Naming disease proteins

In this issue, we start a series of reviews with a focus on proteins that are responsible for and named after a particular disease. When we discussed this idea at a recent editorial meeting, I was quite taken by the harmonic symbolism of this concept, which nicely combines medicine and progress in biological research. The first review describes parkin and the advances that have been made in studying this protein and finding a possible treatment for Parkinson's disease (pp TKTK). It will be followed by others, including reviews on huntingtin and Huntington's disease and dystrophin and muscular dystrophy. This series will further emphasize the field of molecular medicine that we wish to support in *EMBO reports*, in recognition of the increasing need for cooperation between basic research biologists and clinical researchers.

Many will remember the excitement in 1993 when the genetic analysis of a family suffering from Huntington's disease eventually pointed to a culprit located in a very narrow region of chromosome 4. When I read that paper, I expected that the protein causing this debilitating disease was already known and that a manuscript was ready to go to a quality journal. And if so, the researchers would surely follow up rapidly with a good guess at the function of the protein, which would soon lead to the first compounds to block or stimulate it in order to treat Huntington's disease. Wrong! The delays were much greater than anyone expected and progress to a cure is still a road being travelled by researchers, rather than a case history for students of biology.

This scenario is now too familiar and we have in turn become much more sanguine when another 'disease gene' is discovered. We know that this is not the bell signalling the last lap of a long-distance race but rather the first milestone of a marathon with many more obstacles ahead. The pathway "a disease gene is discovered, which leads to the disease-causing protein,

which leads to a definition of the molecular basis of the disease, which enables researchers to develop compounds to cure the disease" is in fact characterized by very slow kinetics and many steps that require massive inputs of labour and financial 'ATP'. This is not surprising; after all, we have known for more than 100 years which organisms cause certain infectious diseases, but some cures are still elusive.

What was surprising when we discussed this review series was the fact that there are relatively few proteins that are named after the disease in which they are involved. This surprise highlights the flawed linear view among many biologists about genes and their consequences: that one gene/one faulty protein leads to a third inevitability—disease. But life is more complex. Each month we learn that we cannot draw straight lines between the molecular components of the cellular machinery. There are redundancies; various parallel pathways can lead to the same effect. Compensatory elements allow a cell to carry on its tasks even when an apparently essential component is missing and allow for multiple choices for the use of metabolites. To think that the simple dysfunction of one gene product is the cause of a disease is to expect that evolution dared to risk all on a single option. This does not happen frequently, hence the limited number of proteins that carry the name of the disease for which they are responsible. A recent and learned viewpoint by Eduard Kellenberger in this journal (Kellenberger E (2004) *EMBO reports* 5: 546–549) points to these changes in our understanding of the rules of biology and helps us to look at our research from a fresh perspective. It also challenges the appropriateness of using simple model systems to understand human biology. In some cases, the information that can be gleaned from yeast, *Drosophila* or *Caenorhabditis elegans* is essential as a guide to similar elements in humans, but too often we are seduced by

the information from model organisms and allow ourselves to extrapolate erroneously to the different and much more complex human organism.

As most diseases are not caused by a single gene defect, we should think in terms of communities of genes and interacting products as the culprits. Luckily, DNA array technology has promoted this mindset and many researchers now mix approaches that focus on a single gene with the analysis of changes in expression patterns when a system is disturbed. Increasingly, we now equate a disease state with an expression fingerprint and use this information to search for treatments. It is still premature to use this approach for analysing disease, and we should be cautious when using it as the sole diagnostic technique, but with time we will learn how robust such patterns are and perhaps their complexity will enable a better diagnosis rather than relying on a single factor. This will have an impact on the development of diagnostic tests, on the training of physicians and on monitoring programmes to follow treatment. It will also have an impact on the biotech industry—the concept of a drug target becomes blurred if the illness is due to multiple starting points.

Although we increasingly recognize and embrace the complexity of life—and of many diseases—there are still the highly debilitating illnesses that are caused by the failure of one single gene, such as Huntington's or cystic fibrosis. And as no cure has been found for most of these, even though we have a pretty good picture of their molecular basis, they remain important topics both for clinical as well as for basic research, not to mention patients. These are at the heart of this new focus in the reviews section. We hope that you will find this series as challenging and interesting as we did when we discussed the topic.

Frank Gannon

doi:10.1038/sj.embor.7400193