THE ROYAL SOCIETY

Introduction

HUNTINGTON'S DISEASE: A CLINICAL, GENETIC AND MOLECULAR MODEL FOR POLYGLUTAMINE REPEAT DISORDERS

It is now well over a century since George Huntington (figure 1), in a brief description of admirable clarity, delineated the main features of the neurodegenerative disorder that has since been known as Huntington's disease (HD) (Huntington 1872). Working as a family practitioner on Long Island, New York State, he drew on not only his own observations of the affected families under his care, but also on those of his father and grandfather in the same practice before him, covering a total period of 60 years.

Although limited to less than two pages, Huntington's description, later cited by William Osler as a model of clinical observation (Osler 1894), covered almost all of the key features of the disease. He noted the involuntary movements or 'chorea', the adult onset, the relentlessly progressive nature and fatal course of the condition and the progressive motor disability, as well as the loss of mental function and behavioural problems, these last often occurring from an early stage of the disease. He also recognized its familial nature, correctly stating that it did not normally skip generations.

It is worth noting in the context of a scientific meeting that not only was George Huntington's classical description brief, but it was the only scientific paper that he ever wrote. Despite this, his name is rightly remembered as having provided the foundations for the later detailed studies that have made, and continue to make, HD a model not only for the study of other neurodegenerative disorders but also for a wide range of aspects involving research and practice relating to genetic disorders, and now as the principal model for the study of the newly recognized group of polyglutamine repeat disorders.

During the century after Huntington's description, a wealth of detail was recorded on the clinical and neuropathological aspects of HD (Harper 1996). These studies showed clearly that it was a primary brain degeneration, with a rather characteristic distribution involving cell loss in parts of the basal ganglia and cerebral cortex, notably the caudate nucleus. No clear indications could, however, be obtained from pathological, clinical or experimental studies as to what might be the specific underlying cause of the disorder; increasingly, workers involved in HD research looked towards the developing field of genetics to provide the answers.

The inheritance of HD had been recognized as following an autosomal dominant pattern from the time that Mendel's laws had been rediscovered. Affecting both sexes equally, transmitted only by affected individuals (apart from those dying young), and with 50% of offspring of an affected parent developing the disorder, HD provided a striking example of this mode of inheritance, whereas its high frequency in many rapidly expanding immigrant populations of European origin made it a major problem in relation to genetic counselling and to prevention of the condition. However, puzzling genetic features were noted that did not fit with conventional Mendelian inheritance, notably the paternal transmission of the rare juvenile form of the disease (Merrit *et al.* 1969), and the recognition that 'anticipation', long recognized and debated in relation to another dominantly inherited disorder, myotonic dystrophy (Penrose 1948), applied also to HD, at least in the male line (Ridley *et al.* 1991). Until the gene was isolated, solutions to these questions remained the subject of speculation only.

Gene mapping was first attempted for HD in the 1970s, using blood groups and other protein markers, but was not then successful owing to the lack of power of these markers and the scarcity of large families with a sufficient number of living affected members. It was the advent of DNA polymorphisms, along with recognition of the value of the large and extended Venezuela kindred with HD, that made localization of the gene possible, added to which should be mentioned the foresight of the Hereditary Disease Foundation in supporting the long-term effort needed and in attracting high-quality scientific groups to work on the topic. In fact, the initial localization, to everybody's surprise, came much earlier than expected with the finding (Gusella *et al.* 1983) that one of the first DNA polymorphisms available clearly localized the gene to the short arm of chromosome 4.

From this point on, the eventual isolation and identification of the gene was never in doubt, even though it took ten years to achieve this goal (Huntington's Disease Collaborative Research Group 1993). It should be remembered that at the outset of the work in 1983, the actual isolation of a gene by positional cloning was entirely a conjecture rather than an established fact. Again, the formation of the Huntington's Disease Collaborative Research Group, its funding and its coordination and general nurturing by the Hereditary Disease Foundation through years of apparently slow progress, provide an object lesson on how a truly collaborative major scientific project can be sustained and ultimately succeed.

It should not be thought that the work of the Collaborative Group proceeded on a fixed plan; on the contrary, new techniques were introduced as they became feasible, whereas new ideas from other fields of

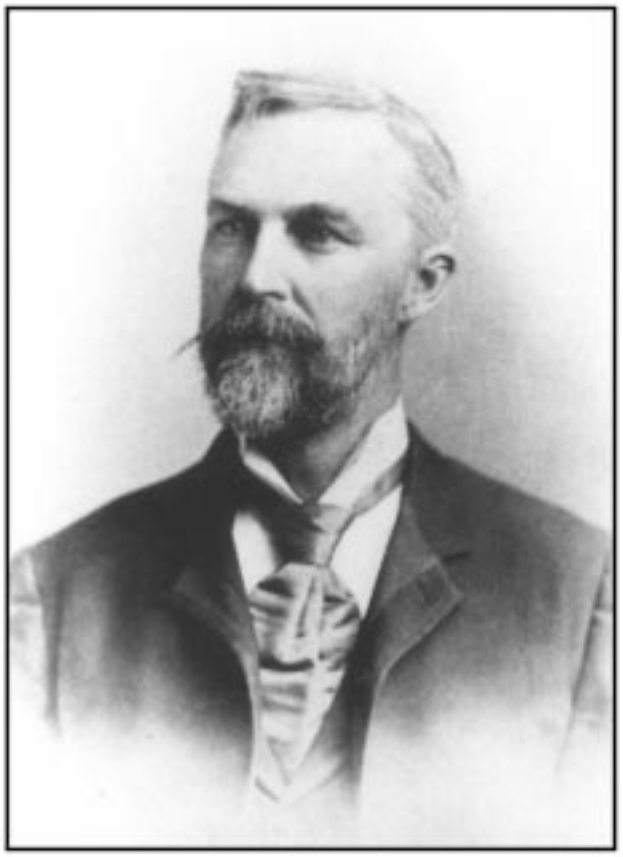

Figure 1. George Huntington in later life. (From Harper (1996), by permission of W. B. Saunders.)

work were also brought in. It was during the final but time-consuming stages of the work, when a series of possible candidate genes in the critical region were being examined for mutations, that the concept of trinucleotide repeat disorders was brought to the scene, helping greatly in the recognition of the gene and then transforming the course of all subsequent HD molecular research.

Although two disorders, Kennedy's disease (LaSpada *et al*. 1991) and fragile X mental retardation (Oberlé 1991), had been shown to result from trinucleotide repeat mutations, this did not have an immediate impact on the search for the HD gene, in part because of the X-linked nature of these disorders. Rather, it was the example of myotonic dystrophy, recognized as a trinucleotide repeat disorder (Brook *et al*. 1992), that had a profound effect, largely because two of the Collaborative Group's member teams had also been directly involved with this disease. Despite the clinical differences between the conditions, the shared genetic features of anticipation and parent of origin effects provided compelling reasons for HD also being a trinucleotide repeat disorder, so that it was the search for and eventual finding of such a repeat, and its expansion in the HD patients, that provided the conclusive proof that the gene and mutation responsible for HD had finally been discovered (Huntington's Disease Collaborative Research Group 1993).

In the immediate aftermath of this discovery attention was focused not so much on the function of the gene or the possible role of the mutation in pathogenesis but on the relation of the variability of the CAG repeat to the clinical and genetic aspects of the disease. As with myotonic dystrophy, so for HD most of the puzzling issues could now be explained (Snell *et al*. 1993): the anticipation was clearly related to genetic instability and intergenerational change in repeat length, whereas the severity of disease, in particular juvenile HD, could also be closely related to this. Parent of origin effects could be related to differences in expansion at male and female meiosis, whereas the immediate origins of the disease could be explained in terms of healthy individuals carrying an expanded allele below the repeat length of *ca*. 38 that is now recognized as critical in terms of clinical disease (Rubinsztein *et al*. 1996).

Meanwhile, research on the nature and function of the protein produced by the gene inevitably proceeded more slowly. This resulted in part from the lack of clues provided by gene sequence information, but it also reflected the fact that the skills and techniques now needed for this work were very different from those possessed by most of the genetic groups who had successfully found and isolated the gene itself. The necessary restructuring of the groups took time, as did adjustment to the fact that the field was now open to many experts who had previously had no involvement with HD. Again, it is a tribute to the loyalty that HD

Table 1. *CAG repeat expansions and inherited neurological degenerations*

disorder	protein	chromosome location	areas of brain predominantly involved
spinobulbar muscular atrophy (Kennedy's disease)	androgen receptor	Xq13	spinal and bulbar motor neurons
Huntington's disease	huntingtin	4p16	caudate nucleus, putamen; also cerebral cortex
dentatorubral–pallidoluysian atrophy (DRPLA)	atrophin	12p13	dentate and red nuclei, cerebellum, brain stem
spinocerebellar ataxia (SCA)			
type 1	ataxin-1	6p23	cerebellum and brain stem especially Purkinje cells
type 2	ataxin-2	12q24	comparable to SCA1
type 3 (Machado–Joseph disease)	ataxin-3	14q32	cerebellum; striato-nigral pathways
type 6	L1A voltage-dependent calcium channel subunit	19p13	cerebellum, brain stem
type 7	ataxin-7	13p12	cerebellum; additional retinal degeneration

research has produced that so many workers chose to stay with HD while radically altering the nature of their work, rather than simply moving on to another genetic disorder. Equally, the formation of new collaborations involving those with completely new relevant approaches has allowed the field of HD molecular research to remain extremely cohesive.

Perhaps the most striking example of this attraction into HD research of people with different skills has been the one that has given rise to the present issue and the discussion meeting that generated it: the involvement of Dr Max Perutz. It has already been noted that the initial focus of research after isolation of the HD gene was on relating the mutation to the genetic features of the disorder, but it was not long before it became clear that within the broader grouping of trinucleotide repeat disorders a subgroup could be defined in which the expanded repeat sequence was CAG, where in all cases the repeat appeared in the protein sequence, and where the clinical features represented progressive degenerations of the central nervous system of a closely similar nature. Kennedy's disease (spinobulbar muscular atrophy) has already been mentioned, but HD was closely followed by other disorders including dentatorubral–pallidoluysian atrophy (DRPLA) and a series of dominantly inherited spinocerebellar ataxias (see table 1) (Orr *et al*. 1993; Koide *et al*. 1994). The fact that all these disorders, despite widely different genes and protein sequences, shared a CAG repeat expansion suggested that the expanded polyglutamine sequence in the various proteins might be of direct relevance in the pathogenesis of the neurodegeneration of this group of disorders, something that greatly increased the interaction of people working primarily on different disorders in this group, whose research had often previously had little contact with the ideas and experimental work being done on the other polyglutamine repeat disorders.

That this concept of a common pathogenetic process turned from being an unfocused idea into a hypothesis with a clear molecular basis that could be tested experimentally can undoubtedly be attributed to Max Perutz's entry into the HD field, and can best be illustrated by quoting from his review of the topic (Perutz 1996): 'These remarkable discoveries posed a great challenge to biomedical research. What is the molecular mechanism of CAG expansion? Can it be prevented? What is the structure and function of the normal glutamine repeats? How does expansion affect them? Why is it toxic to specific neurons in the central nervous system? Can the toxic effects of expanded glutamine repeats be prevented or at least alleviated?'

This clearly set out the challenges to be met, some of which are beginning to be answered by work such as that reported in the various papers in this Discussion. However, Perutz did not stop there but produced a model that could explain the possible role of glutamine repeats in terms of structural biology. This 'polar zipper' model (figure 2) has provided the starting point for much of current HD research; its origin is again best described in his own words from the same review.

'By a strange accident, my attention was drawn to glutamine repeats before the publication of the gene for Huntington's disease. Such repeats have been found, for instance, in some homeodomain proteins of *Drosophila*. A survey of the Swiss Prot data bank showed that 33 out of 40 proteins with 20 or more glutamines in a row are transcription factors, many of them in *Drosophila*, involved mostly in the developmental regulation especially of the nervous system. Wondering what the structure of glutamine repeats

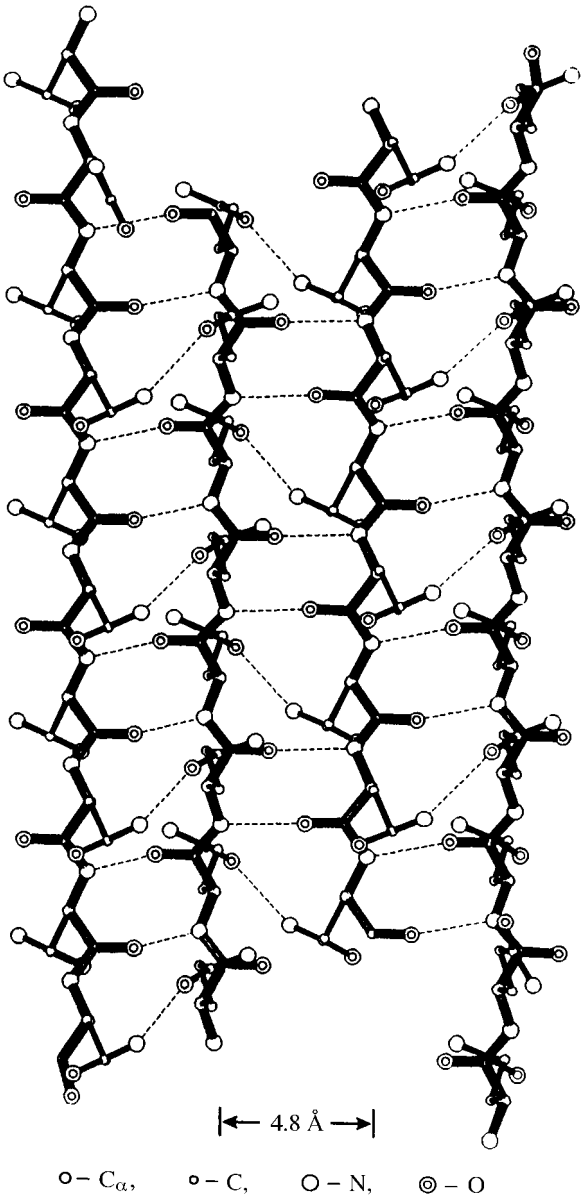

Figure 2. The 'polar zipper' model for polyglutamine repeats. (From Perutz (1996), with kind permission of the author and *Current Opinion in Structural Biology*.)

might be, I built an atomic model which showed that β-strands of poly-L-glutamine could be linked together by hydrogen bonds between both their main-chain and side-chain amides. In other words, they acted as polar zippers, which made me wonder if they attached β-strands of proteins to each other, while leucine zippers had evolved to make α-helices stick together. When I read the astonishing paper in *Cell* on the gene for Huntington's disease, it occurred to me that the polar zipper action of glutamine repeats might furnish a possible clue to the molecular mechanism of the disease, but in view of the medical importance of the problem it seemed essential to test that idea experimentally.'

Testing that idea experimentally has, as this collection of papers shows, already born fruit in terms of our increased understanding of HD and other glutamine repeat disorders, to the extent that possibilities for modifying their course look feasible in a way that would have seemed unrealistic even five years ago. Such a prospect has already further increased the impetus of the research, both basic and clinical, and now allows hope for affected patients and relatives at risk that therapy of real value will become available in the foreseeable future.

February 1999 Peter S. Harper

References

Brook, J. D., McCurrach, M. R., Harley, H. G. *et al*. 1992 Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3′ end of a transcript encoding a protein kinase family member. *Cell* **68**, 799–808.

Gusella, J. F., Wexler, N. S., Conneally, P. M. *et al*. 1983 A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* **306**, 234–238.

Harper, P. S. (ed.) 1996 *Huntington's disease*, 2nd edn. London: Saunders.

Huntington, G. 1872 On chorea. *Med. Surg. Rep.* **26**, 317–321.

Huntington's Disease Collaborative Research Group 1993 A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**, 971–983.

Koide, R., Ikeuchi, T., Onodera, O. *et al*. 1994 Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). *Nature Genet*. **6**, 9–13.

LaSpada, A. R., Wilson, E. M., Lubahn, D. B. *et al*. 1991 Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**, 77–79.

Merrit, A. D., Conneally, P. M., Rahman, N. F. *et al*. 1969 Juvenile Huntington's chorea. In *Progress in neurogenetics*, vol. 1 (ed. A. Barbeau & J. R. Brunette), pp. 645–650. Amsterdam: Excerpta Medica.

Oberlé, I. 1991 Instability of a 550 base pair DNA segment and abnormal methylation in fragile X syndrome. *Science* **252**, 1097–1102.

Orr, N. T., Chung, M., Banfi, S. *et al*. 1993 Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type I. *Nature Genet*. **4**, 221–226.

Osler, W. 1894 Case of hereditary chorea. *Johns Hopkins Hosp. Bull.* **5**, 119–129.

Penrose, L. S. 1948 The problem of anticipation in pedigrees of dystrophia myotonica. *Ann. Eugen*. **14**, 125–132.

Perutz, M. F. 1996 Glutamine repeats and inherited neurodegenerative diseases. Molecular aspects. *Curr. Opin. Struct. Biol*. **6**, 848–858.

Ridley, R. M., Frith, C. D., Farrer, L. A. & Conneally, P. M. 1991 Patterns of inheritance of the symptoms of Huntington's disease suggestive of an effect of genomic imprinting. *J. Med. Genet.* **28**, 224–231.

Rubinsztein, D. C. (and 36 others) 1996 Phenotypic characterisation of individuals with 30–40 CAG repeats in the Huntington's disease (HD) gene reveals HD cases with 36 repeats and apparently normal elderly individuals with 36–39 repeats. *Am. J. Hum. Genet.* **59**, 16–22.

Snell, R. G., MacMillan, J. C., Cheadle, J. P., Fenton, I., Lazarou, L. P., Davies, P., MacDonald, M., Gusella, J. F., Harper, P. S. & Shaw, D. J. 1993 Relationship between trinucleotide repeat expansions and phenotypic variation in Huntington's disease. *Nature Genet*. **4**, 393–397.