Their loss is our gain?

'Knockout' mice have been developed that lack the Huntington's disease gene, in an effort to gain insight into the disorder and into the pathophysiological effects of tri-nucleotide repeat expansion.

The characteristics of the mutations that cause a particular genetic disease can often be predicted simply on the basis of inheritance patterns. Recessively inherited disorders are usually caused by a loss of function of the corresponding gene product, and any mutation which destroys the function of the gene, its RNA or protein product can effect this inheritance pattern. Dominantly inherited disorders, on the other hand, express disease when only one of the two copies of the gene is abnormal. Hence the biochemical defect is usually not a complete loss of function, because one functioning normal copy of the gene is present in each cell. Instead, most dominantly inherited disorders are the result of single amino-acid changes (missense mutations) which alter the protein product in a novel manner to produce a protein with an aberrant (and deleterious) function.

Dominantly inherited diseases thus fall into one of three categories: 'change of function', in which the abnormal protein acquires a novel activity that is deleterious to the cell; 'dominant negative', in which the abnormal protein forms hetero-oligomeric complexes with the protein produced from the normal allele of the gene, thereby knocking out the function of the entire protein complex; or 'haplo-insufficiency', in which a single copy of the gene has lost function, and the 50 % of the normal amount of protein that is produced from the non-mutated allele is not adequate to preclude clinical symptoms. Excellent examples of change-of-function disorders are provided by sodium channel disorders, in which missense mutations create new channel-gating properties and result in periodic paralysis [1]. Dominant-negative protein abnormalities are exemplified by the missense mutations in collagen genes that lead to destruction of collagen fibrils and result in osteogenesis imperfecta type II [2]. An example of haplo-insufficiency is the loss of function in one copy of the low density lipoprotein receptor gene, with insufficient compensation by the good allele, resulting in familial hypercholesterolemia [3].

Tri-nucleotide repeat expansion mutations, like those that cause Huntington's disease, have only recently been identified, and there is considerable interest in determining how abnormally large repeat sequences cause molecular and clinical pathology. Diseases caused by the expansion of a tri-nucleotide motif — in which, for example, the normal sequence $(CAG)_{10}$ expands to the disease-causing $(CAG)_{200}$ — now number over ten, with some of the most common and poorly understood neurological disorders included in this group: Huntington's disease, Fragile-X mental retardation, and myotonic dystrophy. Hopefully, the nature of each disease will provide clues regarding the effect of the expansion on the protein product, aided by the loss-of-function/change-of-function paradigms described above.

Huntington's disease has been a highly visible focus of efforts to define molecular pathogenesis and the way in which tri-nucleotide repeat expansions cause disease. It is a particularly devastating disorder: family members who inherit the faulty gene in this dominantly inherited disease show no symptoms until mid- or late-life, whereupon there is a deterioration of the brain leading to a progressive loss of motor control and cognitive decline. The cause of the disorder at the anatomical level is a loss of neurons, particularly in the corpus striatum and regions of the cerebral cortex [4,5]. Genetic linkage analysis in large Huntington's disease pedigrees was used to localize the disease locus to the tip of chromosome 4 in 1983 [6], with further genetic and physical mapping leading to the identification of a novel gene and corresponding protein, 'huntingtin' or 'IT15' [7]. Conserved over a range of species [8] and expressed in a wide variety of tissues at all developmental stages [9], huntingtin would appear to have a 'housekeeping' role in all cells. Its expression pattern alone does not provide any insight into the neuronal specificity of disease pathogenesis. Lack of homology to other known proteins or functional protein domains has also left Huntington's disease researchers at a loss as to the gene's normal function.

The normal huntingtin gene contains a polymorphic stretch of CAG tri-nucleotide repeats in the beginning of the protein-coding sequence, predicted to result in between 8 and 35 glutamines in the normal protein product. Huntington's disease is caused by one copy of the gene containing an abnormally large expansion of this repeat. Presumably, half of the protein in the huntingtin-expressing cells of a patient contains 40 or more glutamines instead of the 8–35 glutamines present in the other half of the cell's protein (from the normal gene). The inclusion of too many glutamines in the huntingtin protein is suggestive of a change of function — particularly as poly-glutamine stretches have been shown to be important in the function of DNA-binding proteins involved in the regulation of transcription [10] — which would explain the observed dominant inheritance pattern. Until recently, however, no data had been accumulated concerning the normal function of the huntingtin protein; such information is critical to understanding the effect of the tri-nucleotide repeat in the mutant form.

Cross-species conservation of the huntingtin gene has led to the identification of the murine homologue [8]. To

study the normal function of the gene, the murine gene was 'knocked out', using homologous recombination technology to disrupt the gene at exon 5 [11]. The resulting knockout mice were bred to create mutant strains in which either one or both of the huntingtin alleles are inactivated. It is important to note that these mice were not like human patients with Huntington's disease: the modified genes in the mice are presumed to be loss-of-function, as a result of the a targeted deletion, whereas the mutant genes in human patients contain a tri-nucleotide repeat expansion within the coding sequence (Fig. 1a). Interestingly, there is some evidence in the paper [11] that a 20 kDa protein is produced from the disrupted huntingtin gene in the brains of heterozygote mice. The authors assumed, as we will here, that this putative truncated mouse huntingtin is completely inactive, despite the fact that the truncated protein product would be expected to contain the poly-glutamine region. The potential of this truncated protein to disrupt normal neuronal function should probably not be ignored. With this caveat in mind, these mice should provide insights into normal gene function, although their relevance to Huntington's disease pathophysiology — and in particular to the role of the tri-nucleotide expansion in changing the protein's function — is not intuitively clear.

Both the homozygous mutant mice (no good gene copy — complete loss of function) and the heterozygous mice (one good gene copy — 50% normal protein) had distinguishable phenotypes. Homozygotes, but not heterozygotes, showed embryonic lethality at the point of gastrulation, as assessed by morphologic study of genotypically determined embryos during development. This is indicative of a previously undefined role for huntingtin in early embryogenesis.

Most striking was the observation of cognitive impairment and focal neuronal deficiency in heterozygous mice (haplo-insufficient phenotype). The heterozygous mice were noted to be more sensitive to handling and were further assessed for neurological changes at both behavioral and anatomical levels. Four-month-old, sex-matched heterozygote knockout and wild-type control mice were subjected to a battery of tasks designed to measure several aspects of cognition, including: motor activity testing (spontaneous motor activity in a horizontal plane); T-maze alternation testing (short-term memory and discriminative learning); radial-arm maze testing (working memory); and the Morris water maze task (spatial information acquisition). Knockouts performed differently from the controls on two of the tests, showing increased spontaneous motor activity as well as difficulty unlearning previously learned behavior in the Morris water maze task. Morphometric histological comparison of heterozygous knockout and wild-type brain structures revealed a significant decrease in the total number of neurons in the subthalamic nucleus, as well as a suggestive decrease in the globus pallidus, in the knockouts. Histological hallmarks of human Huntington's disease, such as marked decrease in striatal neurons and reactive gliosis, were not features of the knockout brains. The combination of cognitive impairment and regionally specific neuronal deficits in heterozygous knockout mice argues strongly for cross-species conservation of a critical role for huntingtin in brain function.

What insights have these knockout mice provided into the human disorder and the issue of tri-nucleotide repeat disorders? Unfortunately, it seems that the mice have, so far, served only to deepen the enigmatic nature of Huntington's disease. Four observations seem relevant. First, the heterozygous mice have a phenotype that is not terribly dissimilar to humans heterozygous for the expansion mutation that causes Huntington's disease; this is not what one would expect when comparing a dominantly inherited (change-of-function?) disease like Huntington's to a knockout (loss of function), unless having 50 % of huntingtin is not enough for normal brain function

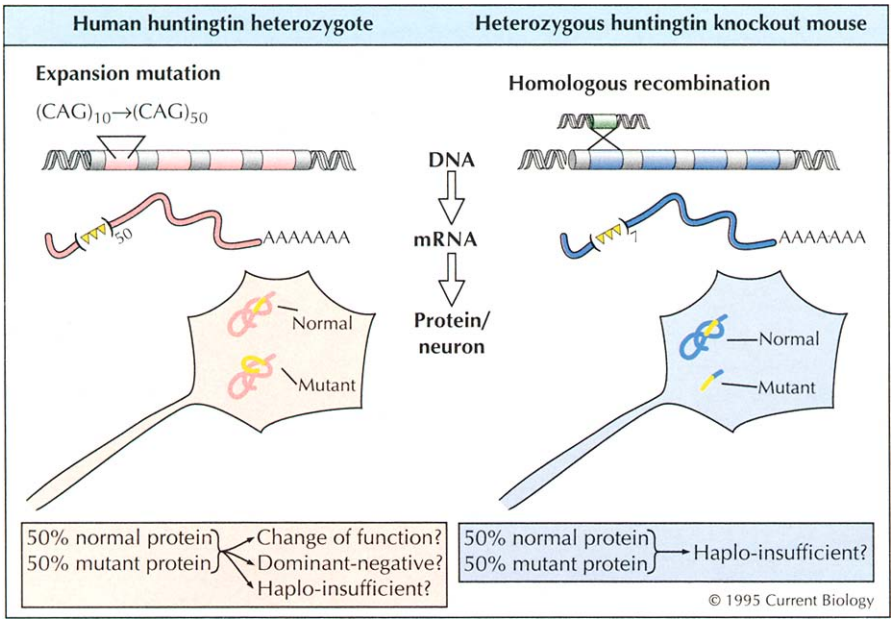

Fig. 1. Human Huntington's disease patients and the mouse knockout heterozygotes differ at the DNA, mRNA and protein levels. The mutant human huntingtin gene contains a CAG repeat expansion that results in the production of an expanded mRNA, as well as a protein containing an expanded poly-glutamine tract. The effect of the expansion on protein function, speculated to be a change of function, remains unknown. The mouse mutant huntingtin gene is disrupted at exon 5, resulting in the production of mRNA and a truncated protein assumed to have lost all function. Normal DNA alleles and mRNA products are not shown.

(haplo-insufficiency). Second, homozygous knockout mice die as embryos, whereas human Huntington's disease homozygotes are externally similar to heterozygotes, and there is no evidence for increased fetal demise [11,12]; this suggests that the excessive number of glutamines has little bearing on function in the human embryo, yet has a profound effect on huntingtin function in the adult nervous system. Third, young huntingtin-deficient homozygous or heterozygous mice seem to be phenotypically abnormal, whereas humans are rarely affected before mid-life, and then are faced with a progressive decline in abilities; this suggests a cumulative toxic effect of mutant huntingtin in humans, again consistent with a change of function as a result of the expansion mutation. Fourth, in heterozygous mice, the focus of the neuronal deficit is in the subthalamic nucleus and globus pallidus, with sparing of the caudate and putamen, whereas in heterozygous humans, the subthalamic nucleus can be relatively spared, while the caudate and putamen are heavily involved [13]; this could be due to species differences, but may also reflect the effects of a change of function in human mutant huntingtin (rather than loss of function, as in the mouse) that results in more widespread neuronal damage.

It seems that in the post-huntingtin-gene knockout era, three pathogenic mechanisms must be considered for the abnormal tri-nucleotide repeat in the huntingtin gene. First, and most probable, is a change of function of the huntingtin protein in Huntington's disease patients. Under this model, the change, due to an increased number of glutamines, has no bearing on fetal development. Second, a dominant-negative effect remains a possibility: the mutant huntingtin may form oligomers with the normal protein and destroy the function of an as-yet unidentified complex. Again, this requires that the dominant-negative effect does not occur during embryonic life. Finally, haplo-insufficiency through tri-nucleotide repeat-induced inactivation of a huntingtin domain critical for neural, but not embryonic, protein function should be considered. This would explain the phenotype observed in heterozygous mice (in homozygotes, the neural effects would never be seen, due to early lethality), as well as the apparent absence of any fetal effect in human homozygotes. Weighing in against this is the observation that an individual human exists who has only a single functioning huntingtin gene (loss-of-function of the other allele resulting from a chromosomal abnormality) and who is normal at nearly 50 years of age [14].

Clearly, sorting out these possibilities will require an understanding of the biochemistry and cell biology of huntingtin, and of the effect of the tri-nucleotide repeat on the gene, RNA and protein function. In this context, it is important to note that tri-nucleotide expansions in other diseases seem to exert their pathogenic effects by novel mechanisms. In fragile X, a common form of inherited mental retardation, the expanded CGG sequence causes hypermethylation of the gene's promoter, loss of transcription, and the expected loss-of-function in this X-linked recessive disorder [15]. In myotonic dystrophy, a highly clinically variable yet common muscular dystrophy, the expansion seems to disrupt RNA metabolism through precipitation of RNA molecules containing the expansion [16,17]. This dominantly inherited disorder has been proposed as the first 'dominant-negative RNA disorder' [17], although this needs to be investigated more thoroughly. Perhaps the next step towards elucidating the effect of the Huntington's disease triplet expansion will be the development of mice transgenic for the expanded gene in the contexts of normal and absent wild-type gene function. Maybe then the enigma of Huntington's disease will shrink to the point where a rational therapeutic strategy can be evolved.

References

1. Hoffman EP, Lehmann-Horn F, Rudel R: **Overexcited or inactive: ion channels in muscle disease.** *Cell* 1995, **80**:681–686.
2. Byers PH: **Disorders of collagen biosynthesis and structure.** In *The Metabolic Basis of Inherited Disease,* 7th edition. Edited by Scriver CR, Beaudet AL, Sly WS, Valle D. New York: McGraw-Hill; 1995: 4029–4077.
3. Brown MS, Goldstein JL: **A receptor-mediated pathway for cholesterol homeostasis.** *Science* 1986, **232**:34–37.
4. Folstein SE: *Huntington's Disease: A Disorder of Families.* Baltimore: The Johns Hopkins University Press; 1989.
5. Hoffman EP, Jaffurs D: **An expanding enigma.** *Curr Biol* 1993, **3**:456–459.
6. Guesella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, Tanzi RE, Watkins PC, Ottina K, Wallace MR, Sakaguchi MR, *et al.* : **A polymorphic DNA marker genetically linked to Huntington's disease.** *Nature* 1983, **306**:234–238.
7. Huntington's Disease Collaborative Research Group: **A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes.** *Cell* 1993, **72**:971–983.
8. Lin B, Nasir J, MacDonald H, Hutchinson GB, Graham RK, Rommens JM, Hayden MR: **Sequence of the murine Huntington disease gene: evidence for conservation, alternate splicing and polymorphism in a triplet (CCG) repeat.** *Hum Mol Genet* 1994, **3**:85–92.
9. Strong TV, Tagel DA, Vades JM, Elmer LW, Boehm K, Swaroop M, Kaatz KW, Collings FS, Albin RL: **Widespread expression of the human and rat Huntington's disease gene in brain and nonneural tissues.** *Nature Genet* 1993, **5**:259–265.
10. Gerber HP, Seipel K, Georgiev O, Hofferer M, Hug M, Rusconi S, Schaffner W: **Transcriptional activation modulated by homopolymeric glutamine and proline stretches.** *Science* 1994, **263**:808–810.
11. Nasir J, Floresco SB, O'Kusky JR, Diewert VM, Richman JM, Zeisler J, Borowski A, Marth JD, Phillips AG, Hayden MR: **Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes.** *Cell* 1995, **81**:811–823.
12. Fahn S: **Huntington disease.** In *Merritt's Textbook of Neurology,* 9th edn. Edited by Roland LP. Baltimore: Williams and Wilkins; 1995:695–699.
13. Oppenheimer DR, Esiri MM: **Diseases of the basal ganglia, cerebellum and motor neurons.** In *Greenfield's Neuropathology,* 5th edn. Edited by Adams JH, Duchen LW. London: Edward Arnold; 1992:988–1045.
14. Ambrose CM, Duyao MP, Barnes G, Bates GP, Lin CS, Srinidhi J, Baxendale S, Hummerich H, Lehrach H, Altherr M *et al.*: **Structure and expression of the Huntington's disease gene: evidence against simple inactivation due to an expanded CAG repeat.** *Somat Cell Mol Genet* 1994, **20**:27–38.
15. Pieretti M, Zhang F, Fu YH, Warren ST, Oostra BA, Caskey CF, Nelsen DL: **Absence of expression of the FMR-1 gene in Fragile-X syndrome.** *Cell* 1991, **66**:817–822.
16. Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH: **Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues.** *J Cell Biol* 1995, **128**:995–1002.
17. Wang J, Pegoraro E, Menegazzo E, Gennarelli M, Hoop RC, Angelini C, Hoffman EP: **Myotonic dystrophy: evidence for a possible dominant-negative RNA mutation.** *Hum Mol Genet* 1995, **4**:599–606.

W. Gregory Feero and Eric P. Hoffman, Departments of Molecular Genetics and Biochemistry, Human Genetics, and Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261, USA.