## Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Integrative ecological genomics</td>
<td>257</td>
</tr>
<tr>
<td>7.1</td>
<td>The need for integration: systems biology</td>
<td>257</td>
</tr>
<tr>
<td>7.2</td>
<td>Ecological control analysis</td>
<td>263</td>
</tr>
<tr>
<td>7.3</td>
<td>Outlook</td>
<td>266</td>
</tr>
</tbody>
</table>

References 271

Index 299

---

### CHAPTER 1

**What is ecological genomics?**

We define ecological genomics as a scientific discipline that studies the structure and functioning of a genome with the aim of understanding the relationship between the organism and its biotic and abiotic environments.

With this book, we hope to contribute to this new discipline by summarizing the developments over the last 5 years and explaining the general principles of genomics technology and its application to ecology. Using examples drawn from the scattered literature, we indicate where ecological questions can be answered, reformulated, or solved by means of genomics approaches. This first chapter introduces the main purpose of ecological genomics. We describe its characteristics, its interactions with other disciplines, and its fascination with model species. We also touch on some of its possible applications.

#### 1.1 The genomics revolution invading ecology

The twentieth century has been called the 'century of the gene' (Fox Keller 2000). It began with the rediscovery in 1953 of the laws of inheritance by Drs. Correns and von Tschermak, laws that had been formulated about 40 years earlier by Gregor Mendel. With the appearance of the Royal Horticultural Society's English translation of Mendel's papers, William Bateson suggested in a letter in 1902 that this new area of biology be called genomics. The word gene followed, coined by Wilhelm Ludwig Johannsen in 1909, and then in 1920 the German botanist Hans Winkler proposed the word genome. The term genomics did not appear until the mid-1980s and was introduced in 1987 as the name of a new journal (McKusick and Ruddle 1987). The century ended with the genomics revolution, culminating in the announcement of the completion of a draft version of the human genome in the year 2000.

Realizing the importance of Mendel's papers, William Bateson announced that genetics was to become the most promising research area of the life sciences. One hundred years later one cannot avoid the conclusion that the progress in understanding the role of genes in living systems indeed has been astonishing. The genomics revolution has now expanded beyond genetics, its impact being felt in many other areas of the life sciences, including ecology. In the ecological arena, the intersection between genomics and ecology has led to a new field of research, evolutionary and ecological functional genomics. Foil and Mitchell-Olde (2003) indicated that this new multidisciplinary 'focuses on the genes that affect evolutionary fitness in natural environments and populations'.

Our definition of ecological genomics given above seems at first sight to include the basic aim of ecology, viewing genomics as a new tool for analyzing fundamental ecological questions. However, the merging of genomics with ecology includes more than the incorporation of a toolbox, because with the new technology new scientific questions emerge and existing questions can be answered in a way that was not considered before. We expect therefore that ecological genomics will develop into a truly new discipline, and will forge a mechanistic basis for ecology that is often felt to be missing. This could also strengthen the relationship between ecology and the other life
sciences, because to a certain extent ecological genomics speaks the same language and read the same papers as molecular biologists.

Fig. 1.1 illustrates the various fields from which ecological genomics draws and upon which it is still growing. First of all, as indicated by Feder and Mitchell-Olend (2003), ecological genomics is closely linked to evolutionary biology and the associated disciplines of population genetics and evolutionary ecology. Another major area supporting ecological genomics is plant and animal physiology, which have their base in biochemistry and cell biology. A special position is held by microbial ecology, the meeting place of microbiology and ecology, where the use of genomics approaches has proceeded further than in any other subdiscipline of ecology. We consider genomics itself as a mainly technological advance, supporting ecological genomics in the same way as it supports other areas of the life sciences, such as medicine, neurobiology, and agriculture.

The genomics revolution is not only due to advances in molecular biology. Three major technological developments that took place in the 1990s also made it possible: microtechnology, computing, and communication.

**Microtechnology.** The possibility of working with molecules on the scale of a few micrometres, given by advances in laser technology, has been very important for one of genomics’ most conspicuous achievements, the development of the gene chip.

**Computing technology.** To assemble a genome from a series of sequences requires tremendous computational power. Extensive calculations are also necessary for the analysis of expression matrices and protein databases. Without the advent of high-speed computers and data-storage systems of vast capacity all this would have been impossible.

**Communication technology.** Consulting genome databases all over the world has become such a normal practice that the scientific progress of any genomics laboratory has become completely dependent on communication with the rest of the World Wide Web. The Internet has become an indispensable part of genomics.

The essence of genomics is that it is the study of the genome and its products as a unitary whole. In biology, the suffix -omes signifies the collectivity of units (Lederberg and McCray 2001), as for example in coelom, the system of body cavities, and biome, the entire community of plants and animals in a climatic region. In aiming to investigate many genes at the same time genomics differs from ecology, which although investigating many phenotypes, usually deals with only a few genes at a time (Fig. 1.2). Ecological genomics borrows from these two extremes, investigating phenotypic

---

**Figure 1.1** The position of ecological genomics in the middle of the other life-science disciplines with which it interacts most intensively.

**Figure 1.2** The playing field of ecological genomics, in between genomics, with its focus on the single genome of a model organism, studying all the genes that it contains, and ecology, studying a few genes in many species.
sciences, because to a certain extent ecological genomics spans the same language and read the same papers as molecular biologists.

Fig. 1.1 illustrates the various fields from which ecological genomics draws and upon which it is still growing. First of all, as indicated by Feder and Mitchell-Odke (2003), ecological genomics is closely linked to evolutionary biology and the associated disciplines of population genetics and evolutionary ecology. Another major area supporting ecological genomics is plant and animal physiology, which have their base in biochemistry and cell biology. A special position is held by microbial ecology, the meeting place of microbiology and ecology, where the use of genomics approaches has proceeded further than in any other subdiscipline of ecology. We consider genomics itself as a main technological advance, supporting ecological genomics in the same way as it supports other areas of the life sciences, such as medicine, neurobiology, and agriculture.

The genomics revolution is not only due to advances in molecular biology. Three major technological developments that took place in the 1990s also made it possible: microbiology, computing, and communication.

Microbiology. The possibility of working with molecules on the scale of a few micrometres, given advances in laser technology, has been very important for one of genomics’ most conspicuous achievements, the development of the gene chip.

Computing technology. To assemble a genome from a series of sequences requires tremendous computational power. Extensive calculations are also necessary for the analysis of expression matrices and protein databases. Without the advent of high-speed computers and data-storage systems of vast capacity all this would have been impossible.

Communication technology. Consulting genome databases all over the world has become such normal practice that the scientific progress of any genomics laboratory has become completely dependent on communication with the rest of the World Wide Web. The Internet has become an indispensable part of genomics.

The essence of genomics is that it is the study of the genome and its products as a unitary whole. In biology, the suffix -ome signifies the collectivity of units (Lederberg and McCray, 2001), as for example in cotsyme, the system of body cavities, and biome, the entire community of plants and animals in a climatic region. In aiming to investigate many genes at the same time genomics differs from ecology, which although investigating many phenotypes, usually deals with only a few genes at a time (Fig. 1.2). Ecological genomics borrows from these two extremes, investigating phenotypic biodiversity as well as diversity in the genome. With this new discipline, ecology is enriched by genomics technology and genomics is enriched by ecological questioning and evolutionary views. Because genomics analyses the genome in its entirety, it transcends classical genetics, which studies genes one by one, relating DNA sequences to proteins and ultimately to heritable traits. Genomics is based on the observation that the impact of one gene on the phenotype can only be understood in the context of the expression of several other genes or, in fact, of all other genes in the genome, plus their products, metabolites, cell structures, and all the interactions between them. This is not to say that every study in genomics deals with everything all the time, but that the mind is set and tools are deployed to maximize awareness of any effects elsewhere in the genome, outside the system under study. Consequently, genomics is invariably associated with unexpected findings. The discovery aspect of genomics is expressed aptly in a public-education project of Genome Canada entitled The GEE! in Genome (www.genomencanada.ca).

The work of Spellman and Rubin (2002) and their discovery of transcriptional territories in the genome of the fruit fly, Drosophila melanogaster, is an example of how the genomics approach can fundamentally alter our way of thinking about the relationship between genes and the environment (see also Weitzman 2002). The authors carried out transcription profiling with DNA microarrays (see Section 2.3) to investigate the expression of almost all of the genes in the fruit fly’s genome under 88 different environmental conditions. Their work was in fact a meta-analysis of transcription profiles collected earlier in six separate investigations. Because the complete genome sequence of Drosophil us is known, it was possible to trace every differentially expressed gene back to its chromosomal position. They concluded that genes physically adjacent in the genome often had similar expression when comparing different environmental challenges. The window of correlated expression appeared to extend to 10 or more adjacent genes and they estimated that 20% of the genome was organized in such ‘expression clusters’. Most astonishingly, genes in one cluster proved to be no more similar in structure or function than could be expected from a random arrangement. Spellman and Rubin (2002) suggested that local changes in chromatin structure trigger the expression of large groups of genes together. Thus a gene may be expressed not because there is a particular need for its product, but because its neighbour is expressed for a reason completely unrelated to the function of the first gene. At the moment it is not known whether such mechanisms lead to unexpected correlations between phenotypic traits, but surely the discovery of transcriptional territories could never have been made on a gene-by-gene basis, and this is due to the genomics approach.

The interactions between the genes within the genome and the dynamic character of the genome on an evolutionary scale have been sketched vividly by Dover (1999) as an internal tangled bank. This idea goes back to Darwin (1859) who, after investigating the banks of hollow roads in the English countryside, was intrigued by the great variety of organisms tangled together.

It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the boughs, with various insects darting about, and with worms crawling through the damp earth.

Darwin considered the way in which all organisms depended on each other as the template for evolution. Inspired by Darwin, Dover (1999) made a distinction between the ‘external tangled bank’ (the ecology) and the ‘internal tangled bank’ (the genome), attributing to them complementary roles in the evolutionary process (Fig. 1.3). The concept of the internal tangled bank emphasizes the role of genetic turbulence (gene duplication, genetic sweeps, exon shuffling, transposition, etc.) in the genome and it illustrates that there is ample scope for innovation from within’. These innovations are then checked against the external tangled bank, and this constitutes the process of evolution. This agrees with François Jacob’s famous description of ‘evolution through tinkering’ (Jacob, 1977). It should not surprise us that genetic turbulence leaves many traces in the genome that do not have
direct negative phenotypic consequences; these traces from the past provide a valuable historical record for genome investigators to discover.

1.2 Yeast, fly, worm, and weed

A striking feature of genomics is its focus on a limited number of model species with fully sequenced genomes and large research networks organized around them. The genomes of these model species have been sequenced completely and the information is shared on the Internet, allowing scientists to take maximal advantage of progress made by others. This explains the extreme speed with which the field is developing. Ecology does not have a strong tradition in standardized experimentation with one species. Thus the genomics approach is all the more striking to an ecologist, who is often more fascinated by the diversity of life than by a single organism, and engaged in a very wide variety of topics, systems, and approaches. In this section we examine the arguments for introducing model species in ecological genomics.

The best-known completely sequenced genomes, in addition to those of mouse and human, are those of the yeast Saccharomyces cerevisiae, the ‘fly’ Drosophila melanogaster, the ‘worm’ Caenorhabditis elegans and the ‘weed’ Arabidopsis thaliana. Investigations into the genomes of these model organisms are supported by extensive databases on the Internet that provide a wealth of information about genome maps, genomic sequences, annotated genes, allelic variants, cDNAs, and expressed sequence tags (ESTs), as well as news, upcoming events, and publications. These four model genomes and their relationships with evolutionary related species will be discussed in more detail in Chapter 3. The genomics of the mouse and human are not discussed at length in this book because the model status of these two species has mainly a medical relevance.

The first genome to be sequenced completely was that of Haemophilus influenzae (Fleischmann et al. 1995). This bacterium is associated with influenza outbreaks, but is not the cause of the disease, which is a virus. Although several years earlier the ‘genome’ of bacteriophage ΦX174 had
been sequenced (Sanger 1977a), 1995 is considered by many as the true beginning of genomics as a science, not in the least because the *H. influenzae* project demonstrated the usefulness of a new strategy of sequencing and assembly (whole-genome shotgun sequencing; see Chapter 2). With 1.8 Mbp the genome of *H. influenzae* was about 10 times larger than that of any virus sequenced before, but still two to four orders of magnitude smaller than the genome of most eukaryotes. Genome sequences of many other prokaryotes soon followed, including that of *Methanococcus jannaschii*, a archaeon living at a depth of 2,600 m near a hydrothermal vent on the floor of the Pacific Ocean (Bult et al. 1996). The genome of this extemophile was interesting because of the many genes that were completely unknown before. In 1999, a large network of scientists embarked on a project for sequencing the yeast genome, which was completed in 1996 and was the first eukaryotic genome to be elucidated (Goffeau et al. 1996). Thus, by 1996, the first genomic comparisons were possible between the three domains of life: Bacteria, Archives, and Eucarya.

The International Human Genome Project initiated by the US National Institutes of Health and the US Department of Energy, was launched in 1990 with completion due in 2005. However, in the meantime a private enterprise, Celera Genomics, embarked on a project with the same aim but a different approach and actually overtook the Human Genome Project. The competition was settled with the historic press conference on 26 June 2000, when US President Bill Clinton, J. Craig Venter of Celera Genomics, and Francis Collins of the National Institutes of Health jointly announced that a working draft of the human genome had been completed (Fig. 1.4). Many commentators have...
Table 1.1 List of complete and published genomes (not including viruses) by June 2005

<table>
<thead>
<tr>
<th>Taxonomic group</th>
<th>No. of genomes</th>
<th>Remarks on species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria total</td>
<td>211</td>
<td>Many common laboratory models and pathogens</td>
</tr>
<tr>
<td>Archaea total</td>
<td>21</td>
<td>Several methanogens and extremophiles</td>
</tr>
<tr>
<td>Eukarya*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myxomycota</td>
<td>1</td>
<td>Dictyostelium discoideum (slime mould)</td>
</tr>
<tr>
<td>Entamoeba</td>
<td>1</td>
<td>Entamoeba histolytica (amoeba causing dysentery)</td>
</tr>
<tr>
<td>Apicomplexa</td>
<td>6</td>
<td>Four Plasmodium and two Microsporidium species</td>
</tr>
<tr>
<td>Kinetoplastida</td>
<td>2</td>
<td>Trypanosoma brucei, Leishmania tropica (parasites)</td>
</tr>
<tr>
<td>Cryptomonadina</td>
<td>1</td>
<td>Guillardia theta (flagellated unicellular algal)</td>
</tr>
<tr>
<td>Bacillariophyta</td>
<td>1</td>
<td>Thalassiosira pseudonana (marine diatom)</td>
</tr>
<tr>
<td>Rhodophyta</td>
<td>1</td>
<td>Cyanidioschyzon merolae (small unicellular red algal)</td>
</tr>
<tr>
<td>Plants</td>
<td>4</td>
<td>Chlamydomonas reinhardtii (green alga), Populus trichocarpa (black cottonwood), Arabidopsis thaliana (thale cress), Oryza sativa var. japonica, var. indica (rice)</td>
</tr>
<tr>
<td>Fungi</td>
<td>14</td>
<td>Including Saccharomyces cerevisiae (baker’s yeast)</td>
</tr>
<tr>
<td>Animals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nematoda</td>
<td>2</td>
<td>Caenorhabditis elegans (free-living roundworm), Caenorhabditis briggsae</td>
</tr>
<tr>
<td>Insecta</td>
<td>4</td>
<td>Bombyx mori (silkworm), Drosophila melanogaster (fruit fly), Anopheles gambiae (mosquito, malaria vector),Apis mellifera (honey bee)</td>
</tr>
<tr>
<td>Tunicata</td>
<td>1</td>
<td>Ciona intestinalis (sea squirt)</td>
</tr>
<tr>
<td>Pisces</td>
<td>3</td>
<td>Takifugu rubripes (puffer or fugu fish), Tetraodon nigroviridis (puffer fish), Danio rerio (zebrafish)</td>
</tr>
<tr>
<td>Aves</td>
<td>1</td>
<td>Gallus gallus (red jungle fowl)</td>
</tr>
<tr>
<td>Mammalia</td>
<td>5</td>
<td>Rattus norvegicus (brown rat), Mus musculus (house mouse), Canis familiaris (domestic dog), Pan troglodytes (chimpanzee), Homo sapiens (human)</td>
</tr>
<tr>
<td>Animals; total</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Eukarya: total</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>279</td>
<td></td>
</tr>
</tbody>
</table>

Sources: from www.genomesonline.org, genomewebnewsnetwork.org, GenBank Nucleotide Sequence Database, and sundry sources.

qualified this announcement as more a matter of public communication than scientific achievement. At that time the accepted criterion for completion of a genome sequence, namely that only a few gaps or gaps of known size remained to be sequenced and that the error rate was below 1 in 10,000 bp, had not been met by far. The euchromatin part of the genome was not completed until mid-2004, although that milestone was again considered by some to be only the end of the beginning (Stein 2004). Nevertheless, the Human Genome Project can be regarded as one of the most successful scientific endeavours in history and the assembly of the 3.12 billion bp of DNA, requiring some 500 million trillion sequence comparisons, was the most extensive computation that had ever been undertaken in biology.

The number of organisms whose genome has been sequenced completely and published is now approaching 300 (Table 1.1). Bacteria dominate the list, as the small size of their genomes makes these organisms well-suited for whole-genome sequencing. By June 2005, no fewer than 730 prokaryotic organisms and 496 eukaryotes were the subject of ongoing genome sequencing projects. The list in Table 1.1 will certainly be out of date by the time this book goes to press, as new genome projects are being launched or completed every month.

The list of species with completed genome sequences does not represent a random choice from
was the most extensive comparison that had ever been undertaken in biology.

The number of organisms whose genome has been sequenced completely and published is now approaching 500 (Table 1.3). Bacteria dominate the list, as the small size of their genomes makes them the organisms well-suited for whole-genome sequencing. By June 2005, no fewer than 750 prokaryotic organisms and 46 vertebrates were the subject of ongoing genome sequencing projects. The list in Table 1.3 will certainly be out of date by the time this book goes to press, as new genome projects are being launched or completed every month.

The list of species with completed genome sequences does not represent a random choice from the Earth's biodiversity. From an ecologist's point of view, the absence of reptiles, amphibians, molluscs, and annelids is striking, as also is the scarcity of birds and arthropods other than the insects. How did a species come to be a model in genomics? We review the various arguments below, asking whether they would also apply when selecting model species for ecological studies.

Previously established reputation. This holds for yeast, C. elegans, Drosophila, mouse, and rat. These species and have already proven their usefulness as models before the genomics revolution and were adopted by genomics because so much was known about their genetics and biochemistry, and, perhaps just as important, because a large research community was interested, could support the work, and use the results.

Genome archaeology. One of the first questions that is asked when a species is considered for whole-genome sequencing is, what is the size of its genome? At least in the beginning, a relatively small genome was a major advantage for a sequencing project. The genome size of living organisms ranges across nine orders of magnitude, from \(10^6\) bp (0.001 Mb) in DNA viruses to nearly \(10^9\) bp (1 000 Mb) in some protists, fungi, and amphibians. The puffier fish, Tilapia rubripes, was indeed chosen because of its relatively small genome (one-eighth of the human genome).

Possibility for genetic manipulation. The possibility of genetic manipulation was an important reason why Arabidopsis, Drosophila, and mouse became such popular genomic models. The ultimate answer about the function of a gene comes from studies in which the genome segment is knocked out, downregulated, or overexpressed against a genetic background that is the same as that of the wild type. Also, the introduction of constructs in the genome that can report activity of certain genes by means of signal molecules is very important. This can only be done if the species is accessible using recombinant-DNA techniques. Foreign DNA can be introduced using transposon vectors, for example, modified P-elements that can 'jump' into the DNA of Drosophila, or bacteria such as Agrobacterium tumefaciens that can transfer a piece of DNA to a host plant. DNA can also be introduced by physical means, especially in cell cultures, using electroporation, microinjection, or bombardment with gold particles. Another popular approach is post-transcriptional gene silencing using RNA interference (RNAi), also called inhibitory RNA expression. The question can be asked, should the possibility for genetic manipulation be an argument for selecting model species in ecological genomics? We think that it should, knowing that the capacity to generate mutants and transgenes of ecologically relevant species is crucial for confirming the function of genes. Ecologists should also use the natural variation in ecologically relevant traits to guide their explorations of the genome (Kooistra et al. 2004, Toner et al. 2005). A basic resource for genome investigation can be obtained by using natural varieties of the species, and developing genetically defined culture stocks.

Medical or agricultural significance. Many bacteria and parasitic protozoa were chosen because of their pathogenicity to humans (see the many parasites in Table 1.1). Other bacteria and fungi were taken as genomic models because of their potential to cause plant disease (phytopathogenicity). Obviously, the sequencing of rice was motivated by the huge importance of this species as a staple food for the world population (Adam 2000). Some agriculturally important species have great relevance for ecological questions; for example, the bacterium Sinorhizobium meliloti, a symbiont of leguminous plants, is known for its nitrogen-fixing capacities, but it also makes an excellent model system for the analysis of ecological interactions in nutrient cycling, together with its host Medicago truncatula.

Biotechnological significance. Many bacteria and fungi are important as producers of valuable products, for example antibiotic, medicinics, vitamins, soy sauce, cheese, yoghurt, and other foods made from milk. There is considerable interest in analysing the genome of these microorganisms because such knowledge is expected to benefit production processes.
(Pühler and Selbitschka 2003). Other bacteria are valuable genomic models because of their capacity to degrade environmental pollutants; for example, the marine bacterium *Alcanivorax borkumensis* is a genomic model because it produces surfactants and is associated with the biodegradation of hydrocarbons in oil spills (Röling et al. 2004).

**Evolutionary position.** Whole-genome analysis of organisms at crucial or disputed positions in the tree of life can be expected to contribute significantly to our knowledge of evolution. The sea squirt, *Ci. intestinalis*, was chosen as a model because it belongs to a group, the Urochordata, with properties similar to the ancestors of vertebrates. The study of this species should provide valuable information about the early evolution of the phylum to which we belong ourselves. *Me. jannaschii* was chosen for more or less the same reason, because it was the first sequenced representative from the domain of the Archaea. Many other organisms, although not on the list for a genome project to date, have a strong case for being declared as model species for evolutionary arguments. These include the velvet worm, *Peripatus*, traditionally seen as a missing link between the arthropods and annelids, but now classified as a separate phylum in the Panarthropoda lineage (Nielsen 1995), and the springtail, *Folsomia candida*, formerly regarded as a primitive insect, but now suggested to have developed the hexapod body plan before the insects separated from the crustaceans (Nardi et al. 2003).

**Comparative purposes.** Over the last few years, genomicists have realized that assigning functions to genes and recognizing promotor sequences in a model genome can greatly benefit from comparison with a set of carefully chosen reference organisms at defined phylogenetic distances. Comparative genomics is developing an increasing array of bioinformatics techniques, such as syntenic analysis, phylogenetic footprinting, and phylogenetic shadowing (see Chapter 3), by which it is possible to understand aspects of a model genome from other genomes. One of the main reasons for sequencing the chimpanzee’s genome was to illuminate the human genome, and a variety of fungi were sequenced to illuminate the genome of *S. cerevisiae*.

**Ecological significance.** It will be clear that ecological arguments have only played a minor role in the selection of species for whole-genome sequencing, but we expect them to become more important in the future. Jackson et al. (2002) have formulated arguments for the selection of ecological model species, and we present them in slightly adapted form.

**Biodiversity.** The new range of models should embrace diverse phylogenetic lineages, varying in their physiology and life-history strategy. For example, the model plants *Arabidopsis* and rice both employ the C3 photosynthetic pathway. To complement our genomic knowledge of primary production, new models should be chosen among plants utilizing C4 photosynthesis or crassulacean acid metabolism (CAM). Considering the diversity of life histories, species differing in their mode of reproduction and dispersal capacity should be chosen; for example, hermaphroditism versus unisexualism or parthenogenesis versus bisexual reproduction, etc.

**Ecological interactions.** Species that take part in critical ecological interactions (mutualisms, antagonisms) are obvious candidates for genomic analysis. One may think of mycorrhizae, nitrogen-fixing symbionts, pollinators, natural enemies of pests, parasites, etc. The most obvious strategy for analysing such interactions would be to sequence the genomes of the players involved and to try and understand interactions between them from mutualisms or antagonisms in gene expression.

**Suitability for field studies.** The wealth of knowledge from experienced field ecologists should play a role in deciding about new ‘ecogenomic’ models. Not all species lend themselves to studies of behaviour, foraging strategy, habitat choice, population size, age structure, dispersal, or migration in the field, simply because they are too rare, not easily spotted, difficult to sample quantitatively, impossible to mark and recapture, not easy to distinguish from related
species, or inaccessible to invasive techniques. Thus suitability for field research is another important criterion.

Feder and Mitchell-Olck (2003) developed a similar series of criteria for an ideal model species in evolutionary and ecological functional genomics (Fig. 1.5). These authors point out that there is currently a discrepancy between classical model species and many ecologically interesting species. Models such as Drosophila and Arabidopsis are not very suitable for ecological studies, whereas popular ecological models have a poorly characterized genome and lack a large community of investigators. In some cases a large ecological community is available, but functional genomic studies are difficult for reasons of quite another nature. For example, many ecologists favor wild birds as a study object, but there are ethical objections to genetic manipulation of such species and laboratory experiments are restricted by law.

It is not easy to foresee how the list of genomic model species will develop in the future. Obviously, ecologists taking ecological genomics seriously will need to avail themselves of genomic information on their model species, preferably a whole-genome sequence. This is not to say however, that all questions in ecological genomics require the full-length DNA sequence of a species before they can be answered. Some issues may prove to be solvable with the use of less extensive genomic investigations, for example a gene hunt followed by multiplex quantitative PCR, rather than transcription profiling with microarrays of

---

**Figure 1.5 Criteria for evolutionary and ecological functional genomics for a model species, according to Feder and Mitchell-Olck (2003).** At present few species satisfy all criteria. Reproduced by permission of Nature Publishing Group.
the complete genome (see Section 2.3). In addition, microarray studies with part of the expressed genome are possible even in species lacking a complete DNA sequence. Microarrays can be manufactured at costs that are affordable for small research groups if they are limited to genes associated with a specific function or response pathway (Hed et al. 2004; see also Section 6.4). Still, the number of species with fully characterized genomes is expected to rise dramatically in the coming years; after a while all the major ecological models will also be genomic models and the saturation point could very well be due to the limited number of molecular ecologists in the worldwide scientific community.

Not all ecological models will enjoy the type of in-depth investigations now dedicated to yeast, fly, worm, and weed. Murray (2000) points out that the development of genome-based tools has a strong element of positive feedback; the rich—that is, widely studied organisms—get richer and the poor get poorer. This development has already been felt in the fields of animal and plant physiology, where many of the species traditionally investigated in comparative physiology and biochemistry have been abandoned in favour of models that can be genetically manipulated to study the function of genes. Murray (2000) predicted that ‘the larger its genome and the fewer its students, the more likely work on an organism is to die’. Crawford (2001) has argued, however, that functional genomics should resist this tendency and instead choose species best suited to addressing specific physiological or biochemical processes. For example, the Nobel Prize for Medicine was given to H.A. Krebs for his research on the citric acid cycle, which was conducted on common doves. By modern standards the dove is a non-model species, but it was chosen because its breast muscle is very rich in mitochondria. In animal physiology, Krogi’s principle assumes that for every physiological problem there is a species uniquely suited for its analysis (Grace and Cossins 2003).

According to this principle, genomic standard species are likely to be suboptimal for at least some problems of physiology, because no model is uniquely suited to answering all questions.

DNA microarrays, with their associated massive generation of data on expression profiles (see Section 2.3), are one of the most tangible features of modern genomics and are often seen as holding the greatest promise for solving problems in ecology. However, not all ecologists are convinced that microarray-based transcription profiling is the best way to advance the genomics revolution into ecology. Thomas and Klapo (2004), for example, argued that commercial microarrays are available only for genomic model species, whereas the interest of ecologists is with species that are important in the environment and amenable to ecological studies; these two interests do not necessarily coincide. This leaves ecologists with two options. One is to develop their own microarrays, starting with spotted cDNAs of unknown sequence, doing a lot of tedious sequencing work, and gradually finding out more about the genome of their study species. Another option is to apply transcriptome samples of non-models to microarrays of model species. In these cross-species hybridizations it is assumed that there is sufficient homology between the non-model and the model to allow differential expressions to be assessed reliably. For example, Arabidopsis may function as a model for other species of the Brassicaceae, and Drosophila as a model for other higher insects. Obviously, how useful such an approach is will depend on how far the sequences of model and non-model diverge. This will not be the same for all parts of the genome and therefore there is some doubt on the validity of cross-species hybridization, although there will certainly be situations where it works well.

Other investigators are less hesitant about the prospects of microarrays in ecology. Gibson (2002) emphasized that today it is feasible to establish a 5000-clone microarray resource within 12 months of a commencing project and that neither the estimated expense nor the availability of technology need to be a major obstacle for progress. We share this optimism. Given the fact that the number of almost completely sequenced organisms is increasing month by month, we can expect that the genome of several species of great interest to ecologists may be completed within a few years.
the complete genome (see Section 2.3). In addition, microarray studies with part of the expressed genome are possible even in species lacking a complete DNA sequence. Microarrays can be manufactured at costs that are affordable for small research groups if they are limited to genes associated with a specific function or response pathway (Field et al. 2004; see also Section 6.4). Still, the number of species with fully characterized genomes is expected to rise dramatically in the coming years; while all the major ecological models will also be genomic models the saturation point could very well be due to the limited number of molecular ecologists in the worldwide scientific community.

Not all ecological models will enjoy the type of in-depth investigations now dedicated to yeast, fly, worm, and wood. Murray (2000) points out that the development of genome-based tools has a strong element of positive feedback. Another is, widely studied organisms—get richer and the poor get poorer. This development has already been felt in the fields of animal and plant physiology, where many of the species traditionally investigated in comparative physiology and biochemistry have been abandoned in favour of models that can be genetically manipulated to study the function of genes. Murray (2000) predicted that 'the larger its genome and the fewer its students, the more likely work on an organism is to die'. Crawford (2002) has argued, however, that functional genomics should resist this tendency and instead choose species best suited for addressing specific physiological or biochemical processes. For example, the Nobel Prize for Medicine was given to H.A. Krebs for his research on the citric acid cycle, which was conducted on common drosophila. By modern standards the drosophila is a non-model species, but it was chosen because its brain mass is very rich in mitochondria. In animal physiology, Krogh’s principle assumes that for every physiological problem there is a species uniquely suited to its analysis (Gracey and Coe 2003). According to this principle, genomic standard species is particularly salutary for at least some problems of physiology, because no model is uniquely suited to answering all questions.

DNA microarrays, with their associated massive generation of data on expression profiles (see Section 2.3), are one of the most tangible features of modern genomics and are often seen as holding the greatest promise for solving problems in ecology. However, not all ecologists are convinced that microarray-based transcription profiling is the best way to advance the genomics revolution into ecology. Thomas and Klaper (2004), for example, argued that commercial microarrays are available only for genomic model species, whereas the interest of ecologists is with species that are important in the environment and amenable to ecological studies; these two interests do not necessarily coincide. This leaves ecologists with two options. One is to develop their own microarrays, starting with spotted cDNAs of unknown sequence, doing a lot of tedious sequencing work, and gradually finding out more about the genome of their study species. Another is to apply transcriptome samples of non-models to microarrays of model species. In these cross-species hybridizations it is seemed that there is sufficient homology between the non-model and the model to allow differential expressions to be assessed reliably. For example, Arabidopsis may function as a model for other species of the Brassicaceae, and Drosophila as a model for other insects. Obviously, how well such an approach will depend on how far the sequences of model and non-model diverge. This will not be the same for all parts of the genome and therefore there is some doubt on the validity of cross-species hybridization, although there will certainly be situations where it works well.

Other investigators are less hesitant about the prospects of microarrays in ecology. Gibson (2000) emphasized that today it is feasible to establish a 5000-clone microarray resource within 12 months of a commencing project and that neither the estimated expense nor the availability of technology need to be a major obstacle for progress. We share this optimism. Given the fact that the number of almost completely sequenced organisms in transcriptomics is to obtain a profile of global gene expression in relation to some condition of interest. Which genes are turned on and off during certain phases of the cell cycle? Which genes are upregulated by certain physiological conditions? Which genes change their expression in response to adaptation to the environment? The study of transcriptomics is part of functional genomics, because it does not look at the DNA as such, but at its functions. In general, it is expected that there are more transcripts than there are protein-encoding genes in the genome, even when considering only those genes that are actively transcribed. This is due to the mechanism of alternate splicing: the generation of different mRNAs from the same

---

**Figure 1.6** The relationship between genomics, transcriptomics, proteomics, and metabolomics.

**The genome**

- Genomics
- Transcription
  - RNA splicing
  - RNA editing
- The transcriptome
  - RNA, mRNA, tRNA
  - Translation
- The proteome
  - Enzymes
  - Structural proteins
  - Transcription factors
  - Signal proteins
- The metabolome
  - Sugars
  - Amino acids
  - Lipids
  - Secondary metabolites
  - Hormones

---
pre-miRNA during the removal of introns. RNA editing (post-transcriptional insertion or deletion of nucleotides, or conversion of one base for another) is another reason for incongruence between the genome and the transcriptome.

There are more reasons why a functional analysis of the genome can provide a different picture than an inventory of genes. Obviously, all cells of an organism have the same genome, but not the same transcriptome. Even when looking at cells of the same type, the transcriptome depends on environmental conditions, physiological state, developmental state, etc. So the transcriptome allows a glimpse of the living cell much more than the genome itself. The argument also holds when making comparisons across species. Classical molecular phylogenetics (see Graur and Li 2000) is based on variation of homologous DNA sequences across species. However, the same structural DNA can be regulated in different ways in different species. We illustrate this argument with an example from Enard et al. (2002), who did one of the first studies in what may be called comparative transcriptomics.

Enard et al. (2002) analysed the expression of 18,000 genes in liver, blood leukocytes, and brain tissue of humans, chimpanzee (Pan troglodytes), and rhesus monkey (Macaca mulatta). The expression patterns in human blood and liver turned out to be more similar to chimpanzees than to rhesus monkeys, which is in accordance with the phylogenetic distances between the three primate species; however, the expression profiles in the brain were more similar between chimpanzee and rhesus monkey than between either of the two monkey species and human (Fig. 1.7). So, although chimpanzees share 98.7% of their DNA with humans, the human species expresses that DNA in a different manner, especially in the brain. Gene expression in the brain has undergone accelerated evolution compared to gene expression elsewhere in the body, and evolution has resulted in a divergence of humans from chimpanzees, mostly due to regulatory change rather than structural reorganization of the DNA.

Proteomics is the study of all the proteins in the cell. As with genomics, proteomics arose thanks to technological innovation, which in this case is tandem mass spectrometry (MS/MS) and liquid chromatography coupled to tandem mass spectrometry (LC/MS/MS). The idea is to separate a mixture of soluble proteins by means of chromatography and then to estimate masses, first of the larger peptide and, after a second ionization, of fragments of the same peptide. The fragment patterns provide a fingerprint characteristic of the protein. Interpretation of proteomics data is
pre-mRNA during the removal of introns. RNA editing (post-transcriptional insertion or deletion of nucleotides, or conversion of one base for another) is another reason for incongruence between the genome and the transcriptome. There are more reasons why a functional analysis of the genome can provide a different picture than an inventory of genes. Obviously, all cells of an organism have the same genome, but not the same transcriptome. Even when looking at cells of the same type, the transcriptome depends on environmental conditions, physiological state, developmental state, etc. So the transcriptome allows a glimpse of the living cell much more than the genome itself. The argument also holds when making comparisons across species. Classical molecular phylogenetics (see Gauze and Li 2000) is based on variation of homologous DNA sequences across species. However, the same structural DNA can be regulated in different ways in different species. We illustrate this argument with an example from Emerad et al. (2002), who did one of the first studies in what may be called comparative transcriptomics.

Emerad et al. (2002) analysed the expression of 18,000 genes in liver, blood leukocytes, and brain tissue of humans, chimpanzee (Pan troglodytes), and rhesus monkey (Macaca mulatta). The expression patterns in human blood and liver turned out to be more similar to chimpanzees than to rhesus monkeys, which is in accordance with the phylogenetic distance between the three primate species; however, the expression profiles in the brain were more similar between chimpanzees and rhesus monkeys than between either of the two monkey species and human (Fig. 1.7). So, although chimpanzees share 98.7% of their DNA with humans, the human species expresses that DNA in a different manner, especially in the brain. Gene expression in the brain has undergone accelerated evolution compared to gene expression elsewhere in the body, and evolution has resulted in a divergence of homologs from chimpanzees, mostly due to regulatory change rather than structural reorganization of the DNA.

Proteomics is the study of all the proteins in the cell. As with genomics, proteomics owes thanks to technological innovation, which in this case is usually supported by genomic sequence information, in such a way that an observed peptide fragment pattern may be compared to a database of proteins predicted from the genome. Mass spectrometry may also be used to determine the amino acid sequence of a protein. For this application, the protein is cleaved with a protease, for example trypsin, which generates a collection of fragments characteristic of the protein. These fragments may be compared to an in silico (computer-simulated) digestion derived from the genome and the known digestion sites of the protease.

The proteome provides a different picture of a cell’s activities to the transcriptome. Several authors have indeed wondered about the lack of correlation between mRNA and protein abundances. One of the reasons for this is the existence of control mechanisms at the ribosomes, where mRNA is translated to peptides. Translational control allows the cell to select only certain mRNAs for translation and block others. The selection is often dependent on environmental conditions, so this mechanism allows for physiological adaptation on the level of the proteome, even though the transcriptome remains the same. Another issue is post-translational modifications or protein processing, processes that can greatly affect the function of a protein, for example by acetylation or ubiquitination of the N-terminal residue, hydroxylation of prolines, or cleavage of the covalent into smaller units. The proteome and the genome are linked by many feedback mechanisms, because some proteins are transcription factors necessary for gene activation, others are enzymes involved in transcription or translation, and still others are structural components of chromosomes. So, in a molecular biology context, the living cell can only be understood fully by considering genome, transcriptome, and proteome together.

As an example of a study applying proteomics in an environmental context, consider the work of Shadler et al. (2003). These authors studied protein fingerprinting in embryos of zebrafish (Danio rerio) exposed to environmental endocrine disruptors. The compound p-nonylphenol is a degradation product of certain detergents and is discharged into the aquatic environment through sewage effluent. Because of its structural similarity to vertebrate steroid hormones, especially oestrogens, nonylphenol has been associated with feminization of male fish. Fig. 1.8 shows a two-dimensional gel of differential protein expression of fish exposed to nonylphenol. This so-called protein-expression profile was composed by matching the treatment profile with the control profile and subtracting them from each other. The Venet diagram in Fig. 1.8b provides a pictorial illustration of the number of proteins that are shared between treatments. It is interesting to note that nonylphenol induced several proteins that were also induced by oestradiol (23 in total), but that a

![Figure 1.7](image-url) Figure 1.7 Distance trees showing the similarity of gene-expression profiles in brain, blood leukocytes, and liver of human, chimpanzee, and rhesus monkey. Numbers refer to the rate between the rate of evolution of the human and the chimpanzean lineages, taking the rhesus monkey as an outgroup. Reprinted with permission from Emerad et al. (2003). Copyright 2002 NAS.

![Figure 1.8](image-url) Figure 1.8 (a) Features on images from zebrafish embryos induced by exposure to endocrine-disrupting compounds of a two-dimensional electrophoresis gel, on which proteins are separated by a combination of isoelectric point and molecular mass. Showing only proteins that were differential between the control and nonylphenol-exposed zebrafish. (b) Venet diagram representing the number of proteins shared by two or more treatments. The diagram shows that, from the total of 201 proteins, there were 22 seen only in the control treatment and 23 seen only in the nonylphenol treatment and 23 are in both the nonylphenol and the control treatment, with the natural hormone cortisol. After Shadler et al. (2003) with permission from Springer.
significant number of proteins (32) were specific to nonylphenol. The study suggests that the two compounds have overlapping but otherwise dissimilar modes of action and that it may be too simple to qualify nonylphenol as only mimicking oestradiol. The functional genomics of endocrine disruption will be discussed in more detail in Chapter 6.

*Metabolomics* is the study of all low-molecular-weight cellular constituents. Usually only metabolites belonging to a limited category are included, for example all soluble carbohydrates, or all metabolites that can be measured by a certain analytical technique such as pyrolysis gas chromatography or infrared spectrometry. No single method can measure the thousands of different chemical compounds that may be present at any time in a cell, because of the greatly diverging chemical properties (hydrophilic versus hydrophobic compounds, acids versus bases, reactive versus inert compounds, etc.). The metabolome requires a diversity of analytical approaches to obtain a complete picture.

There are still hardly any studies of proteomics and metabolomics that address a truly ecological question and that is why both of these -omics do not play a major role in this book. Their role could grow in the future, when ecology has absorbed the principles of genomics. In Chapter 7 we will address some aspects of metabolomics when discussing metabolic networks. Finally, Table 1.2 describes some other terms used in connection with genomics.

With the further development of ecological genomics, applications will also come within reach. One can envisage a multitude of issues where a better knowledge of genomes in the environment can support measures to improve ecosystem health, risk assessment of pollution, conservation of endangered species, etc. (Greer et al. 2001). Such applications fall outside the scope of this book; however, we mention two examples below, to sketch the range of possibilities.

Purohit et al. (2003) suggested that multilocus DNA fingerprints prepared from environmental samples could act as an *indicator DNA signature* (IDS); for example, fingerprints of microbial soil communities could be indicative of soil pollution. Their suggestion can be extended to involve transcription profiles that are characteristic of certain environmental conditions or physiological states. Fig. 1.9 illustrates this principle. When an organism is exposed to polluted soil, this will be accompanied by gene expression that has both a general aspect due to the generality of the stress response and a specific aspect which characterizes the challenge (see also Chapter 6). When the expression profile observed for a suspect soil is compared with a database of reference profiles, the type of pollution and its biological effects may be indicated (Fig. 1.9). This may help to support decisions about the urgency of remediation measures.

As a second example of a possible application consider the case of soil-borne pathogens. Many pathogens attacking economically important crops are difficult to control by conventional strategies such as the use of host resistance and synthetic pesticides. However, some soils have an inherent capacity to suppress diseases and such soils need lower rates of pesticide application to combat

<table>
<thead>
<tr>
<th>Term</th>
<th>Object of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathogenomics</td>
<td>Genomes of human pathogens: analysis of genes involved in disease generation</td>
</tr>
<tr>
<td>Pharmacogenomics</td>
<td>Genomic responses to drugs, analysis of expression profiles that indicate similarity of action across compounds, analysis of genetic polymorphisms that determine a person's disposition to drug action</td>
</tr>
<tr>
<td>Toxicogenomics</td>
<td>Mode of action of toxic compounds, development of expression profiles that indicate similarity of toxic action across compounds</td>
</tr>
<tr>
<td>Ecotoxicogenomics</td>
<td>Genomic responses of organisms exposed to environmental pollution</td>
</tr>
<tr>
<td>Ionomics</td>
<td>All mineral nutrients and trace elements in an organism, for example using inductively coupled plasma mass spectrometry (ICP-MS)</td>
</tr>
</tbody>
</table>
significant number of proteins (O2) were specific to nonmycelial. The study suggests that the two compounds have overlapping but otherwise dissimilar modes of action and that it may be too simple to quantify nonmycelial as only reddening coenocidial. The functional genomics of endocrine disruption will be discussed in more detail in Chapter 6.

Metabolomics is the study of all low-molecular-weight cellular constituents. Usually only metabolites belonging to a limited category are included, for example all soluble carbohydrates, or all metabolites that can be measured by a certain analytical technique such as pyrolysis gas chromatography or infrared spectrometry. No single method can measure the thousands of different chemical compounds that may be present at any time in a cell, because of the greatly diverging chemical properties (hydrophilic versus hydrophobic compounds, acids versus bases, reactive versus inert compounds, etc.). The metabolome requires a diversity of analytical approaches to obtain a complete picture.

There are still hardly any studies of proteomics and metabolomics that address a truly ecological question and that is why both of these -omics do not play a major role in this book. Their role could grow in the future, when ecology has absorbed the principles of genomics. In Chapter 7 we will address some aspects of metabolomics when discussing metabolic networks. Finally, Table 1.2 describes some other terms used in connection with genomics.

With the further development of ecological genomics, applications will also come within reach. One can envisage a multitude of issues where a better knowledge of genomes in the environment can support measures to improve ecosystem health, risk assessment of pollution, conservation of endangered species, etc. (Greer et al. 2001). Such applications fall outside the scope of this book; however, we mention two examples below, to sketch the range of possibilities.

Purol et al. (2003) suggested that multilocus DNA fingerprints prepared from environmental samples could act as an indicator DNA signature (IDS); for example, fingerprints of microbial soil communities could be indicative of soil pollution. Their suggestion can be extended to involve transcription profiles that are characteristic of certain environmental conditions or physiological states. Fig. 1.9 illustrates this principle. When an organism is exposed to polluted soil, this will be accompanied by gene expression that has both a general aspect due to the generality of the stress response and a specific aspect which characterizes the challenge (see also Chapter 6). When the expression profile observed for a suspect soil is compared with a database of reference profiles, the type of pollution and its biological effects may be indicated (Fig. 1.9). This may help to support decisions about the urgency of remediation measures.

As a second example of a possible application consider the case of soil-borne pathogens. Many pathogens attacking economically important crops are difficult to control by conventional strategies such as the use of host resistance and synthetic pesticides. However, some soils have an inherent capacity to suppress diseases, and such soils need lower rates of pesticide application to combat them. Disease-suppressive capacity is due to the presence of genes involved with antibiotic production by antagonistic microorganisms (Van Elsas et al. 2002; Weller et al. 2002; Garbera et al. 2004). In several cases, specific microbial populations have been identified that contribute to disease suppressiveness; however, for most soils, we have little understanding of the consortium of microorganisms and the corresponding genes that are responsible for this critical function. Natural disease-suppressive soils can be regarded as a largely untapped resource for the discovery of new antagonistic microorganisms and antibiotics. We will see several examples of this in Chapter 6. Management strategies can be developed that involve selective stimulation and support of populations of antagonistic microorganisms in the rhizosphere. Genomic methods of soil diagnosis could be used as feedback on agricultural management decisions.

<table>
<thead>
<tr>
<th>Term</th>
<th>Object of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathogenomics</td>
<td>Genomes of human pathogens; analysis of genes involved in disease generation</td>
</tr>
<tr>
<td>Pharmacogenomics</td>
<td>Genomic responses to drugs, analysis of expression profiles that indicate toxicity of action across compounds, analysis of genetic polymorphisms that determine a person’s disposition to drug action</td>
</tr>
<tr>
<td>Toxicogenomics</td>
<td>Mode of action of toxic compounds; development of expression profiles that indicate similarity of toxic action across compounds</td>
</tr>
<tr>
<td>Ecotoxicogenomics</td>
<td>Genomic responses of organisms exposed to environmental pollution</td>
</tr>
<tr>
<td>Limnogenomics</td>
<td>All mineral nutrients and trace elements in an organism, for example using inducibly coupled plasma mass spectrometry (ICP-MS)</td>
</tr>
</tbody>
</table>

1.4 The structure of this book

We have organized this book to address fundamental questions in three areas of ecology where we believe ecological genomics can make important contributions. Having given a broad introduction to genomics, and ecological genomics in particular, in this chapter, two more specific introductory chapters follow. Chapter 2 explores the most important genomic methodologies, and Chapter 3 gives a survey of what can be learnt from comparing the genomes of model organisms with each other and with those of evolutionarily related species. We also discuss the various properties of both prokaryotic and eukaryotic genomes. Chapters 2 and 3 form the methodological and evolutionary basis for the rest of the book. In the next three chapters, questions relate to different levels of integration, from community ecology down to population ecology, ending with
physiological ecology. Each of these chapters ends with an appraisal of how the genomics achievements contribute to answering the basic question of the chapter.

Community structure and function. In Chapter 4 the genomics approach is used to discuss a question fundamental to community ecology, of how biodiversity supports ecosystem function. Most of the examples in this chapter are taken from microbial ecology. We review the ways in which microbiologists use genomics to estimate species diversity in the environment and how functions of uncultured species can be reconstructed from environmental genomes.

Life-history patterns. Chapter 5 discusses the genomic aspects of life-history evolution, an important theme in population ecology. Questions of longevity, reproductive effort, sex, and diapause are discussed, as well as the issue of trade-offs between life-history traits. We show that progress in mechanistic studies of plasticity and optimal timing of reproduction has considerable relevance to ecology.

Stress responses. The many genomic studies of mechanisms that allow plants and animals to survive in harsh environments form the subject of Chapter 6. The way in which plants and animals transduce stress signals into gene expression shows many commonalities across species, as well as stress-specific signatures. We argue that insights in these mechanisms is needed to define the ecological niche of the species.

Integrative ecological genomics. We conclude the book with a short chapter on integrative approaches, discussing some aspects of network analysis and ecological control analysis. These two approaches belong to the realm of systems biology, a new field of research, linking genomics, proteomics, and metabolomics with biochemical modelling. Chapter 7 suggests that integrative approaches are also required in ecological genomics and it discusses some examples to support this claim. Finally, a number of emerging issues are discussed in the outlook section of Chapter 7.