BBF **BEHAVIORAL AND BRAIN FUNCTIONS**

SHORT PAPER   **Open Access**

Correlation of CAG repeat length between the maternal and paternal allele of the Huntingtin gene: evidence for assortative mating

Peg Nopoulos[1,2,3*], Eric A Epping[1], Tom Wassink[1], Bradley L Schlaggar[4,5,6,7] and Joel Perlmutter[4,6,7,8]

Abstract

Triplet repeats contribute to normal variation in behavioral traits and when expanded, cause brain disorders. While Huntington's Disease is known to be caused by a CAG triplet repeat in the gene Huntingtin, the effect of CAG repeats on brain function below disease threshold has not been studied. The current study shows a significant correlation between the CAG repeat length of the maternal and paternal allele in the Huntingtin gene among healthy subjects, suggesting assortative mating.

Introduction

Huntington's Disease (HD) is a neurodegenerative disorder caused by a triplet repeat expansion of the gene Huntingtin (*HTT*, OMIM 613004). Triplet repeats are an example of simple sequence repeats (SSRs) which are scattered throughout the genome and can increase or decrease in length between generations. Triplet repeats can be referred to as 'dynamic mutations' and they make up a large class of genomic variants that contribute to a wide variety of disorders, mostly affecting the brain [1]. More importantly, there is increasing evidence that dynamic mutations serve important functions (namely regulation of gene expression) and may play a substantial role in modulating brain development and brain function [1-9]. For instance, SSRs are particularly abundant in genes involved in brain development and have been shown to contribute to normal variation in behavioral traits in animals and humans [3]. These SSRs therefore may have provided the variability needed for the changes of brain development and function in the primate lineage leading to human evolution [4].

In sexually reproducing populations, mating does not occur randomly, but in relationship to certain characteristics - either with a positive correlation (a male pairs with a similar female) or a negative correlation (a male pairs with a dissimilar female). This phenomenon, termed assortative mating, has been widely reported in humans with positive correlations for characteristics such as intelligence [10,11], body size [12-15] education [16], personality characteristics [17-19] and mental disorders [20,21]. A recent review outlined a number of human behaviors that are associated with SSRs including anxiety related traits, novelty seeking behavior, cognitive function, and altruism [3]. Therefore, behaviors modified by SSRs may undergo assortative mating, as shown for the allelic variation of the dopamine receptor D4 (DRD4) gene, an SSR associated with novelty seeking behavior [22].

A better understanding of the function of *HTT* below disease threshold may be important for understanding the disease process of HD. For instance, if there is phenotypic variation in brain structure or function associated with *HTT* CAG length below disease threshold, it may help to define a possible spectrum of phenotype ranging from normal to pathologic. This phenotypic spectrum extends the concept of disease pathology beyond the classic dichotomous categorization between normal and diseased brain. Furthermore, some literature supports the notion of important relationships between the normal HTT allele and the expanded allele, manifesting as differences in disease expression [23].

Although *HTT* has not directly been associated with variance in behavior, it is critical for brain development [24,25] and therefore may be associated with variance in brain structure and function. We sought to evaluate the possibility of assortative mating in a group of subjects recruited from the community and with no family history of Huntington's.

* Correspondence: peggy-nopoulos@uiowa.edu
[1]Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, Iowa, USA
Full list of author information is available at the end of the article

Methods

As a control group for a study on children at risk for Huntington's Disease, healthy children are recruited from the surrounding community of our hospital via advertising. Through a screening interview with parents, children are excluded if they have a history of significant medical neurologic, or psychiatric history. All participants signed informed consent prior to enrolling in the protocol, which was approved by the local Investigational Review Board (IRB). Participants ages 6-12 years signed both an assent form (language geared toward children) and the consent form.

Several children who participated were siblings and to avoid the confound of genetic relatedness, in the cases in which there were 1 or more siblings, the participants were randomly deleted and only 1 child from each family was included. The demographics of the group included 47 healthy children, including 31 girls and 16 boys, ranging from 6-18 years of age. A total of 40 of the 47 children were Caucasian (85% of the sample), 4 were African American (8.6%) and 3 were multiracial (6.4%).

Each child provided either blood or saliva for genetic analysis. All analyses were done through the University of Iowa Molecular Diagnostics Laboratory. Size of the CAG repeat region of *HTT* on chromosome 4p16.3 was determined with PCR analysis of genomic DNA. PCR primers that exclude the adjacent polymorphic CCG tract were used to amplify the CAG region. A second set of primers that includes the CCG polymorphism is routinely used to assist in differentiating two alleles with an identical CAG repeat number. The CAG repeat length for each subject is determined by comparing the PCR products to sizing standards. By convention, the longest allele is designated

as Allele1 and the shorter allele as Allele2. Parent DNA was not available to determine maternal or paternal transmission of each allele.

Statistical analysis

Normality of distribution of CAG lengths of both alleles was tested using the Shapiro-Wilk test. If either allele was found to be non-normally distributed, then non-parametric analysis was used (Spearman Correlation) to assess the association between length of Allele1 and length of Allele2.

Results

Distribution of CAG repeat lengths were not normally distributed. For Allele1, the range was from 15 to 30 with mean of 20.20, s.d. of 3.88. Shapiro-Wilk statistic was significant (0.873, $p < 0.0001$) indicating non-normality of distribution. For Allele2, the range was 4 to 29, mean of 17.29 and s.d. of 3.67. Shapiro-Wilk statistic was significant (0.848, $p < 0.0001$) again indicating a non-normal distribution.

The Spearman correlation between Allele1 and Allele2 was highly significant at $r = 0.511$, $p = 0.0002$. This observation confirms that longer Allele1 lengths are positively associated with longer Allele2 lengths. Figure 1 shows the scatter plot of the data with regression line displayed.

Discussion

This simple analysis of a unique data set shows evidence that there is assortative mating in regard to CAG length of *HTT*. That is, the length of CAG repeat in the maternal allele of *HTT* strongly correlates with the length of
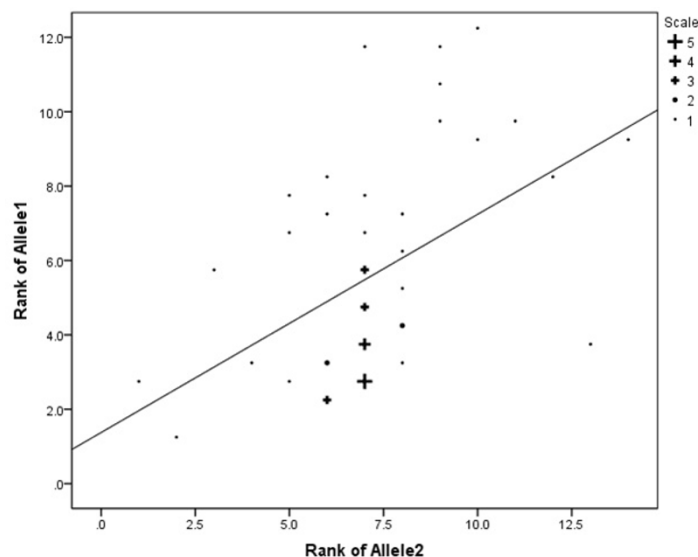

Figure 1 Relationship between CAG repeat length of Allele1 and CAG repeat length of Allele2.

the CAG repeat of the paternal allele of *HTT*, suggesting that the male and female pair mated based on the common feature of having a similar genotype (length of CAG repeat).

One interpretation of these findings would be that variability of CAG length is manifest by variation in phenotype of brain structure and function. This notion supports a report in which measures of mitochondrial energy metabolism (ATP/ADP) directly correlated to *HTT* CAG repeat lengths below disease threshold [26]. Although in that study the genetic variation was associated with metabolic phenotypic variation, brain structure/function phenotypic variation also may be associated with CAG repeat length in *HTT*. As is seen with other SSRs, variation in multiple types of behavior and cognitive functions have been shown to be associated with variations in repeat sizes of these genetic elements. With expansion of *HTT* CAG repeat length beyond 36 repeats, disease is manifested and the brain region most heavily affected is that of the basal ganglia. Basal ganglia circuits include widespread connections from and to multiple cortical regions, including frontal lobes. These frontal circuits influence numerous complex functions including cognitive and personality traits [27-32]. Potential influence of CAG repeat length on these basal ganglia frontal circuits could influence behaviors that provide the basis for assortative mating.

An alternative explanation for the current findings could be that a post-meiotic recombination between CAG repeat domains of the two *HTT* alleles 'equilibrate' the 2 allele sizes, thereby producing a correlation between the sizes of the *HTT* alleles [33]. To distinguish whether the mechanism responsible for the reported correlation is assortative mating or post-meiotic recombination, future studies will need to analyze parental DNA along with the proband's DNA.

Racial or ethnic relationships within the sample is an important consideration in the current study since there are differences in the distribution of normal *HTT* allele sizes in different ethnic groups [34-38]. Furthermore, different haplotypes associated with different distribution of the normal *HTT* allele sizes may influence the prevalence of HD in certain regions of the world [39,40]. Therefore, individuals choosing mates based on race or ethnicity may explain our current findings of relationships consistent with assortative mating. However, this explanation requires that a substantial number of subjects represent more than one racial or ethnic group. In the current sample, the vast majority (85%) are Caucasian with a small numbers of African Americans (n = 4) or multiracial subjects (n = 3). Moreover, if the correlation between the ranks of Allele1 and Allele2 are calculated within the 40 Caucasians, the relationship remains significant (Spearman's r = 0.408, p = 0.009). Thus, it seems unlikely that the current findings represent assortative mating based on ethnic group. Yet, recent reports of distribution of normal *HTT* allele sizes suggest that the prevalence of modifier genes may be different even *among sub-groups* of Caucasians [40]. Again, this explanation requires multiple subjects within several discreet sub-groups within this sample of 40 Caucasians which, although possible, seems less likely. Thus, assortative mating based not on ethnic group but on some other human feature remains a viable explanation for the findings reported here. Nevertheless, given the preliminary nature of the findings, follow-up in larger samples and further exploration of the functions of the variance of normal CAG length in *HTT* are warranted.

Author details
[1]Department of Psychiatry, University of Iowa Carver College of Medicine, Iowa City, Iowa, USA. [2]Department of Pediatrics, University of Iowa Carver College of Medicine, Iowa City, Iowa, USA. [3]Department of Neurology, University of Iowa Carver College of Medicine, Iowa City, Iowa, USA. [4]Department of Radiology, Washington University School of Medicine, St. Louis, MO, USA. [5]Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA. [6]Department of Anatomy & Neurobiology, Washington University School of Medicine, St. Louis, MO, USA. [7]Department of Neurology, Washington University School of Medicine, St. Louis, MO, USA. [8]Department of Occupational Therapy, Washington University School of Medicine, St. Louis, MO, USA.

Authors' contributions
PN secured funding, completed the analysis, and drafted the manuscript. EA provided feedback and revisions to manuscript. TW provided feedback and revisions to manuscript. BS provided feedback and revisions to manuscript. JP provided feedback and revisions to manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

References
1. Richards RI, Sutherland GR: **Dynamic mutations: a new class of mutations causing human disease.** *Cell* 1992, **70(5)**:709-12.
2. Fondon JW, Garner HR: **Molecular origins of rapid and continuous morphological evolution.** *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101(52)**:18058-63.
3. Fondon JW, *et al*: **Simple sequence repeats: genetic modulators of brain function and behavior.** *Trends in neurosciences* 2008, **31(7)**:328-34.
4. Hannan AJ: **Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for 'missing heritability'.** *Trends in genetics: TIG* 2010, **26(2)**:59-65.
5. Hannan AJ: **TRPing up the genome: Tandem repeat polymorphisms as dynamic sources of genetic variability in health and disease.** *Discovery medicine* 2010, **10(53)**:314-21.
6. Kashi Y, King DG: **Simple sequence repeats as advantageous mutators in evolution.** *Trends in genetics: TIG* 2006, **22(5)**:253-9.
7. Li YC, *et al*: **Microsatellites within genes: structure, function, and evolution.** *Molecular biology and evolution* 2004, **21(6)**:991-1007.
8. Nithianantharajah J, Hannan AJ: **Dynamic mutations as digital genetic modulators of brain development, function and dysfunction.** *BioEssays: news and reviews in molecular, cellular and developmental biology* 2007, **29(6)**:525-35.

9. Preuss TM, *et al*: **Human brain evolution: insights from microarrays.** *Nature reviews Genetics* 2004, **5**(11):850-60.
10. Epstein E, Guttman R: **Mate selection in man: evidence, theory, and outcome.** *Social biology* 1984, **31**(3-4):243-78.
11. Hur YM: **Assortative mating for personaltiy traits, educational level, religious affiliation, height, weight, adn body mass index in parents of Korean twin sample.** *Twin research: the official journal of the International Society for Twin Studies* 2003, **6**(6):467-70.
12. Garrison RJ, Anderson VE, Reed SC: **Assortative marriage.** *Eugenics quarterly* 1968, **15**(2):113-27.
13. Salces I, Rebato E, Susanne C: **Evidence of phenotypic and social assortative mating for anthropometric and physiological traits in couples from the Basque country (Spain).** *Journal of biosocial science* 2004, **36**(2):235-50.
14. Schwartz CR, Mare RD: **Trends in educational assortative marriage from 1940 to 2003.** *Demography* 2005, **42**(4):621-46.
15. Speakman JR, *et al*: **Assortative mating for obesity.** *The American journal of clinical nutrition* 2007, **86**(2):316-23.
16. Ho HZ: **Assortative mating in unwed birth parents, adoptive, and nonadoptive parents.** *Social biology* 1986, **33**(1-2):77-86.
17. Baron-Cohen S: **The hyper-systemizing, assortative mating theory of autism.** *Progress in neuro-psychopharmacology & biological psychiatry* 2006, **30**(5):865-72.
18. Galbaud du Fort G, *et al*: **Spouse similarity for antisocial behaviour in the general population.** *Psychological medicine* 2002, **32**(8):1407-16.
19. McCrae RR, *et al*: **Personality trait similarity between spouses in four cultures.** *Journal of personality* 2008, **76**(5):1137-64.
20. Low N, Cui L, Merikangas KR: **Spousal concordance for substance use and anxiety disorders.** *Journal of psychiatric research* 2007, **41**(11):942-51.
21. Mathews CA, Reus VI: **Assortative mating in the affective disorders: a systematic review and meta-analysis.** *Comprehensive psychiatry* 2001, **42**(4):257-62.
22. Eisenberg DT, *et al*: **Assortative human pair-bonding for partner ancestry and allelic variation of the dopamine receptor D4 (DRD4) gene.** *Social cognitive and affective neuroscience* 2010, **5**(2-3):194-202.
23. Aziz NA, *et al*: **Normal and mutant HTT interact to affect clinical severity and progression in Huntington disease.** *Neurology* 2009, **73**(16):1280-5.
24. Mehler MF, Gokhan S: **Mechanisms underlying neural cell death in neurodegenerative diseases: alterations of a developmentally-mediated cellular rheostat.** *Trends in neurosciences* 2000, **23**(12):599-605.
25. Molero AE, *et al*: **Impairment of developmental stem cell-mediated striatal neurogenesis and pluripotency genes in a knock-in model of Huntington's disease.** *Proceedings of the National Academy of Sciences of the United States of America* 2009, **106**(51):21900-5.
26. Seong IS, *et al*: **HD CAG repeat implicates a dominant property of huntingtin in mitochondrial energy metabolism.** *Hum Mol Genet* 2005, **14**(19):2871-80.
27. Bonelli RM, Cummings JL: **Frontal-subcortical dementias.** *The neurologist* 2008, **14**(2):100-7.
28. Frank MJ, *et al*: **Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation.** *Nature neuroscience* 2009, **12**(8):1062-8.
29. Hikosaka O, Nakamura K, Nakahara H: **Basal ganglia orient eyes to reward.** *Journal of neurophysiology* 2006, **95**(2):567-84.
30. Lebreton M, *et al*: **The brain structural disposition to social interaction.** *The European journal of neuroscience* 2009, **29**(11):2247-52.
31. Rosen HJ, Levenson RW: **The emotional brain: combining insights from patients and basic science.** *Neurocase* 2009, **15**(3):173-81.
32. Yu R, *et al*: **Insula and striatum mediate the default bias.** *The Journal of neuroscience: the official journal of the Society for Neuroscience* 2010, **30**(44):14702-7.
33. Aziz NA, *et al*: **Parent-of-origin differences of mutant HTT CAG repeat instability in Huntington's disease.** *Eur J Med Genet* 2011, **54**(4):e413-8.
34. Garcia-Planells J, *et al*: **Ancient origin of the CAG expansion causing Huntington disease in a Spanish population.** *Hum Mutat* 2005, **25**(5):453-9.
35. Kartsaki E, *et al*: **Late-onset and typical Huntington disease families from Crete have distinct genetic origins.** *Int J Mol Med* 2006, **17**(2):335-46.
36. Rubinsztein DC, *et al*: **Haplotype analysis of the delta 2642 and (CAG)n polymorphisms in the Huntington's disease (HD) gene provides an explanation for an apparent 'founder' HD haplotype.** *Hum Mol Genet* 1995, **4**(2):203-6.
37. Scholefield J, Greenberg J: **A common SNP haplotype provides molecular proof of a founder effect of Huntington disease linking two South African populations.** *Eur J Hum Genet* 2007, **15**(5):590-5.
38. Wang CK, *et al*: **DNA haplotype analysis of CAG repeat in Taiwanese Huntington's disease patients.** *Eur Neurol* 2004, **52**(2):96-100.
39. Warby SC, *et al*: **HTT haplotypes contribute to differences in Huntington disease prevalence between Europe and East Asia.** *Eur J Hum Genet* 2011, **19**(5):561-6.
40. Warby SC, *et al*: **CAG expansion in the Huntington disease gene is associated with a specific and targetable predisposing haplogroup.** *Am J Hum Genet* 2009, **84**(3):351-66.