete version was announced: 2.85 billion nucleotides had been sequenced.

The HGP team collected blood from females and sperm from males in a large sample of anonymous donors. Only a few samples were processed as DNA resources, and the source names remain confidential, so neither the donors nor scientists know whose DNA was sequenced. Celera collected samples from five donors who identified themselves only by race and sex.

human genome the sequence of DNA nitrogenous bases found on the 23 sets of chromosomes in humans

coding DNA a region of DNA that contains a sequence of nucleotides that will be expressed; a gene

non-coding DNA a region of DNA that does not contain a sequence of nucleotides that will be expressed

The **human genome** is made up of about 3 billion base pairs. A base pair is a nucleotide on one strand of the DNA molecule bonded weakly to its complementary nucleotide on the other strand of the DNA molecule. Not all base pairs encode genes, but the regions of the genome (or sequence of bases) that do code for a gene are called **coding DNA**. Coding DNA may make up as little as 2 % of the human genome. The remaining 98 % consists of highly repetitive DNA sequences called **non-coding DNA**, or "junk" DNA. A small portion of this non-coding DNA may serve important functions, but scientists are almost certain most of it does not.

Other Genomes

The largest genome that has been sequenced so far is a freshwater amoeboid that is made up of 670 billion base pairs. A large percentage of the DNA in this organism is non-coding DNA. The number of genes is not proportional to genome size—a larger genome does not necessarily mean a greater number of genes. At present, it is believed that humans have about 20 000 genes. **Table 1** shows a comparison of the number of genes in other organisms.

Table 1 Number of Genes Believed Present in Different Genomes

Organism	Size of genome	Approximate number of genes
amoeba	670 billion	unknown
newt (salamander)	84 billion	unknown
wheat	17 billion	unknown
human	3 billion	20 000
mouse	2.6 billion	25 000
Asian rice	446 million	50 000
fruit fly	137 million	13 602
yeast	12.1 million	6 034
E. coli	4.6 million	3 200
influenza virus	1.8 million	1 700

Functional Genomics

The HGP was the first step in understanding humans at the molecular level. Now that we know the sequence of nitrogenous bases, further research is being conducted to isolate genes and determine their function. Information gathered about genes and their functions is known as **functional genomics**.

functional genomics the study of the relationship between genes and their function

The mapping of the human genome has led to some medical advances. Genetic tests can predict such diverse matters as whether people with breast cancer need chemotherapy or whether individuals are at risk for eye disease. The information that the HGP has provided is invaluable as medical researchers search for the mutated genes that cause many genetic disorders. Yet researchers still have many questions. The human genome is complex. Scientists study other organisms that are genetically similar to humans. An organism that is used in place of another organism to study biological functions is a **model organism**. Surprisingly, mice are ideal model organisms for humans. Mice share many similarities to humans with respect to nitrogenous base sequence, gene location, and coding and non-coding regions.

model organism an organism that can be used to study biological functions of another organism, due to its genetic similarity

Mice and humans have roughly the same number of base sequences in their genome as other mammals, such as dogs, cats, rabbits, and monkeys (**Figure 2**). Mice and humans also share a similar number and type of genes, and most genes found in humans also exist in mice. Variation between the gene products may exist, of course, but the type of gene is similar. For example, both mice and humans have genes that code for hemoglobin, but mouse hemoglobin differs slightly from human hemoglobin. However, it is obvious that the interaction of the genes must differ, since phenotypically mice are significantly different from humans.

Figure 2 Mammals contain roughly the same number of nitrogenous bases—3 billion—in their genome and about the same number of genes.

Knowing that a single change in the nitrogenous base sequence can result in disorders such as cystic fibrosis, sickle-cell disease, and breast cancer, it is not surprising that there is not much difference between the genomes of humans and apes. (**Figure 3**).

Figure 3 There is a 95 % to 98 % similarity between the genomes of humans and apes.

CAREER **LINK**

Bioethicist

Bioethicists examine the controversies and concerns that arise as advances in science and technology are applied in fields such as medicine, genetics, and biotechnology. To learn more about being a bioethicist,

GO TO NELSON SCIENCE

DNA Identification

When the Human Genome Project was in the planning stages, valid questions about the implications of the project arose. Who would have access to an individual's genome? Would this lead to mandatory DNA sampling in the future when applying for employment? University? College? Would governments begin to take DNA samples from all new infants for storage and possible use?

DNA information is, for the most part, still personal and private (unless you are suspected of performing criminal acts). DNA information is not required for making any of the above decisions. Bioethicists provide information to society so that informed decisions may be made on difficult questions.

People can now have their DNA screened for numerous genetic disorders, such as Huntington's disease and cystic fibrosis. Many private companies offer genetic testing and DNA banking. When DNA is stored in a **DNA bank**, it is not studied or sequenced; the DNA is simply frozen and stored for future use. To use a DNA bank, a person must supply a blood sample. The DNA is extracted from the blood and then stored in a freezer for an amount of time depending on company policy.

DNA bank a database of DNA sequences; the sequences can be from plants, animals, or humans

DNA banks also store the DNA of endangered species. Scientists are not able to bring back extinct species with current technology, but no one is sure what the future holds. So far, geneticists have sequenced fragments of DNA from extinct animals and have used bacteria to produce the protein that the fragment codes for. As you will learn in the next unit, geneticists have even sequenced the entire genome of a Neanderthal—an extinct hominid that lived more than 30 000 years ago. Some DNA banks store DNA samples of plants and fish. Researchers have access to the information in these banks that may be used to develop more viable agricultural species.

DNA Fingerprinting

In order to compare the DNA of a suspect to the DNA at a crime scene, a technique known as DNA fingerprinting is used. **DNA fingerprinting** produces a pattern of bands on a gel that is characteristic of each individual but highly variable from person to person. Forensic detectives perform multiple tests and compare many bands from each sample. If every band on the suspect's gels match every band from the crime scene sample, then it is highly likely that the suspect was at the crime scene (**Figure 4**).

DNA fingerprinting a pattern of bands on a gel that is unique to each individual

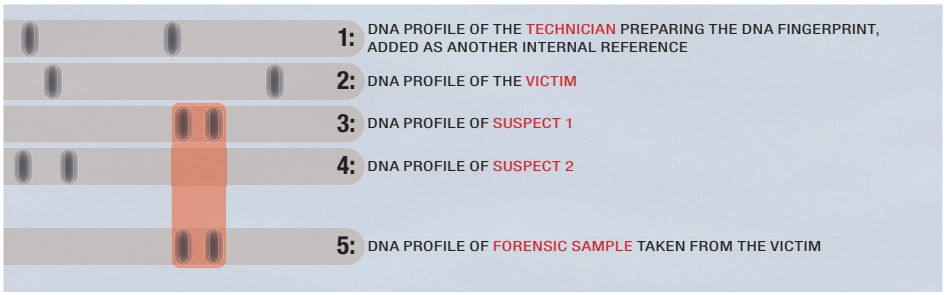

Figure 4 The DNA fingerprint of suspect 1 matches the DNA fingerprint of the sample (possibly blood, semen, or hair) that was found at the crime scene. Forensic detectives make multiple gels and often compare 10 or more bands.

DNA fingerprinting is used to illustrate innocence as well as guilt. For example, David Milgaard spent more than 20 years in prison in Saskatchewan before DNA fingerprinting technology was available. When the DNA evidence provided by Milgaard was compared to stored crime scene samples, it proved that the DNA at the crime scene was not his. In this case, DNA evidence showed that it was impossible for Milgaard to have committed the crime, and he was released from prison in 1999, having been wrongfully convicted.

DNA fingerprinting may be used to settle paternity suits, determine ancestry, locate long-lost relatives, or identify remains, such as those of soldiers in unmarked graves. Currently, paternity suits are the most common application of DNA fingerprinting. DNA fingerprinting proves the genetic relationship between children and their biological parents, siblings, or even grandparents.

DNA fingerprinting applications are very widespread, including some in consumer affairs. Recently, researchers at the University of Guelph found that many fish products labelled as one type of fish were in fact another type of fish. Wildlife enforcement officers use DNA fingerprinting to prosecute hunters or fishers who hunt or fish certain animals out of season, or endangered species. Animal meat that has been packaged and stored in a freezer can be difficult to identify. DNA fingerprinting can reveal the true source.

National DNA Data Bank

In Canada, it is legal for a judge to request blood or hair samples from suspects in criminal cases and convicted criminals to be added to Canada's DNA Data Bank. The bank holds the DNA sequence of thousands of convicted criminals. The information in the bank allows the police to

- link crimes together when there are no suspects
- help identify suspects
- eliminate suspects when there is no match between the crime scene DNA and the suspect's DNA
- determine whether a serial offender is involved

Many cases have been solved, and criminals have been prosecuted, with the use of forensic DNA evidence. In addition, innocent people accused of a crime have been exonerated by DNA evidence.

WEB **LINK**

Using DNA Data to Solve Crimes

To read about how Canada's DNA bank has been used to solve crimes,

GO TO NELSON SCIENCE

Investigation 6.3.1

Comparing DNA Fingerprints (page 254)

Now that you have learned about DNA fingerprinting, you can complete Investigation 6.3.1.

In this investigation you will compare two samples of DNA to a known sample.

6.3 Summary

- The Human Genome Project and Celera Genomics successfully sequenced the 3 billion base pairs in the human genome.
- Different genomes contain different numbers of genes. The human genome contains approximately 20 000 genes.
- The mouse genome is very similar to the human genome in gene function. Mice are good model organisms for human genetic studies.
- DNA banks can be used to store DNA for future use. DNA data banks also store the DNA of endangered species and plants, as well as fish.
- DNA fingerprinting is a technique that may link a suspect to a crime scene.
- Forensic DNA evidence is frequently used in Canadian criminal trials and even in fish and wildlife enforcement.
- The National DNA Data Bank in Canada stores the DNA of convicted criminals.

6.3 Questions

1. Differentiate between the human genome and the Human Genome Project. K/U T/I
2. When did the Human Genome Project commence? When was it completed? K/U
3. Compare the mouse genome to the human genome. How are they similar? How are they different? K/U T/I
4. Why are mice a good model for studying human genes? K/U
5. Using the Internet and other sources, research the term "knockout mice." Use a flowchart to explain how knockout mice are used in gene research. T/I C
6. What is a DNA data bank? List the different ways a data bank may be used. K/U
7. What is DNA fingerprinting? Why is it an effective forensic tool? K/U
8. The DNA fingerprint shown in **Figure 5** was produced after a crime had been committed. Which of the suspects could have committed the crime? T/I

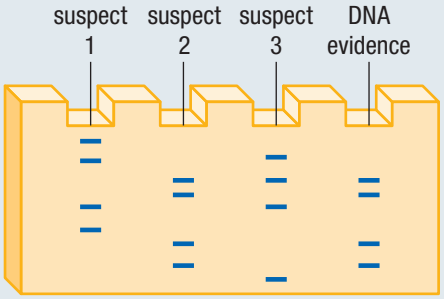

Figure 5

9. Use the Internet and other sources to research a Canadian case where DNA evidence was used to exonerate or convict a suspect in a human crime or a fish and wildlife crime. T/I A
10. Use the Internet and other sources to research the number of species for which entire genomes have been sequenced. T/I A
    (a) List at least eight species that you researched.
    (b) Comment on the types of organisms that have been sequenced. Suggest reasons as to why these types of species may have chosen for genome sequencing.
11. Visit the website of the Human Genome Project (HGP). Read the website, and conduct research to answer to the following questions. T/I A
    (a) Do all chromosomes have the same number of genes?
    (b) Do genes seems to be spread out evenly along each chromosome?
    (c) How was the HGP funded?
    (d) Who has access to the information?

GO TO NELSON SCIENCE