The frontiers of current biological research

Michel Morange

What are the frontiers?

In this lecture, I will try to define the current frontiers of knowledge in biology. There are different ways to have view of them. The frontiers of knowledge can be conceived of as the place where scientists are actively digging, the questions they are presently asking. If one compares scientific progress to the displacement of a cell like a macrophage, sliding on a substrate of knowledge, extending its pseudopods towards new unexplored territories, the frontier of knowledge can be seen as this tiny territory between the most advanced pseudopods and the *terra incognita* in front of them.

But the frontiers of knowledge can also be seen as the limits of knowledge, as the obstacles that have to be overcome, as the gaps that have to be filled. If you ask biologists, they will give you a list of unanswered questions, a series of current descriptions that are considered insufficient. The limits of knowledge also dramatically emerge when biological knowledge is used for practical issues, as in fighting diseases. Progress is being made in the fight against cancer, but

at too slow a pace. Neurodegenerative diseases are increasingly well described, but no one therapeutic strategy has yet proved compelling. Here, the frontiers of knowledge coincide with the limits of our action. The way to overcome these limits is presently unknown.

As François Jacob said "Yet, while it is part of our nature to produce a future, the system is geared in such a way that our predictions have to remain dubious. We cannot think of ourselves without a following instant, but we cannot know what this instant will be like. What we can guess today will not be realized. Change is bound to occur anyway, but the future will be different from what we believe. This is especially true in science. The search for knowledge is an endless process and one can never tell how it is going to turn out. Unpredictability is in the nature of the scientific enterprise. If what is to be found is really new, then it is by definition unknown in advance. There is no way of telling where a particular line of research will lead" (Jacob 1982).

So, the frontiers of knowledge are different from the future of biology, and it would be unreasonable to hope to say what biology will be tomorrow. All we can do is to give a picture of present trends in research. Some trends correspond to efforts made to overcome the limits of knowledge, but others are simply the consequence of the progress made in one technology, in one new experimental approach.

I will organize this lecture in three parts. In the first, I will show you that molecular descriptions still have a strong explanatory power, and that these descriptions will continue to have an important place in the future of biological research. In the second, I will argue that they must be complemented, and extended. The rather static descriptions given so far will become more and more dynamic, and a global picture will replace the present piecemeal descriptions. Finally, a strong trend in biological research, already clearly visible, is the progressive encounter between two branches of biological research which hitherto have remained far apart: functional biology, i.e., biochemistry, molecular and cell biology, and physiology, all disciplines in which scientists describe the mechanisms at work within organisms, and evolutionary biology, in which researchers seek to provide scenarios for the evolution of organisms. This current trend fleshes out the hitherto abstract scenarios of evolutionary biologists, and provides explanations for the emergence of complex macromolecular nanomachines. I am convinced that the exchange of models between the two branches of biology will considerably enrich biological knowledge.

I Molecular descriptions conserve their explanatory power

It is frequently said today that biology must abandon the reductionist approach which had been dominant during the era of molecular biology, when the secret of life was looked for in the characterization of isolated macromolecules. A global vision should replace the previous reductionist one. Molecular biology is dead, and the period when it dominated the landscape of biological disciplines should be seen as a sad period in the history of biology.

Such statements are at odds with the present situation of biology (Morange 2008). The "paradigm" of molecular biology consisted in looking for the explanation of functions in the structural descriptions of macromolecules. The structure of DNA was emblematic of the way structure explains the function of macromolecules. The structure of DNA immediately showed how this molecule was able to bear genetic information. It explained how it was able to generate two identical copies of itself, by separation of the two strands, and synthesis of the complementary strand. It also showed, as Jim Watson and Francis Crick immediately understood, how the sequence of nucleotides might be the code by which genes had their effects within organisms. The same heuristic value of the structural description is also clearly visible in the case of proteins and enzymes. The structures explain how the molecules are able to fulfil their functions. For enzymes, the description of the amino acids in the active site explains how these enzymes are able to catalyze specific chemical reactions. Other functions of proteins, such as the capacity to be a receptor for a signal or a channel for ions, can also been

explained by a precise structural description of these proteins and of the way they behave as nanomachines. We will consider one of many possible examples. The potassium channel, located in the cell membrane, is involved in the production of the nerve influx, i.e., the way nerve cells communicate with other nerve cells, and activate target cells such as muscle cells. It was demonstrated more than fifty years ago that the nerve influx results from the occurrence of transient transmembrane currents, due to the passage of ions across the cell membrane. It was later shown that the passage of ions was permitted by the existence of protein channels. The structure of these channels has been characterized, and it has been fully explained how the channels are able to fulfil their three functions: to open transiently when the transmembrane voltage is altered by the propagation of the nerve influx, to close after a short time, and to be specific for one particular ion (Jiang et al. 2003). To provide this explanation, the static structure revealed by X-ray diffraction studies is used to provide a scenario of the internal movements of the macromolecule, how its different parts move one relative to the other. This dynamic description explains how these proteins are able to work as superb nanomachines to fulfil their functions.

The heuristic power of structural determination, its capacity to provide satisfactory explanations of the behaviour of macromolecules, is not decreasing: the opposite is true. The huge progress made in these methods, the development of new methods providing a more dynamic description of the internal movements of these nanomachines, and the possibility of using the information gathered to design new therapeutic agents make it highly improbable that the role of structural information in the explanations of biologists will diminish. The description at higher levels of organization will probably expand (see later), but these new levels will not replace the molecular level. What was learned from the description of macromolecules will remain at the core of our knowledge of organisms. The macromolecular level is not one among many other levels: it is the level at which information is encoded in the genome. This gives it a preeminent role.

The founders of molecular biology, such as Francis Crick and Jacques Monod, said that they had discovered the secret of life. Clearly, many questions remain unsolved in biology, and much exciting work awaits future biologists! The development of an organ as complex as the brain is clearly not understood. And the way to fight many diseases is unknown. But Crick and Monod were not wrong. Some fundamental principles explaining the characteristics of present-day organisms have been understood – the existence of genetic information, of a genetic code. The advances in understanding organisms have been so dramatic that it is now reasonable to conceive of synthesizing a living, totally artificial organism, as some synthetic biologists now do.

II Molecular explanations must be complemented, and extended

There are different ways to complement the current molecular description of organisms. The first is to follow the same path as during the last decades. Recently, totally new phenomena have been discovered. They do not abolish the previous observations, but they add a layer of complexity. Whereas the regulation of gene expression was attributed to proteins, the so-called transcription factors, microRNAs are increasingly seen to play a part in gene regulation. Similarly, regulation of gene expression by the control of proteins surrounding DNA, the histones and more generally the chromatin, appears more and more important. These epigenetic marks can be transmitted through cell division, and in some cases as in plants, through generations. They can be modified by the environment, and they give organisms a capacity to stably adapt to new environments by modifications that do not alter genetic information. In addition, the numerous studies done on the different molecular networks in cells – gene regulatory and signalling networks – unveil interactions and regulations never seen before.

Even more significant are the technological developments that complement the structural determination, and make it more precise and more dynamic. The first consists in studying

isolated macromolecules, by using tricks permitting their micromanipulation. The result is a better and more physical description of the way they act as nanomachines. The second consists in observing individual molecules directly in cells. This has been made possible by the development of molecular tools – the use of fluorescent proteins which can be coupled with the molecules under study – and a parallel dramatic progress in the sensitivity of the physical devices allowing the detection of these weak signals.

New phenomena have been discovered, or at least revealed, by the use of these new technologies, and they raise important and still unresolved questions. It was anticipated, due to the low numbers of molecules in cells, and the slow rate of some of the most important reactions, such as the initiation of transcription, that many processes in cells would not be regular, but vary in a random way. The existence of these stochastic variations, called "noise", was rapidly demonstrated (Raser and O'Shea 2004). For instance, transcription of the same gene can differ from one cell to another, as well as between the two copies of the same gene.

The discovery of this phenomenon, thanks to the development of the new technologies that I described previously, raises at least two important questions. The first concerns the way organisms are able to cope with these stochastic variations. Is the architecture of the molecular networks specially designed to buffer these stochastic variations? The latter can generate a diversity of phenotypes unrelated to the diversity of genotypes. Is this phenotypic plasticity exploited by organisms to adapt to changing environments? These questions have received preliminary answers, but much remains to be learnt about these newly discovered phenomena.

There is a marked tendency in present biological research to collect information on individual molecules in order to predict the behaviour of the complex systems of which these macromolecules are a part. Traditional descriptions in molecular and cell biology were qualitative. Components of the networks, and relations between these components, were represented on a simple diagram, and the global behaviour of the system was interpreted with the help of this diagram. But the complexity of these networks, with the existence of multiple positive and negative feedback loops, has become such that interpretations become more and more problematic, and the predictions incorrect. These limits were clearly revealed by the knockout experiments initiated by biologists at the beginning of the 1990s. One gene, whose individual function was believed to be well known, was selectively inactivated, and the effects on the organism of this inactivation were highly different from those expected. Similarly, in cancer, the networks involved in the control of cell division have been fully described, but this description remains insufficient to predict – and sometimes even to explain – the effects observed following the modifications of one or other of the components of these networks that occur during the formation of tumours.

Biologists are convinced that these limits will be overcome by the use of formal, mathematical models reproducing the behaviour of these complex networks. The task is gigantic, due to the huge number of different components, the heterogeneity of the medium in which they work, and the still poor knowledge of their *in vivo* concentrations. Whatever these difficulties, models have an increasing place in the work of biologists. They play roles that they traditionally had in physics, but that are new in biological research. They can help biologists to construct new regulatory circuits and functional modules. This is the case in synthetic biology, when new functional devices are introduced into recipient (bacterial) organisms. Formal models can also help biologists to check whether they have correctly described a system, the macromolecular components involved, and the relations between them. If the behaviour of the system is stable, and correctly reproduces the *in vivo* behaviour, the answer will be positive. Otherwise, the results afforded by the model will help biologists to look for these missing components and relations. The model can also be used to test a simple hypothesis, to see whether it is consistent with current knowledge. It can avoid useless (wet) experiments and force experimenters to make their

hypotheses more precise. These changes in the way of doing research, what is called epistemology, are probably more important than the increase in knowledge or the development of new technologies in the design of what will be biology in the mid and late 21st century.

III The encounter between functional and evolutionary biology

In a famous article published in 1961, the great evolutionist Ernst Mayr noticed that there were two highly different categories of biologists: those interested in the way cells and organisms function, in the elucidation of the complex mechanisms behind this perfect functioning, and those more interested in the raison d'être of these complex functions and the adaptation they provide to organisms harbouring them. The first category includes physiologists, molecular and cell biologists, biochemists; the second, evolutionists, but also ecologists and zoologists. Geneticists are at the boundary between the two groups: they can be molecular geneticists or population geneticists. Ernst Mayr underlined the differences between the two approaches, and their complementarity (Mayr 1961). But this complementarity was one of principle, not of facts. Functional biologists worked as if the complex structures they studied had no history, were not the product of evolution; and evolutionary biologists paid no attention to the mechanisms behind adaptation, considering that organisms had multiple possibilities to adapt, and that the description of the mechanisms by which they specifically adapted would add nothing to our understanding of the evolutionary process. The gap between the two forms of biology was the niche in which the supporters of Intelligent Design found their arguments: they stated that there are no natural explanations for the emergence of these splendid functional devices in organisms: they may only have been designed by a superior intelligence.

During the last century there were some attempts to fill the gap between the two forms of biology. But these efforts did not lead to the development of research programmes. One reason, but not the only one, was the difficulty of the task, and the lack of an appropriate methodology. Another reason was mutual ignorance, resulting from the different university training of functional and evolutionary biologists.

The situation is rapidly changing, for different and converging reasons. The first is the progress of transdisciplinarity within biology. The trigger was the development of genomics and post-genomics, which required the skills of computer scientists, mathematicians, physicists, and engineers. This new group of "biologists" had not been trained to find the separation between the two branches of biology "natural". When they started to work in biology, they rapidly moved from functional to evolutionary questions, without having the feeling of committing a transgression.

This is particularly evident in systems biology, where researchers seek to describe the organization of macromolecules in complex networks in cells and organisms, and the behaviour of these networks. Leaders in the field, such as Uri Alon and Stanislas Leibler, rapidly moved from a description of these systems to questions about their origin, and the adaptive value they provide. Some of the scenarios imagined for their origin and/or adaptive value were sometimes naive, or supported by inadequate observations (Keller 2005). These researchers clearly demonstrated the value, but also the difficulty, of bringing together the two forms of biology.

The second reason for the encounter between these two branches of biology is also a natural consequence of the development of gene sequencing programmes. A sequence is not informative *per se*. One of the only ways to extract information from a sequence is to compare it with other sequences. This comparison can be used to discover the function of hitherto unknown genes, and the question is therefore limited to functional biology. But, in general, comparison of sequences, or of the organization of genes in the genome, immediately leads to questions about the evolution of the systems under comparison. These questions can be limited to a description of what happened, for

instance a characterization of the genes that were duplicated or deleted during the evolution of one or other species, but questions soon arise regarding the selective pressures that have moulded evolutionary history.

The discovery of the genes controlling development, the "master genes" such as the homeotic genes, and of their conservation during evolution, has attracted the attention of evolutionists. More and more studies in a new discipline dubbed "Evo-Devo" aim to relate modifications in structure and/or expression of these developmental genes to the evolutionary transitions revealed by the work of palaeontologists. Molecular mechanisms for these transitions are proposed, as well as evolutionary scenarios for their emergence.

In the previous case, the efforts to fill the gap between functional and evolutionary biology were made by evolutionists. But the opposite can be true, and more and more evolutionary questions spontaneously emerge from the work of molecular and cell biologists. One reason is that the molecular descriptions have been pushed so far that molecular explanations reveal their limits. The specific characteristics of a functional device, for instance a multi-molecular system, can only be found in the complex evolutionary process which has generated it. Consider, for instance, a superb nanomachine like a chaperonin, in charge of the correct folding of proteins in cells - the process by which the long polypeptide chains attain their native structures. Why are only some proteins of the cell the targets of the chaperonin? The explanations in terms of differences in the physico-chemical characteristics of the target proteins are not wrong, but they must be complemented by other explanations putting the complex relations between proteins and their chaperonins in an evolutionary perspective, and by trying to explicate the selective pressures exerted on these complex systems (Kerner *et al.* 2005).

The recent possibility of studying evolution *in vitro*, to put "the Beagle in a bottle" as beautifully said in an article in *Nature* (Buckling *et al.* 2009), was a major event in moving functional and evolutionary biology closer one together. The possibility of studying evolution in the laboratory is not new. Pioneering work was done, for instance, on the fruit fly *Drosophila*. But these early studies rapidly reached their limits in the number of organisms included, the number of generations that could be studied, and in the practical possibility of relating the transformations observed to the genetic and molecular mechanisms underlying them. Such limits were overcome by the adoption of bacteria (or viruses) as model systems, and also by the *in vitro* study of the evolution of macromolecules, RNAs and proteins. The path followed by evolution, the constraints on the system, the trade-off between antagonistic changes are no longer "abstract words", but can be identified with precise molecular events. And the same strategies can now be applied to naturally evolving more complex organisms, such as the rapidly evolving cichlid fishes in the East-African great lakes (Kocher 2004).

Synthetic biology adds a new dimension to these efforts. It will be possible in the future to test molecular functional devices that have not been selected by evolution. In this way, it will be possible to discriminate in the ensemble of possibilities to which life has not yet had access those that are forbidden and those that have not yet been exploited by organisms, or to which access was made more difficult by the choices initially made.

Needless to say, this work reduces the gap between functional and evolutionary biology. More and more work is being done in this direction and represents a strong trend in biology.

One discipline, epidemiology, was a precursor in the encounter between the two forms of biology. To explain the emergence of an epidemic or a pandemic requires a description of the nature of the pathogenic agent, but also the characteristics of the disease and of its transmission in the human population. The capacity of a pathogen to evolve and its complex relations with its hosts have been progressively described. The studies have become more and more precise. The genomes of most pathogens being small, a full description of them and of their evolution in parallel with the development of an epidemic has been possible, as in the case of HIV and AIDS. The resistance

of pathogens to treatments, such as antibiotics, has been fully explored: the mechanisms involved have been described in the smallest detail, as has the propagation of resistance.

The study of diseases like cancer is also benefitting from these combined efforts of functional and evolutionary biologists. Instead of being considered as the simple result of the addition of somatic mutations, the formation of tumours and metastases is now seen as a long evolutionary process in which cancer cells progressively adapt to new niches within the organism. More generally, the study of diseases is giving increasing scope to evolutionary considerations. To explain pathology affecting human beings, one must take into account the recent evolutionary history that has generated modern humans, and the ecological niche in which this evolutionary history took place.

By comparing the models and results of both branches of biology, scientists will be able to elaborate a less naive vision of what happened during evolution. Such a naive vision is obvious in the case of the evolutionary origin of modern humans, perhaps because so much is at stake! For the moment, one has the choice between the naive models of evolutionists trying to describe how our ancestors left the branches of the trees and stood up in the savanna, and those of geneticists and molecular biologists outlining the crucial change in one gene, the "language gene" (Vargha-Khadem *et al.* 2005) or the "jaw gene" (Stedman *et al.* 2004). Human evolution was far more complex, more tortuous, and so much remains to be discovered!

Conclusion

The closer relations progressively established between functional and evolutionary biology will deeply affect the way biology is studied. Consider, for instance, the importance of model organisms in 20th century biology: the fruit fly (*Drosophila*), bacteria (*Escherichia coli* and its bacteriophages), the nematode, mice. Most efforts by biologists were focused on these systems. They were not useless, they permitted the characterization of the most fundamental mechanisms operating in organisms. The new biology will probably be more open to diversity, to a plurality of models, to what happens in nature and not in the test tube or in the laboratory.

Fundamental progress was made during the 20th century in the description of macromolecular mechanisms. The complex evolutionary history of these mechanisms, and the diversity it generated, have yet to be described. This shift in interest is clearly visible when one considers the question of life. As I mentioned previously, the founders of molecular biology were convinced that they had discovered the secret of life. And they were not wrong! But what remains to be described is the complex process which generated life and its different forms. From a question of principles, the question of life has been transformed into a historical question. To reproduce extant forms of life artificially is an objective which is no longer beyond the reach of synthetic biologists. But to understand how life emerged is a different question, which is far from being solved. How the complex systems in organisms progressively emerged, and how they were gradually coupled during the long prehistory of life will require the work of many biologists. I hope that many of you who attend these lectures will participate in this exciting adventure!

Bibliography:

Angus Buckling, R. Craig Maclean, Michael A. Brockhurst and Nick Colegrave (2009) "The *Beagle* in a bottle"; *Nature* 457: 824-829

François Jacob (1982) *The possible and the actual* (Seattle: University of Washington Press)
Youxing Jiang *et al.* (2003) "X-ray structure of a voltage-dependent K^+ channel"; *Nature* **423**: 33-41

Evelyn Fox Keller (2005) "Revisiting 'scale free' networks"; *BioEssays* **27**: 1060-1068
Michael J. Kerner *et al.* (2005) "Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*"; *Cell* **122**: 209-220

Thomas D. Kocher (2004) "Adaptive evolution and explosive speciation: the cichlid fish model"; *Nature Reviews/Genetics* **5**: 288-298

Ernst Mayr (1961) “Cause and effect in biology”; *Science* **134**: 1501-1506

Michel Morange (2008) “The death of molecular biology?”; *Hist. Phil. Life Sci.* **30**: 31-42

Jonathan M. Raser and Erin K. O’Shea (2004) “Control of stochasticity in eukaryotic gene expression”; *Science* **304**: 1811-1814

Hansell H. Stedman *et al.* (2004) “Myosin gene mutation correlates with anatomical changes in the human lineage”; *Nature* **428**: 415-418

Faraneh Vargha-Khadem, David G. Gadian, Andrew Copp and Mortimer Mishkin (2005) “*FOXP2* and the neuroanatomy of speech and language”; *Nature Reviews/Neuroscience* **6**: 131-138