

Biochemistry 118Q – Bioinformatics and Genomics  
Professor Doug Brutlag  
Final Paper

Denise Woo  
3/8/1998

## Huntington's disease, an example of expanded glutamine repeats in neurodegeneration -- current situation and new possibilities

### Introduction

Tandem repeats, that is simple sequence repeats, occur commonly in the human genome, and they have long been used as markers in linkage studies. In this decade, it has also been found that tandem repeats underlie an entirely new class of human mutations. The expansion of a group of trinucleotide repeats is now known to cause several inherited diseases, all of which are neurological disorders. These trinucleotide repeat diseases share several common features: 1. Inheritance is autosomal dominant or X-linked. 2. Disease severity correlates directly with increasing repeat length. 3. Expansions arise from CG-rich triplet repeats that are polymorphic in the normal population. 4. The expanded trinucleotide repeats are unstable and change in size when transmitted to successive generations. 5. The diseases show anticipation, that is, disease severity increases in successive generations of a family. 6. Parent-of-origin effects are common, where the most severely affected individuals usually inherit the disease from the father.

These diseases can be further classified into two main groups: the first type are those caused by small, constrained CAG repeat expansions that encode polyglutamine in the disease gene protein. This group currently includes eight neurodegenerative diseases: Huntington's disease (HD), dentatorubral-pallidoluysian atrophy (DRPLA), spinocerebellar ataxia (SCA-1, SCA-2, SCA-3, SCA-6, SCA-7), and spinal and bulbar muscular atrophy (SBMA). The mutations involved are all believed to alter one or more properties of the disease protein, leading to a gain of function that is deleterious to neurons. The second group of diseases are multisystem disorders that are caused by much larger expansions of non-CAG repeats that fall outside the protein coding region, and these include myotonic dystrophy (DM) and Fragile X Syndrome (FRAXA and FRAXE). The underlying mechanism of this type of diseases appears to be altered expression of the gene.

In this paper, the focus will be on the Type I diseases associated with expanded glutamine repeats, and particular attention will be paid to Huntington's disease as an example. The characteristics of the disease will first be discussed, followed by proposed molecular mechanisms that may underlie the neuropathology of the disease, and finally, current diagnostic and therapeutic techniques, and also prospective advances using genomics and bioinformatics.

### Features of polyglutamine diseases

Although the eight polyglutamine diseases are caused by genes that have no homologies with each other except for the CAG repeats, they are similar in several important aspects, besides those common to all expanded trinucleotide repeat diseases: 1. Cell death is neuron specific even though the repeat-containing genes are widely expressed in non-neuronal cells. 2. Pathogenesis is at the protein level, since in all cases the disease gene is expressed to give the protein product. 3. Distinct sets of neurons die in each disorder, even though the disease gene product is expressed throughout the brain. 4. The gene products are all intracellular proteins, suggesting that the effect of the disease gene product is exerted intracellularly. 5. Degeneration occurs slowly, with little abnormality before clinical onset of the disease.

### Huntington's disease (HD) – clinical and genetic aspects

HD is the most common and most studied of the polyglutamine disorders, with an incidence of 1: 10,000. It is an autosomal dominant neurodegenerative disorder, and clinical symptoms compose a three-part picture with motor symptoms – characterized by choreic movements and bradykinesia, psychiatric disturbances -- with aggressivity and depression, and cognitive impairment. These symptoms usually appear in the third to fifth decades, but may begin in childhood or after age 50. Patients die on average within 17 years of onset of the disease, affected by profound dementia.

HD illustrates the extent of selective neuronal vulnerability in the polyglutamine disorders. Degenerative changes are most marked in the striatum, particularly in the caudate nucleus and putamen. There is a severe loss of medium spinal projection neurons, while several populations of large interneurons are relatively spared. Cell death follows a gradient in three axes, occurring first in the tail of the caudate nucleus, then proceeding mediolaterally through the caudate and dorsolaterally through the putamen. In advanced states, neuronal loss becomes widespread, involving the cortex and cerebellum.

In 1983, the HD gene was mapped to chromosome 4 using positional cloning, but it was not until ten years later in 1993 that the defective gene was identified by Huntington's Disease Collaborative Group. The defect was found to be an expanded CAG repeat within the coding region of a large gene (IT15) which encodes the protein named huntingtin, whose function is yet unknown. The CAG repeat encodes a polyglutamine tract that is 11 to 34 amino acids long in normal individuals, but 36 to 121 in HD patients. In addition, there was found to be a smaller polymorphic CGG repeat lying immediately 3' to the CAG repeat, encoding a polyproline stretch.

### Molecular mechanisms of polyglutamine diseases

The mechanism of neurodegeneration in these diseases is believed to be the same and yet is still unknown, and the main hypothesis is that the polyglutamine expansions lead to a gain of function that may be toxic to neurons. This hypothesis is supported by the dominant inheritance of these disorders, and the fact that the disease gene is always expressed in the disease state. Moreover, individuals missing one copy of the normal HD gene do not develop HD, and transgenic mice models have also provided evidence for the gain of function mechanism.

Three mechanisms have so far been proposed. The first is the aberrant

functionally important in regulatory proteins. This model proposes that expanding the polyglutamine tract in a regulatory protein alters the strength of its interactions with other coregulatory proteins, thereby disrupting the orderly regulation of a molecular cascade. As a result, neurons receive inappropriate signals, resulting in neuronal apoptosis. The specific neurotoxicity may be due to the fact that only one or more regulatory cascades that are neuron specific are disturbed.

In the protein aggregation model, it is believed that expanded polyglutamine tracts tend to form an extended polar zipper involving tight noncovalent interactions between proteins, since glutamine is a polar amino acid with an amide group at the end of its side chain. These interactions can be nonproductive ones that alter the function of any involved proteins and interfere with the degradation of the aggregated proteins. The resulting accumulation of the insoluble proteinaceous material may be deleterious to neurons, and the long-lived postmitotic nature of neurons makes them particularly susceptible to the accumulation of the aggregated protein.

The third model is the transglutamine cross-linking model. It proposes that expanded polyglutamine tracts become active substrates for transglutaminase activity, leading to the cross-linking of proteins within neurons. Either the cross-linked proteins themselves or the insoluble isodipeptide produced during subsequent proteolysis may be toxic to cells. As in the protein aggregation model, long-lived postmitotic cells like neurons will be especially susceptible to the accumulation of this cross-linked material.

No evidence however, has yet been found for the presence of any aggregated, precipitated or cross-linked protein. To fully understand the molecular mechanism, more research needs to be done using transgenic animal models, in vitro expression systems, as well as biochemical analyses of the normal and expanded proteins.

### Diagnosis of Huntington's disease

Ever since 1987, linked DNA markers have been used for presymptomatic testing of HD, but the need to test many family members, together with inaccuracies from possible recombination, made testing very difficult. The mapping of the HD gene in 1993 allowed development of a polymerase chain reaction (PCR)-based repeat test that is highly sensitive and specific, and is also technically simpler and cheaper than the linkage analysis, so that presymptomatic testing for HD is now readily available.

In the repeat test, blood samples are obtained from the patient, and DNA is extracted from peripheral blood leucocytes. PCR is then used to amplify the region containing the CAG repeat polymorphism. A repeat length of 38 or greater is in general considered to be disease related.

Besides direct testing of adults, this method also enables prenatal diagnosis. This can be done by obtaining single cells from the embryo using amniocentesis, followed by fluorescent PCR using polymorphic DNA probes linked to the HD locus. Using this method, it is possible to determine the at-risk status of the fetus with 96% accuracy. However, prenatal diagnosis of HD is still very controversial, since many argue that selective abortion is unacceptable. Moreover, the test result may also have implications for the parents and their other offspring who may have decided not to be tested, but as a result of testing of the fetus know they may be at risk. In addition, there

termination, and the known carrier status may have an important impact on the child's later life.

Non-disclosure pre-implantation genetic diagnosis (PGD) offers an alternative to prospective parents who are at high risk of carrying HD but do not want to incur the emotional and social burdens resulting from the presymptomatic disclosure of their own carrier status. In this method, couples are offered the option of having IVF with preimplantation biopsy and testing of their embryos for HD without ever being informed of the test results, and only embryos found not to carry the affected gene are used in transfer.

Another option is exclusion PGD testing, in which rather than direct testing, the grandparental allele is tested for by linkage, and embryos which contain the allele will be excluded for transfer, although there is only a 50% chance that these embryos will contain the affected allele.

The advantage of PGD over direct testing is that first, since only embryos found not to carry the affected gene are used for transfer to the uterus, couples need not face the difficult decision about terminating a fetus. Second, parents do not need to be burdened by information regarding their own carrier status. However, the main disadvantage of PGD is the low success rate of obtaining a pregnancy. Other disadvantages include cost, safety for the mother, and possible inaccuracy of the test results from single-cell PCR. Moreover, there are inherent difficulties in maintaining the test results secret, such as when no unaffected embryos are available.

### Therapy for Huntington's disease

Up till now, there has been no treatment which successfully halts the progression of HD, and only symptoms have been able to be treated. Drugs have been used to control excessive involuntary movement, to reduce depression and to combat obsessive behavior. Patients are currently treated in a team management approach, where each team includes a neurologist, a dietician, a genetic counselor, a neuropsychologist, an occupational therapist, a physical therapist, a psychologist, a social worker and a speech pathologist, who provide care and assistance not only to the patient but to family members as well.

Meanwhile, research is underway regarding therapeutic techniques against HD, but since nothing is known about the way the mutant protein huntingtin causes degeneration of the neurons, current therapeutics being investigated rely upon the two general features of the disease. First, HD is related to a localized neurodegeneration in the striatum, and second, this neuronal loss is progressive over years. This has led to two basic therapeutic strategies. The first is to substitute missing neurons in patients by homologous neurons that can replace them functionally and anatomically, and current studies are based upon the use of intracerebral grafting of fetal neural tissue. The second method is to strengthen the natural defense of neurons and allow threatened neurons to survive longer by using neuroprotective drugs or trophic factors. These two strategies in fact often complement each other, since though neural transplants may substitute for lost neurons, they cannot protect host neurons from degeneration, and while neurotrophins can preserve neurons that are not yet lost, they cannot revive dead neurons and recover function.

For the first method, the use of intrastriatal transplantation of fetal striatal tissue has been tested out in animal models, and results show that full neural circuit reconstruction in the striatum can be obtained, at least partially. Afferent axons have been shown to persist for a long period of time in an area of neurodegeneration, and axonal regeneration has been demonstrated after grafting. Regenerated connections are functional, and fine structural analysis confirmed the establishment of appropriate synaptic contacts. There was corresponding improvement in motor behavior, though the effect on cognitive functions is unknown since these are difficult to study in the rat models.

For neuroprotective therapy, different trophic factors are being tested in animal models on their ability to protect specific neuronal populations against natural cell death during fetal and post natal development. However, since all of these molecules are proteins, there are several inherent difficulties involved in their therapeutic use. First, they are difficult to produce in large amounts and genetic engineering would be necessary. Second, they have a wide array of action and so may lead to unwanted alteration of cellular activity in peripheral organs. Also, as proteins, they may lead to immune responses when identified as foreign substances. Moreover, the instability of the molecules as well as their potential inactivation by binding to plasmatic proteins has to be considered. Lastly, the blood-brain barrier acts as a very tight filter against the transfer of proteins, so only a very small proportion of the administered protein will be able to reach neural target cells.

One of the main problems to be solved in this method concerns the administration of the neurotrophic drug to the desired area in the brain. A mechanical mini-pump is used in delivering drugs to the ventricles, but neurotrophic drugs are different in that they are proteins and hence larger, and proteins also tend to interact with each other and so may lead to clogging. Another method is using biological mini-pumps, which are based on the intracerebral transfer of the gene encoding the protein of interest. In vivo gene transfer makes use of vectors, and suitable vectors have yet to be found, since neutral vectors lack efficiency of transduction, while for viral vectors, retroviruses cannot be used due to the post-mitotic nature of neural cells, while other vectors require maintaining a large portion of the viral genome and hence are less safe. Moreover, the area reached by these vectors is generally very limited. On the other hand, ex vivo gene transfer relies upon the intracerebral injection of previously genetically engineered cells. The cells selected for this purpose would need to proliferate actively in vitro, survive intracerebral transplantation, be harmless to the surrounding tissue, and be able to indefinitely and massively synthesize the product of the transgene. As yet, none of the cells analyzed have fulfilled all of these prerequisites.

Two main cell types are currently being developed for this ex vivo gene transfer: cutaneous fibroblasts and myoblasts, but problems remain regarding their long-term survival. Another option is the use of tumoral cell lines, by macroencapsulating the engineered cells in a membrane that allows nutrients, oxygen and the product to circulate, but prevents the cells from invading the host tissue. For added protection, a genetic construct can be introduced to destroy the implanted cells if side effects are observed. This technique appears to be a relatively safe method of administering trophic factors, yet the therapeutic value of delivering a trophic factor has still to be demonstrated.

Recently, it has been suggested that Sertoli cells can serve as a vehicle to co-transplants of fetal tissue, since they can express many nutritive, regulatory, trophic and immunosuppressive factors, which will be able to enhance the recovery benefits associated with the fetal neural cell transplant. This is based on the fact that Sertoli cells can provide a microenvironment for the development of germ cells, which are highly antigenic, in the testis. Studies have so far been done on rat models, and it was found that co-transplanted neurons had more extensive neuritic outgrowth and were larger than neurons in the control group. However, obtaining human Sertoli cells would rely on an already over-burdened donor system to supply pre-pubertal cells. Thus, xenografts may be a more realistic strategy, though many problems still need to be addressed before animals can be used as a tissue source for human transplantation. This does seem a promising form of therapy, yet much research remains to be done.

### Applications of genomics and bioinformatics in Huntington's disease

Genomics has led to giant strides forward through the development of molecular diagnostics for HD using mutation testing. However, the availability of presymptomatic testing has led to both positive and negative effects, for though testing has lifted the shadow under which many people have lived for years, the lack of treatment available for the disease has also to many ethical issues regarding testing. The most serious is the possibility of testing without giving information, counseling and support. Also, though there is some correlation between the repeat size and the age of onset, disease severity also depends on other factors such as genetic background, and the range of variation is too great to be useful for predictions of the age of onset. Testing also brings the possibility of stigmatization in employment, insurance and personal relationships, and as mentioned, direct mutation testing may disclose information concerning the carrier status of other family members who did not wish to be tested. It remains to be seen how such potential conflicts of interest will be resolved.

One of the major obstacles in the development of a treatment for HD is the lack of knowledge regarding the function of the huntingtin protein and how it leads to neurodegeneration. Bioinformatics may be useful by helping to determine the function of the protein through sequence information, while the interactions of huntingtin with other proteins may be elucidated through protein-protein docking, leading to the unraveling of the molecular mechanism of the disease, which may be applicable to all polyglutamine diseases. Once the mechanism has been determined, therapy targeted at halting the events leading to neurodegeneration can be developed, so that many of the present dilemmas will disappear, and true prevention of Huntington's disease may at last become a reality.

### Conclusion

Experimental research in the pathogenesis of HD is still in its infancy, and it is likely that entirely novel therapies will result from research on the basic disease mechanism. A lot of questions remain to be answered, but the tools available in the 21<sup>st</sup> century are vastly improved over those available to previous generations, and this should provide both physicians and patients with enthusiastic hope for the future.

## REFERENCES

- Bachoud-Levi, A.C. et. al. Prospectives for cell and gene therapy in Huntington's disease. *Prog. Brain Res.* 1998. 117: 511-24.
- Bates, G, Expanded glutamines and neurodegeneration – a gain of insight. *Bioessays.* 1996. 18(3): 175-8.
- Braude, P.R. et. al. Non-disclosure preimplantation genetic diagnosis for Huntington's disease: practical and ethical dilemmas. *Prenat. Diagn.* 1998. 18(13): 1422-6.
- Harper P.S. Clinical consequences of isolating the gene for Huntington's disease. *BMJ.* 1993. 307 (6901) : 397-8.
- Jones, A.L. et. al. Huntington disease: advances in molecular and cell biology. *J. Inherit. Metab. Dis.* 1997. 20(2): 125-38.
- Lindvall, O. Prospects of transplantation in human neurodegenerative diseases. *Trends Neurosci.* 1991. 14(8): 376-84.
- MacMillan, J.C. et. al. Molecular diagnostic analysis for Huntington's disease: a prospective evaluation. *J. Neurol. Neurosurg. Psychiatry.* 1995. 58(4): 496-8.
- Nance, M.A. Huntington disease: clinical, genetic and social aspects. *J. Geriatr. Psychiatry Neurol.* 1998 Summer. 11(2): 61-70.
- Paulson, H.L. and Fischbeck, K.H. Trinucleotide repeats in neurogenetic disorders. *Annu. Rev. Nerosci.* 1996. 19: 79-107.
- Ross, C.A. et. al. Pathogenesis of neurodegenerative diseases associated with expanded glutamine repeats: new answers, new questions. *Prog. Brain Res.* 1998. 117: 397-419.
- Sermon, K. et. al. Preimplantation diagnosis of Hungtinton's disease: clinical application and analysis of the HD expansion in affected embryos. *Prenat. Diagn.* 1998. 18(13): 1427-36.
- Willing, A.E. et. al. Sertoli cell transplants: their use in the treatment of neurodegenerative disease. *Mol. Med. Today.* 1998. 4(11): 471-7.