Huntington's Disease Gene Hunters: An Expanding Tale

MacDonald ME. *A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes.* Cell 1993;72:971–983.

It is 28 years since the Huntington's Disease (HD) gene and mutation were identified and published in *Cell* by the Huntington's Disease Collaborative Research Group (HD-CRG; Fig. 1A).[1] The genetic defect causing HD had been assigned to chromosome 4 in 1983 in one of the first successful linkage analyses using polymorphic DNA markers in humans,[2] but it took another ten years to pinpoint the gene and determine the mutation. The long lag was largely because this research was conducted before the human genome was mapped, and was the culmination of a painstaking process involving repeatedly refining the location of the gene, based on locating markers and cloning transcripts from the genome across six independent laboratories. The nature of the genetic mutation—an expanded CAG repeat sequence—was also instrumental in the resolution of this detective story. Expanded repeats in DNA had already been associated with several diseases that had features in common with HD, such as genetic anticipation, including fragile X syndrome,[3] spinal and bulbar muscular atrophy,[4] and myotonic dystrophy.[5–7] This meant the HD-CRG were actively looking for length mutations that segregated with disease that might indicate the presence of an expanding repeat tract. As we enter a new phase of HD research, with the advent of trials of potential disease-modifying treatments, it seems a good time to reflect on the legacy of the HD-CRG publication.

Understanding the precise nature of the mutation started to throw light on a number of important aspects of disease phenotype. In the HD-CRG paper, the authors commented on a possible link between CAG repeat number and age of onset, noting the very long repeats seen in juvenile cases. They also noted instability of the repeat when transmitted to the next generation as underlying the phenomenon of genetic anticipation in which subsequent generations can have progressively earlier disease onsets, and demonstrated the link to paternal transmission, with the longest expansions see in juvenile cases where transmission occurred through the paternal line. It raised the possibility of de novo mutations, noting inter-generational expansion from an allele at the upper end of the normal range, into the disease range in two families in which new mutations had been assumed to have arisen.

The HD-CRG paper was based on data from 75 HD families from a range of ethnic backgrounds, particularly the large Venezuelan kindreds: it should be recognized that collecting these families was a herculean task that took many years.[8] It reported that the Huntington's gene (*HTT*) was large and encoded a previously undescribed protein, which they named huntingtin (HTT), and which was widely expressed in many body tissues despite the major impact of HD being in specific regions of the brain. The expanded CAG repeat was located in exon 1 of a 67 exon transcript. Follow up studies established that age at onset of HD was significantly correlated with the length of the CAG tract in *HTT*.[9–11] Recognizing this relationship has helped to accelerate our understanding of how the molecular genetics relates to key aspects of disease phenotype.

The publication of the gene sequence was rapidly followed by the availability of a genetic test which revolutionized pre-clinical testing by eliminating the need for complicated linkage analysis and provided near 100% accuracy (full penetrance for expansions of 40 and above, with incomplete penetrance for repeats of 36–39). This has made confirmation of diagnosis in symptomatic individuals easier and also raised awareness of the HD phenocopy conditions, the genetic basis for many of these conditions having now been established. Importantly, the gene test allowed predictive testing of an at-risk individual, without the need for samples from multiple family members that had previously been a limitation for testing based on linkage analysis. The work done on developing protocols for presymptomatic testing in HD provided a route map for testing in other, less common, disorders.[12] Knowledge of the mutation also makes uterine and pre-implantation testing feasible; although the uptake of these options remain low,[13] they can be life-changing for some families by removing the risk of HD in the next generation.

The value of being able to diagnose the condition with certainty in life also has profound implications for clinical research. The emergence of a reliable genetic test underpinned a number of important large longitudinal observational studies, starting with the North American's Huntington's Study group's Cooperative Huntington Observational Research Trial (COHORT),[14] followed by the European HD network

REGISTRY study,[15] both of which were superseded by Enroll-HD.[16] Enroll-HD is a global multicenter study with clinical sites in Europe, North America, Australasia and Latin America that provides systematic longitudinal clinical phenotype assessment and a linked biobank for gene positive individuals across all stages of HD (ranging from individuals with no symptoms and no clinical indicators of disease through to those with clear clinical signs). There are now over 20,000 participants in Enroll-HD and, along with the data collected through REGISTRY and COHORT, it has provided data and biosamples to address a wide range of questions about the clinical phenotype and the biology of HD, in addition to providing important data for developing biomarkers, clinical outcome measures and working out disease-staging parameters for clinical trials. Collecting large numbers of participants with a rare disease requires international cooperation and multicenter studies, both of which would likely have been too difficult and too expensive in the absence of a readily available genetic test.

The availability of a genetic test, and knowledge that the mutant gene is fully penetrant, had an even more profound effect on the capacity to study human HD before the onset of overt disease, by allowing the identification of individuals carrying the gene before they develop clinical disease. In addition to numerous smaller studies of premanifest individuals, there have been two well-powered longitudinal premanifest studies; PREDICT-HD[17] and TRACK-HD.[18] PREDICT-HD commenced in 2001 and set out to prospectively characterize clinical, neurobiological, and behavioral markers occurring before clinical diagnosis of HD. Less than a decade later PREDICT-HD was followed by Track-HD which also aimed to detect clinical and biological markers of HD progression, but with a focus on enabling power calculations and defining the most powerful outcome measures for future trials in participants with very early disease or before formal diagnosis of disease onset. It has also been possible to establish a cohort of "far from onset" young premanifest HD gene carriers, which is starting to identify the very earliest brain changes in HD, more than 20 years before clinical diagnosis, and will be invaluable in defining the earliest stages at which intervention with disease-modifying strategies may be useful.[19,20]
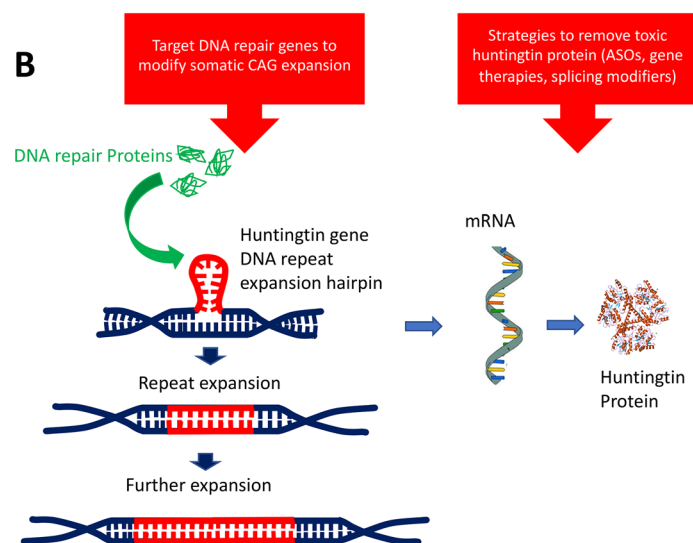
The discovery of the gene also enabled the generation of genetically accurate models in cells and animals. Models have been generated in species from yeast through to large animals and many of these models have offered useful insights into the pathophysiology of HD.[21–24] Most widely used have been mouse models, the first of which, the R6 series, were generated by Gillian Bates.[25] Her models provided the seminal observation that the HTT protein containing an expanded polyglutamine tract formed the characteristic pathological hallmark of insoluble cellular aggregates in mouse brain, leading to their detection in human brain[26]—although in fact these had been previously observed in human HD brain without their importance being understood.[27] The R6 model series was generated using a human *HTT* exon 1 transgene, with expansions of 115–150 CAG and ~1000 bp of 5' sequence, expressed from the human promoter. They represent one of many transgenic models with truncated fragments of *HTT* expressed from various promoters.

Transgenic models containing full length HTT either as a cDNA or the full human sequence in an artificial chromosome
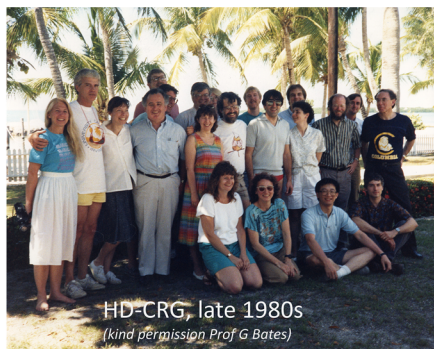
FIG. 1. (A) Members of the Huntington's Disease Collaborative Research Group (HD-CRG), taken at a workshop in Islamorada in the late 1980s (*kindly supplied by Professor G Bates*). (B) The discovery of the gene for HD has led to multiple potential disease-modifying targets, such as targeting DNA repair proteins with the aim of reducing CAG expansion in the brain, and strategies to reduce levels of huntingtin protein, which aims to reduce the toxic effects of the mutant protein. Fig1 A supplied by Professor Gillian Bates, UCL, London UK.

and knock-in models with human-mouse chimaeric sequences of various types have also been generated. Such models have been instrumental in both understanding the biology of HD and in the search for novel therapeutics.[28] Most recently, human stem cells derived directly from patients have been developed yielding further insights into HD pathogenesis.[29–31]

The HD-CRG paper also discussed two key areas of understanding which have underpinned disease-modifier research that are currently in clinical translation. The first of these is huntingtin lowering therapies (Fig. 1B). The paper was able to confirm the previous observation that homozygotes for the disease allele did not differ clinically from heterozygotes, suggesting that HD results from a gain-of-function mutation. They considered gain of function mechanisms "in which either the mRNA product or the protein product of the disease allele would have some new property or would be expressed inappropriately" and suggested that the expanded trinucleotide repeats could be translated to an increased poly-glutamine stretch near the N-terminus—something that has since been proven to be the case. Although loss of function mechanisms almost certainly play some role in the pathophysiology of HD,[32] the concept of the mutant protein having gain-of-function toxic properties has underpinned the current predominant therapeutic strategy of huntingtin lowering; aiming to reduce the burden of toxic huntingtin protein. Huntingtin lowering in animal models has been shown very clearly to reduce the disease load in the brain and improve function in transgenic HD animals.[33,34] Various huntingtin lowering strategies are currently in, or close to, clinical trials. These use multiple mechanistic approaches including ASOs interfering with the translation of huntingtin mRNA, delivered into the CSF,[35] microRNA delivered directly into the striatum using viral vectors, with DNA editing and agents interfering with mRNA splicing currently approaching clinical trial.[36]

A second disease modifying approach that depends on understanding the genetics is that of targeting post-mitotic expansion of the CAG repeat in the brain (Fig. 1B). The large observational and prospective studies of Predict, Registry and Enroll-HD have provided large numbers of participants' DNA and matched clinical data for genome-wide association studies (GWAS) aimed at detecting genetic modifiers of HD. As was outlined by the HD-CRG paper, as well as being unstable when transmitted to the next generation, the HD trinucleotide repeat is unstable in the absence of recombination or replication. It has since been demonstrated that CAG repeat number continues to increase in certain tissues, in particular the striatum; and the greater the rate of expansion, the lower the age at disease onset.[37] Identification of the gene and its mutation allowed further genetic research looking for modifiers of age at onset of disease, which has now identified that DNA repair is central in this process, flagged previously in a number of animal studies but not known to be central in human disease onset.[38] This raises the exciting possibility that targeting DNA repair proteins could reduce HTT CAG expansion and may be capable of slowing the disease progression.

Finally, but critically, clinical trials are much more powerful when a condition can be diagnosed with certainty using a reliable and straightforward test, as has been available for HD since 1994. This is especially important for a rare disease where multicenter studies are necessary. Over the last few years, trials of potential disease modifiers of HD have gradually moved to recruiting patients who are earlier in the disease course, and it is likely that this will soon extend to recruitment of patients prior to the onset of motor symptoms; something that would have been extremely difficult, if not impossible, prior to identification of the gene. In summary, the earlier painstaking work of the teams involved in identifying the HD gene have set in motion a whole field of discovery, and hopefully this will be eventually rewarded by identification of a disease-modifying treatment for this devasting condition.

Author Roles

1) Research project: A. Conception, B. Organization, C. Execution; 2) Statistical Analysis: A. Design, B. Execution, C. Review and Critique; 3) Manuscript: A. Writing of the first draft, B. Review and Critique.

AER: 3A, 3B
LJ: 3A, 3B

Disclosures

Anne E. Rosser MA, PhD, MB.BChir,[1,2,3]
Lesley Jones BSc, PhD[1]
[1]*MRC Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff*

University School of Medicine, Cardiff, United Kingdom, ²Cardiff Brain Repair Group, Cardiff University School Biosciences, Cardiff, United Kingdom; ³Wales Brain Research And Intracranial Neurotherapeutics (BRAIN) Unit, Cardiff, United Kingdom

References

1. MacDonald ME. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 1993;72:971–983.

2. Gusella JF, Wexler NS, Conneally PM, et al. A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* 1983;306: 234–238.

3. Fu Y et al. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 1991;67: 1047–1058.

4. La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 1991;352:77–79.

5. Harley H, Brook JD, Rundle SA, et al. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 1992;355:545–546.

6. Buxton J, Shelbourne P, Davies J, et al. Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* 1992;355:547–548.

7. Aslanidis C, Jansen G, Amemiya C, et al. Cloning of the essential myotonic dystrophy region and mapping of the putative defect. *Nature* 1992;355:548–551.

8. Wexler NS. Three decades of caring for the Venezuelan Huntington's disease families. *Lancet Neurol* 2013;12:738.

9. Duyao M, Ambrose C, Myers R, et al. Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nat Genet* 1993;4: 387–392.

10. MacMillan JC, Snell RG, Tyler A, et al. Molecular analysis and clinical correlations of the Huntington's disease mutation. *Lancet* 1993;342: 954–958.

11. Andrew SE, Paul Goldberg Y, Kremer B, et al. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington;s disease. *Nat Genet* 1993;4:398–403.

12. Harper PS, Lim C, Ten Craufurd D. years of presymptomatic testing for Huntington's disease: The experience of the UK Huntington's Disease Prediction Consortium. *J Med Genet* 2000;37:567–571.

13. Schulman JD, Stern HJ. Low utilization of prenatal and pre-implantation genetic diagnosis in Huntington disease—risk discounting in preventive genetics. *Clin Genet* 2015;88:220–223.

14. Wang Y, Garcia TP, Ma Y. Nonparametric estimation for censored mixture data with application to the cooperative Huntington's Observational Research Trial. *J Am Stat Assoc* 2012;107:1324–1338.

15. Orth M, Handley OJ, Schwenke C, et al. Observing Huntington's disease: The European Huntington's disease network's registry. *PLoS Curr* 2010;2:RRN1184. https://doi.org/10.1371/currents.RRN1184.

16. Landwehrmeyer GB, Fitzer-Attas CJ, Giuliano JD, et al. Data Analytics from Enroll-HD, a Global Clinical Research Platform for Huntington's Disease. *Mov Disord Clin Pract* 2017;4:212–224.

17. Paulsen JS, Hayden M, Stout JC, et al. Preparing for preventive clinical trials: the predict-HD study. *Arch Neurol* 2006;63:883–890.

18. Tabrizi SJ, Langbehn DR, Leavitt BR, et al. Biological and clinical manifestations of Huntington's disease in the longitudinal TRACK-HD study: cross-sectional analysis of baseline data. *Lancet Neurol* 2009; 8:791–801.

19. Johnson EB, Parker CS, Scahill RI, et al. Altered iron and myelin in premanifest Huntington's Disease more than 20 years before clinical onset: Evidence from the cross-sectional HD Young Adult Study. *EBioMedicine* 2021;65:103266.

20. Scahill RI, Zeun P, Osborne-Crowley K, et al. Biological and clinical characteristics of gene carriers far from predicted onset in the Huntington's disease Young Adult Study (HD-YAS): a cross-sectional analysis. *Lancet Neurol* 2020;19:502–512.

21. Rivetti di Val Cervo P, Besusso D, Conforti P, Cattaneo E. hiPSCs for predictive modelling of neurodegenerative diseases: dreaming the possible. *Nat Rev Neurol* 2021;17:381–392.

22. Crook ZR, Housman D. Huntington's Disease: Can Mice Lead the Way to Treatment? *Neuron* 2011;69:423–435.

23. Bates GP, Dorsey R, Gusella JF, et al. Huntington disease. *Nat Rev Dis Prim* 2015;1:15005.

24. Howland D, Ellederova Z, Aronin N, et al. Large Animal Models of Huntington's Disease: What We Have Learned and Where We Need to Go Next. *J Huntington Dis* 2020;9:201–216.

25. Mangiarini L, Sathasivam K, Seller M, et al. Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* 1996;87:493–506.

26. DiFiglia M, Sapp E, Chase KO, et al. Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* 1997;277:1990–1993.

27. Roizin L, Stellar S, Willson N, Whittier J, Liu JC. Electron microscope and enzyme studies in cerebral biopsies of Huntington's chorea. *Trans Am Neurol Assoc* 1974;99:240–243.

28. Bates G, Landles C. Huntington's disease. In: Tabrizi SJ, Jones L, eds. *Huntington's Disease.* UK: (Oxford University Press (OUP); 2014:410–461. https://doi.org/10.1093/med/9780199929146.003.0016.

29. HD iPSC Consortium & The HD iPSC Consortium. Induced Pluripotent Stem Cells from Patients with Huntington's Disease Show CAG-Repeat-Expansion-Associated Phenotypes. *Cell Stem Cell* 2012; 11:264–278.

30. Victor MB, Richner M, Olsen HE, et al. Striatal neurons directly converted from Huntington's disease patient fibroblasts recapitulate age-associated disease phenotypes. *Nat Neurosci* 2018;21:341–352.

31. Ruzo A, Croft GF, Metzger JJ, et al. Chromosomal instability during neurogenesis in Huntington's disease. *Development* 2018;145: dev156844.

32. Saudou F, Humbert S. The Biology of Huntingtin. *Neuron* 2016;89 (5):910–26.

33. Yamamoto A, Lucas JJ, Hen R. Reversal of neuropathology and motor dysfunction in a conditional model of Huntington's disease. *Cell* 2000;101:57–66.

34. Kordasiewicz HB, Stanek LM, Wancewicz EV, et al. Sustained therapeutic reversal of Huntington's disease by transient repression of huntingtin synthesis. *Neuron* 2012;74:1031–1044.

35. Tabrizi SJ, Leavitt BR, Landwehrmeyer GB, et al. Targeting Huntingtin Expression in Patients with Huntington's Disease. *N Engl J Med* 2019;380:2307–2316.

36. Tabrizi SJ, Flower MD, Ross CA, Wild EJ. Huntington disease: new insights into molecular pathogenesis and therapeutic opportunities. *Nat Rev Neurol* 2020;16:529–546.

37. Swami M, Hendricks AE, Gillis T, et al. Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Hum Mol Genet* 2009;18:3039–3047.

38. Wheeler VC, Dion V. Modifiers of CAG repeat instability: insights from model systems. *J Huntingtons Dis* 2021;10:123–148. https://doi.org/10.3233/JHD-200426.