1 Introduction

The classical approach to genetics starts with the identification of variants which have a specific *phenotype*, i.e. they are altered in some way that can be seen (or detected in other ways) and defined. For Mendel, this was the appearance of his peas (e.g. green *vs* yellow, or round *vs* wrinkled). One of the postulates he arrived at was that these characteristics assorted independently of one another. For example, if you cross a strain that produces yellow round peas with another strain that produces green wrinkled peas, the first generation (F_1) are all round and yellow (because round is dominant over wrinkled, and yellow is dominant over green). Of course Mendel did not know why this happened. We now know that if two genes are located on different chromosomes, which will segregate independently during meiosis, the genes will thus be distributed independently amongst the progeny. The same can happen if the two genes are on the same chromosome, but are so far apart that the recombination between the homologous chromosomes will be sufficient to reassort them independently. On the other hand, if they are quite close together, they will tend to remain associated during meiosis, and will therefore be inherited together. We refer to genes that do *not* segregate independently as *linked*; the closer they are, the greater the degree of linkage, i.e. the more likely they are to stay together during meiosis. Measuring the degree of linkage (*linkage analysis*) is a central tool in classical genetics, in that it provides a way of mapping genes, i.e. determining their relative position on the chromosome. Furthermore, it provides us with an important method for correlating genetic and physical maps.

Bacteria and yeasts provide much more convenient systems for genetic analysis, because they grow quickly, as unicellular organisms, and on defined media. You can therefore use chemical or physical mutagens (such as ultraviolet irradiation) to produce a wide range of mutations, and can select

From Genes to Genomes, Second Edition, Jeremy W. Dale and Malcolm von Schantz

specific mutations from very large pools of organisms – remembering that an overnight culture of *E. coli* will contain some 10^9 bacteria per millilitre. These mutations may simply affect the ability to produce a specific amino acid, manifested as a requirement for that amino acid to be added to the growth medium, or to use a particular carbon source such as lactose. Or it may be a more complex phenotype, such as loss of motility, or inability to divide into two cells, leading to the production of filaments. Therefore we can use genetic techniques to investigate detailed aspects of the physiology of such cells, including identifying the relevant genes by mapping the position of the mutations. Although the techniques in bacteria differ from those in higher organisms, forms of linkage analysis still play a major role.

For multicellular organisms, the range of phenotypes is even greater, as there are then questions concerning the development of different parts of the organism. However animals have much longer generation times than bacteria, and using millions of animals (especially mammals) to identify the mutations you are interested in is either cumbersome, impossible or indefensible. Human genetics is even more difficult as you cannot use selected breeding to map genes; you have to rely on the analysis of real families who have chosen to breed with no consideration for the needs of science. Nevertheless, classical genetics has contributed extensively to the study of developmental processes, notably in the fruit fly *Drosophila melanogaster*, where it is possible to study quite large numbers (although nothing like the numbers that can be used in bacterial genetics), and to use mutagenic agents to enhance the rate of variation.

These methods suffered from a number of limitations. In particular, they could only be applied, in general, to mutations that gave rise to a phenotype that could be defined by morphological, physiological, biochemical or behavioural means. Furthermore, there was no easy way of characterizing the nature of the mutation. The situation changed radically in the 1970s, with the development of techniques that enabled DNA to be cut precisely into specific fragments, and to be joined together, enzymatically – techniques that became known variously as genetic manipulation, genetic modification, genetic engineering or recombinant DNA technology. The term ‘gene cloning’ is also used, since joining a fragment of DNA with a vector such as a plasmid that can replicate in bacterial cells enabled the production of a bacterial strain (a clone) in which all the cells contained a copy of this specific piece of DNA. For the first time, it was possible to isolate and study specific genes. Since this applied equally to human genes, the impact on human genetics was particularly marked.

The revolution also depended on the development of a variety of other molecular techniques. The earliest of these (actually predating gene cloning) was *hybridization* which enabled the identification of specific DNA sequences on the basis of their sequence similarity. Later on came methods for deter-

mining the sequence of these DNA fragments, and the polymerase chain reaction (PCR), which provided a powerful way of amplifying specific DNA sequences. Combining those advances with automated techniques and the concurrent advance in computer power has led to the determination of the full genome sequence of many organisms, including the human genome, as well as enormous advances in understanding the roles of genes and their products.

This revolution does not end with understanding how genes work and how the information is inherited. Genetics, and especially modern molecular genetics, underpins all the biological sciences. By studying, and manipulating, specific genes, we develop our understanding of the way in which the products of those genes interact to give rise to the properties of the organism itself. This could range from, for example, the mechanism of motility in bacteria to the causes of human genetic diseases and the processes that cause a cell to grow uncontrollably, giving rise to a tumour. In many cases, we can identify precisely the cause of a specific property. We can say that a change in one single base in the genome of a bacterium will make it resistant to a certain antibiotic, or that a change in one base in human DNA could cause debilitating disease. Understanding the cause of a genetically determined disease leads firstly to genetic diagnosis, and ultimately to remedying it, using gene therapy.

Furthermore, since these techniques enabled the cloning and expression of genes from any one organism (including humans) into a more amenable host, such as a bacterium, they allowed the use of genetically modified bacteria (or other hosts) for the production of human gene products, such as hormones, for therapeutic use. This principle was subsequently extended to the genetic modification of plants and animals – both by inserting foreign genes and by knocking out existing ones – to produce plants and animals with novel properties.

As you are probably aware, the construction and use of genetically modified organisms (GMOs) is not without controversy. In the early days, *E. coli* strains carrying recombinant DNA molecules were treated with extreme caution. *E. coli* is a bacterium which lives in its billions within our digestive system, and those of other mammals, and which will survive quite easily in our environment, unfortunately including our food and our beaches. So there was a lot of concern that the introduction of foreign DNA into *E. coli* would generate bacteria with dangerous properties. Fortunately, this is one fear that has been shown to be unfounded. Some natural *E. coli* strains *are* pathogenic – in particular the O157:H7 strain, which can cause severe disease or death. In contrast, the strains used for genetic manipulation are harmless disabled laboratory strains that will not even survive in the gut. Working with genetically modified *E. coli* can therefore be done very safely (although work with *any* bacterium has to follow some basic safety rules). However, the most

commonly used type of vector, plasmids, is shared readily between bacteria; the transmission of plasmids between bacteria is behind much of the natural spread of antibiotic resistance. What if our recombinant plasmids were transmitted to other *bacterial* strains that *do* survive on their own? This, too, has turned out not to be a worry in the majority of cases. The plasmids themselves have been manipulated so that they cannot be readily transferred to other bacteria. Furthermore, expressing a gene such as that coding for, say, dogfish insulin, or carrying an artificial chromosome containing 100 000 bases of human genomic DNA, is a great burden to an *E. coli* cell, and confers no reward whatsoever. In order to make them accept it, we have to create conditions that will kill all bacterial cells *not* carrying the foreign gene. Although we have to recognize the possibility that some unscrupulous individual might evade the regulations and produce a harmful bacterium, in practice it is not that easy. In general, genetically modified bacteria are not well able to cope with life outside the laboratory, and nature is quite capable of producing pathogenic organisms without our assistance.

The debate has now largely moved on to issues relating to genetically modified plants and animals. It is important to distinguish the *genetic modification* of plants and animals from *cloning* plants and animals. The latter simply involves the production of genetically identical individuals; it does not involve any genetic modification whatsoever. (The two technologies can be used in tandem, but that is another matter.) There are ethical issues to be considered, but cloning plants and animals is not the subject of this book.

The debate largely revolves around two factors: food safety and environmental impact. The first thing to be clear about is that there is no imaginable reason why genetic modification, *per se*, should make a foodstuff hazardous in any way. There is no reason to suppose that cheese made with rennet from a genetically modified bacterium is any more dangerous than similar cheese made with 'natural' rennet. It is possible to imagine a risk associated with some genetically modified foodstuffs, due to unintended stimulation of the production of natural toxins – remembering for example that potatoes are related to deadly nightshade. However, this can happen equally well (or perhaps is even more likely) with conventional procedures for developing new strains, which are not subject to the same degree of rigorous testing for safety.

The potential environmental impact is more difficult to assess. The main issue here is the use of genetic modification to make plants resistant to herbicides or to insect attack. When such plants are grown on a large scale, it is difficult to be certain that the gene in question will not spread to related wild plants in the vicinity (although measures can be taken to reduce this possibility). However, this may be an exaggerated concern. As with the bacterial example above, these genes will not spread significantly unless there is an evolutionary pressure favouring them. We would not expect widespread

resistance to weedkillers unless the plants are being sprayed with those weedkillers. There might be an advantage in becoming resistant to insect attack, but the insects concerned have been around for a long time, so the wild plants have had plenty of time to develop natural resistance anyway. We also have to balance the use of genetically modified plants against the use of chemicals. If genetic modification of the plants means a reduction in the use of environmentally damaging chemicals, then that must be a good thing.

Nonetheless, this issue is by no means as clear-cut as that of genetically modified bacteria. We cannot test these organisms in a contained laboratory. They take months or years to produce each generation, not 20 min as *E. coli* does. Thus, this is an important and complicated issue, and to understand it fully you need to know about evolution, ecology, food chemistry, nutrition and molecular biology. We hope that reading this book will be of some help for the last of these. We also hope that it will convey some of the wonder, excitement, and intellectual stimulation that this science brings to its practitioners. What better way to reverse the boredom of a long journey than to indulge in the immense satisfaction of constructing a clever new screening algorithm? Who needs jigsaw and crossword puzzles when you can figure out a clever way of joining two DNA fragments together? And how can you ever lose the fascination you feel about the fact that the drop of enzyme that you are adding to your test tube is about to manipulate the DNA molecules in it with surgical precision?