

Distribution of the HTT gene A1 and A2 haplotypes worldwide: A systematic review

Thays Andrade Apolinário, PhD; Dionatan Costa Rodrigues, MSc; Mayra Braga Lemos, BSc;
Carmen Lúcia Antão Paiva, PhD; and Luciana Andrade Agostinho, PhD

Running title: Haplotypes A1-A2 HTT gene

Word count of the text, excluding abstract, tables, figures, legends and references: 3387

Word count of the Abstract, excluding headings: 280

Corresponding Author:

Carmen Lúcia Antão Paiva, PhD
Department of Genetics and Molecular Biology
Instituto Bimédico
UNIRIO, RJ, Brazil
Email: clapaiva1@gmail.com
doi:10.3121/cmr.2020.1523

Received: August 13, 2019
1st Revision: February 15, 2020
2nd Revision: May 31, 2020
Accepted: August 4, 2020

Abstract

Background: Huntington's disease (HD)(MIM:143100) is an severe autosomal dominant neurodegenerative disease caused by the dynamic expansion of CAG trinucleotides (> 35) in the *HTT* gene [Genomic Coordinates- (GRCh38):4:3,074,680-3,243,959]. **Objectives:** The aim of this systematic review was to investigate the reported associations between the frequencies of the A1 and A2 haplotypes in HD-affected and non-affected populations from different countries on different continents, in order to demonstrate the overall profile of these haplotypes worldwide, pointing towards the most frequent haplotypes that could be useful for *HTT* mutant-specific allele silencing in different populations. **Methods:** Publications in MEDLINE (PubMed) and Embase from the last 10 years (PROSPERO CRD42018115282) were assessed. **Results:** A total of 20 articles from 113 were selected for evaluation in their entirety, and eight were eligible for this study. **Conclusion:** Regardless of the size of the CAG tract, the articles included in this review demonstrate that populations with high HD prevalence present higher frequencies of the A1 or A2 haplotypes than populations exhibiting low HD prevalence, even when similar average CAG numbers are noted. Based on the presented articles, we suggest that the haplotypic profile is more closely related to the ancestral origin than to the size of the CAG tract. The identification of populations presenting a higher frequency of high-risk genotypes can contribute to more accurate genetic counseling, in addition to providing knowledge on HD epidemiology. According to the continued progress in the development of specific genetic silencing therapies by different research groups and pharmaceutical companies, such as haplotype targeting strategies for allele-specific *HTT* suppression, we conclude that the definition of haplotypes in phase with CAG expansions will contribute to the design of gene-silencing drugs specific for different populations worldwide.

Keywords: Huntington's Disease; *HTT* gene; Haplotypes

1. INTRODUCTION

Huntington's disease (HD) (MIM:143100) is a severe, progressive autosomal dominant neurodegenerative disease characterized by motor, cognitive and behavioral disorders. The cause of this disease is the dynamic expansion of CAG trinucleotides beyond their normal threshold (> 35 CAG repeats) within the *HTT* gene on chromosome 4 (4p16.3) [Genomic Coordinates- (GRCh38):4:3,074,680-3,243,959]. As a result of the expansion of CAG trinucleotide repeats, the huntingtin protein encoded by this gene is expressed with a longer-than-normal polyglutamine tail, which is associated with neuronal death^{1,2,3}, although other pathogenic mechanisms have also been suggested^{4,5,6,7,8}.

Normal alleles exhibit fewer than 27 CAG repeats⁹. Alleles with 27 to 35 replicates are classified as intermediate alleles (IAs), which usually lead to a normal phenotype but are genetically unstable during DNA replication and, therefore, may generate *de novo* HD mutations⁹. Alleles consisting of 36 to 39 CAG units exhibit reduced penetrance and may or may not cause HD symptoms⁹. Alleles with over 40 CAG repeats show complete penetrance and will cause the HD phenotype⁹.

HD symptoms begin between 40-50 years of age in most cases, but juvenile cases in which the symptoms of the disease appear before the age of 20 have been reported. Patients with HD survive for 15 to 20 years after the age of onset of the first symptoms¹⁰.

According to Harper¹¹, the founding effect of the HD mutation occurred in Europe and subsequently spread to other continents, and the prevalence of HD varies according to ancestral origin¹¹. In Eastern and Central Europe, prevalence was reported as 5.16 per 100,000 in 2010¹²; in the UK, 12.3 per 100,000¹³; and in Italy, 10.85 per 100,000¹⁴. In three studies conducted in Asia, the mean prevalence was estimated at only 0.40/100,000

individuals¹⁵, while in Africa HD prevalence ranges from 1-3.56/100,000, depending on ancestry¹².

The instability of CAG trinucleotides across generations can be influenced by the sex or age of the allele carrier, the size of the CAG alleles, environmental factors and *trans*- and *cis*-factors. Genetic variants such as single nucleotide polymorphisms (SNPs) in the *HTT* gene have been described as the main interfering elements^{16,17,18,19,20,21}, although, this interference has been contested by²².

A haplotype (haploid genotype) is a group of alleles in linkage disequilibrium inherited from a single parent; i.e., alleles that tend to always occur together and are therefore statistically associated. An *HTT* haplogroup is a cluster of similar haplotypes, such as haplogroup A (variants A1-A5), B (variants B1a, B1b and B2) and C (C1-C8). Specific sets of SNPs have been investigated in different HD populations for categorization into specific haplogroups.

Different categories of *HTT* gene haplotypes have been assessed by Warby et al.^{20,23}, Pulkes et al.²⁴, Baine et al.²⁵, Kay et al.²⁶ and Lee et al.²⁷. References^{20, 25, 26 and 27} provide figures showing sets of SNPs of different haplotypes that define *HTT* haplogroups.

The investigation of haplotypes in a population is paramount, since HD prevalence and incidence are related to haplotypes that are observed more frequently in the normal population^{20,23}. The identification of populations presenting a higher frequency of A1- A2 genotypes for HD can contribute to more accurate genetic counseling, in addition to providing knowledge about HD prevalence. Furthermore, it allows public health policies to be followed and may aid in the design of specific genetic silencing therapies, representing an allele-specific treatment strategy for these patients.

The goal of this systematic literature review is to investigate the association of the frequencies of the A1 and A2 haplotypes in HD-affected and non-affected populations from different countries on different continents to reveal the overall profile of these haplotypes worldwide. A review of genotypes (haplotypes and haplogroups in conjunction with *HTT* CAG expansions) will add important information to support future *HTT* mutant-specific allele silencing in different populations worldwide.

2. METHODOLOGY

2.1. Protocol registry

We registered the study protocol at the PROSPERO database (<http://www.prisma-statement.org/Protocols/Registration>), under registration number CRD42018115282. We also followed the PRISMA statement recommendations in writing this systematic review³⁰.

2.2. Search strategy and search sources

The research question was structured based on the PICO strategy:

- P (Population): Normal population (non-HD individuals) and affected HD population.
- I (Intervention): Presence of A1 and A2 haplotypes in HD affected and non-affected populations in different countries.
- C (Comparison): Not applicable.
- O (Outcomes): Describe the frequency of individuals with A1 and A2 haplotypes in HD affected and non-affected populations in different countries.

The eligibility criteria were as follows:

- Observational epidemiological studies investigating the frequencies of A1 and A2 haplotypes in normal individuals (individuals not suffering from HD) and in the affected population.
- Articles published in the last 10 years (2008-2018).
- Studies in humans.
- Studies published in English, Portuguese or Spanish.

The exclusion criteria were as follows:

- Studies that investigated individuals without categorizing the frequencies of affected and non-affected individuals.
- Studies conducted only in HD families.
- Studies that did not determine the A1 and A2 haplotypes.
- Narrative literature reviews, case reports or case series.

2.2.2. Search sources and the research strategy

MEDLINE (via PubMed) and Embase were used as the search databases. Additional records were searched in bibliographic references following the same selection and/or exclusion criteria.

- PubMed:

("Huntington Disease"[Mesh] OR "Huntington Chorea" OR "Chorea, Huntington" OR "Huntington's Disease" OR "Chronic Progressive Hereditary Chorea (Huntington)" OR "Huntington Chronic Progressive Hereditary Chorea" OR "Progressive Chorea, Chronic Hereditary (Huntington)" OR "Progressive Chorea, Hereditary, Chronic (Huntington)" OR "Huntington's Chorea" OR "Chorea, Huntington's" OR "Chorea, Chronic Progressive

Hereditary (Huntington) " OR "Huntington Disease, Late Onset" OR "Late-Onset Huntington Disease" OR "Huntington Disease, Late-Onset" OR "Late Onset Huntington Disease" OR "Juvenile Huntington Disease" OR "Juvenile-Onset Huntington Disease" OR "Juvenile Onset Huntington Disease" OR "Huntington Disease, Juvenile-Onset" OR "Huntington Disease, Juvenile Onset" OR "Huntington Disease, Juvenile" OR "Akinetic-Rigid Variant of Huntington Disease" OR "Akinetic Rigid Variant of Huntington Disease" OR "Huntington Disease, Akinetic-Rigid Variant" OR "Huntington Disease, Akinetic Rigid Variant")) AND ("Haplotypes"[Mesh] OR "Haplotype" OR "A1" OR "A2" OR "Haplogroup")

- Embase:

'huntington disease'/exp AND 'haplotype'/exp AND[embase]/lim NOT [medline]/lim AND [2008-2018]/py

2.3. Study selection and data collection process

First, a title analysis of the articles was carried out according to the objective of this review, and articles of interest were included in a second round of analysis through an abstract selection. Articles were then selected for the third stage based on the relevant abstracts, in which the full texts were analyzed for the final selection. Articles that appeared in more than one database were considered only once.

Article selection was performed by two independent reviewers who extracted data such as the authors, year of publication, number of chromosomes and individuals (affected and non-affected by HD), average numbers of CAG repeats and A1 and A2 haplotype frequencies (%) in affected and unaffected individuals.

A critical evaluation of the possible biases of the chosen articles was carried out by both independent reviewers based on the main types of bias that can be found in observational studies. These biases were classified into the following categories: selection, information and confounding bias. In the event of non-agreement, the final decision was made by a third reviewer who performed an independent selection.

3. RESULTS

A total of 116 titles were found, 78 in PubMed, 35 in Embase and three obtained through searching lists of bibliographical references. Among the 116 titles, three had been published in more than one database, and their duplicates were withdrawn, resulting in 113 articles. After applying the exclusion criteria, 20 articles were selected for evaluation in their entirety, eight of which were considered eligible and included in this systematic review. The flowchart of the study selection process is displayed in Figure 1 and data from the studies included in this review are displayed in Table 1.

Figure 1. Article selection algorithm

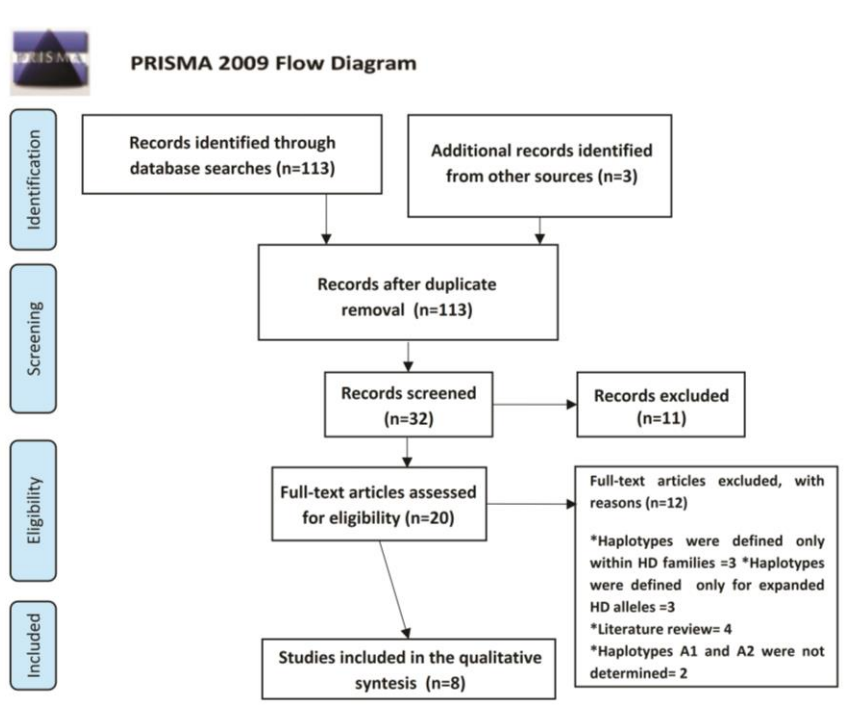


Table 1. Articles which met the inclusion criteria

Haplotypes A1-A2 HTT gene

Author/ Year	Country/ Continent	Markers for mounting the haplotypes/ Methods	Total number of alleles/chromosomes genotyped for A1 and A2 in individuals not affected by HD	Total number of expanded chromosomes genotyped for A1 and A2	Average CAG repeats in individuals not affected by HD*	Average number of CAG repeats in expanded chromosomes*	Haplotypes A1 and/or A2 in individuals not affected by HD	Haplotypes A1 and/or A2 in HD-affected individuals
Warby <i>et al.</i> , 2009 ²⁰	Canadian of European descent	22 TAG SNPs/ Sequencing	182 (n=116 normal alleles and 66 intermediate alleles)	65	Normal alleles: 19 Intermediate alleles: 30	44	Normal allele:53% A1 and/or A2 Intermediate alleles: 83% A1 and/or A2	95% A1 and/or A2
Warby <i>et al.</i> , 2011 ²³	China + Japan/ East Asia and Europe	21 TAG SNPs / Sequencing	501 (n = 73 China and Japan and 428 Europe)	230 (n=21 East Asia and 199 Europe)	East Asian: 16.9 European: 17.8	East Asia: 45.1 European: 45	East Asian: 0% European: 0.07% A1 and 0.13% A2	East Asian: 0% A1 and A2 European: 50% A1 and 29% A2.
Baine <i>et al.</i> , 2013 ²⁵	South Africa /Africa	21 SNPs/ Sequencing	311 (n=52 Caucasian, 65 Mixed and 194 Black)	72 (n=18 Caucasian, 19 Mixed and 35 Black)	Caucasian: 18.23 ± 3.17; Mixed:17.40 ± 3.19; Black: 16.91 ± 2.73	Caucasian: 44.77±4.09; Mixed: 46.19 ±4.75; Black: 44.51 ±4.02	Caucasian: 6% A1 and 9% A2. Mixed: 5% A1 and 1.5% A2 Black: 0% A1 and A2	Caucasian: 39% A1 and 50% A2; Mixed: 21% A1 and 53% A2; Black: 0% A1 and A2
Semaka <i>et al.</i> , 2013 ²⁸	Canada/North America	22 SNPs/ Sequencing	3188	135	Normal alleles: Most frequent CAG with 17 repeats intermediate alleles (mean): 28.9, repeats	NI	Normal alleles: 61.2% A1 or A2 intermediate alleles: 72.4% A1 or A2	71.1% A1 or A2
Pulkes <i>et al.</i> , 2014 ²⁴	Thailand/Asia	22 SNPs/ Sequencing	430	19	16.49	43.5	0.2% A1 and 1.4% A2	0% A1 and A2
Kay <i>et al.</i> , 2015 ²⁶	Canada/ North America Sweden, Finland, France, and Italy/Europe	63 SNPs / Sequencing + InDel for A1 Fragment Analysis	829	283	NI	NI	Canadian: 8% A1 and 16.4% A2 European: Swedish: 8.3% A1 and 12.8% A2 Finnish: 4.8% A1 and 14.3% A2 French:9.7% A1 and 15.3% de A2 Italian: 8.5% A1 and 24.4% A2	Canadian: 48.1% A1 and 32.2% A2 European: Swedish: 51% A1 and 17.6% A2 Finnish: 60% A1 and 40% A2 French:45.3% A1 and 22.6% de A2 Italian: 19.4% A1 and 58.2% de A2
Chao <i>et al.</i> , 2017 ³²	Europe	21 SNPs/ Sequencing + InDel for A1 Fragment Analysis	2184	4078	NI	NI	European: 3.08% A1a, 12.54% A2a and A2b*	European:42.22% A1a, 24.84% A2a and A2b.
Kay <i>et al.</i> , 2017 ³¹	Peru + other Latin American countries: Mexico, Colombia, Ecuador, El Salvador, Argentinaand Venezuela.	79 SNPs/ Sequencing	279 (n=84 Amerindians from Peru, 135 Mestizo Peruvian and 60 other Latin Americans with HD familial history)	80 (n=63 Mestizo Peruvian and 17 Latin American)	NI	NI	Amerindians from Peru: 6% A1 Amerindian and 0% A1 European Mestizo Peruvian: 1% A1 European, 16% A1 Amerindian and 4% A2. Latin American: 7% A1 Amerindian, 5% A1 European and 13% A2	Mestizo Peruvian: 70% A1 Amerindian, 3% A1 European and 11% A2 Latin American : 47% A1 Amerindian, 12% A1 European and 17% A2

A total of 7,904 normal chromosomes (<35 CAG repeats in *HTT*) and 5,018 expanded chromosomes (> 39 CAG) were analyzed among the eight studies included in this review.

The average size of the CAG tract in normal chromosomes for HD individuals ranged from 16.49²⁴ to 26²⁰ repeats, while the mean value of expanded CAG repeats varied from 43.5²⁴ to 46.19²⁵ in mutated *HTT*.

Among the studies included in this review, 37.5% (n = 3) involved European populations; 37.5% (n = 3) involved Canadian populations; 25% (n = 2) involved Asian populations; 12.5% (n=1) involved South African populations; and 12.5% (n = 1) involved Latin American populations. Two studies analyzed populations from different continents^{23,26}.

In two studies, the authors stratified allele types into normal, intermediate, reduced penetrance and complete penetrance alleles^{20,28}. The other studies categorized the alleles into normal (without full penetrance) and affected (with full penetrance).

One study involving populations from European countries demonstrated that recombination between haplotypes A1 and A2 rarely occurs within the transcribed *HTT* gene sequence. Only 9/283 (3.2%) expanded chromosomes, and 25/829 (3%) control chromosomes presented intragenic haplotype recombination²⁶.

A1 and A2 frequencies in individuals not affected by HD, who exhibited only normal alleles, were of 0% in the Asian (Thailand)²⁴, African (South African)²⁵ and Amerindian populations (Peru)³¹; on the other hand, the frequency of both A1 and A2 haplotypes was 61.2% in Canadians of European descent²⁸. In two studies involving normal Canadian individuals of European origin from the general population, A1 and A2 frequencies in the intermediate alleles were 72.4%²⁸ and 83%²⁰, respectively. In individuals affected by HD, however, this frequency ranged from 0% in Africans (South African)²⁵ and Asians (Thai)²⁴ to 95% in Canadians of European descent²⁰.

The highest A1 and A2 haplotype frequencies in individuals not affected by HD (both in normal and IAs) and HD carriers were observed in populations of European origin^{20,28}.

Two studies, one conducted in South Africa²⁵ and the other in Latin America (Peru)³¹, segmented participants by ancestral origin. In the South Africa study, the authors investigated Caucasians, mestizos and black people²⁵, while the Latin America (Peru) study assessed mestizo and Amerindian individuals³¹. In the Caucasian population of the South Africa study, variants A1 and A2 were associated with HD alleles (CAG > 35), and in the admixed

population, similarly to the Caucasian results, variants A1 and A2 were associated with HD alleles. Concerning the alleles of the two non-HD subpopulations (Caucasian and admixed), A1 haplogroups were associated with larger CAG repeats (26-30 in Caucasians and 21-25 in the mixed population)²⁵.

In the Latin America study (mestizo Peruvians), the expanded mutation probably originated in individuals with the native Amerindian A1 haplotype in most of the chromosomes (73%) of affected admixed individuals and those of Amerindian origin³¹. Thus, most of the HD-associated mutations in Latin America may occur on chromosomes with Amerindian descent haplotypes rather than haplotypes resulting from miscegenation with European individuals³¹.

It is worth mentioning that Chao et al.³², using the numerical *HTT* haplotyping system developed at Harvard, included a brief comparison of the two haplotype systems (e.g., letters or numbers) and demonstrated that the letter and number haplotyping systems do not map 1:1³². However, despite categorizing haplogroups with the Hap nomenclature²³, one study was included in this review because the authors compared and integrated the Hap haplotype system with the haplotypes A1 and A2³². Thus, the article did not meet any of the exclusion criteria of this review.

4. DISCUSSION

The HD A1-A2 haplotypes can be found in the general European population and are commonly associated with longer CAG expansions present in IAs and in normal alleles with 17-20 CAG repeats²⁰. Therefore, it has been suggested that *de novo* HD mutations originate in a step-by-step manner, giving rise to new HD cases, after which expansions arise via a step-by-step mechanism in which large normal alleles can generate IAs during

intergenerational transmission, which are then responsible for further generating *de novo* HD mutations²⁰.

Some authors have noted that the A1 and A2 haplotypes are associated not only with HD carriers but also with other normal formats with larger CAG expansions by chance^{22,27}. It has, therefore, been suggested that haplotypes do not influence predisposition to CAG trinucleotide expansion and that the major influencing factor is the size of the CAG tract itself, as observed in a sperm analysis performed by Semaka et al.²⁸, which indicated instability differences due to CAG tract length in IAs.

Targeting SNPs of both the A1 and A2 haplotypes is essential to achieve treatment of most HD patients in populations in which HD is most prevalent. The most common HD haplotypes (A1, A2, and A3a) define mutually exclusive sets of polymorphisms for allele-specific therapy in the greatest number of patients. Therefore, a maximum of 80% patients from Canada, Sweden, France, and Italy may be treatable by targeting these three haplotypes²⁶. The preclinical development of potent, selective antisense oligonucleotides (ASOs) targeting SNPs of the A2 HTT haplotype has been proposed as an allele-specific treatment strategy for the disease²⁶. In the past, it was believed that new HD mutations were very rare events and that the disease was restricted to families with a history of HD. However, it is now known that *de novo* expansions of CAG are more frequent than expected and that sporadic cases of HD correspond to 10% or more of total reported cases^{11, 33,34}.

Differences in HD prevalence rates within populations are closely linked to individual haplotypes, CAG trinucleotide sizes, IA frequency and the rate of new HD mutations³⁵. The prevalence of HD is usually lower in populations in which A1-A2 haplotypes are rare^{23,24}. This situation is noted in the results of the studies included in this review^{20,24,25,28,32}. In the

present study, the prevalence of A1-A2 haplotypes in Asian and African continent populations was found to be considerably lower than in European continent populations^{23,24,25}. In Asia (Thailand)²⁴, these haplotypes were shown to be absent in the affected population and present in only 0.2-1.4% of the general population²⁴. On the other hand, the highest frequency of A1 and A2 haplotypes in Canada, which is known to exhibit a high prevalence of HD, was found in the general population (up to 61.2% in normal alleles), and a frequency of up to 95% was found in expanded alleles^{23,28,30}. Rawlins et al. reported an increase in the prevalence of HD in Europe over the last 50 years in a systematic review, leading the authors to conclude that one of the causes of this increase was an increase in the mutation rate of the *HTT* gene, with a corresponding increase in HD incidence¹². Interestingly, HD incidence in the UK remained unchanged between 1990 and 2010³⁶.

CAG tract transmission instability was investigated in patients with HD and the normal Finnish population²⁹. The authors concluded that the low frequency of HD in Finland was partly explained by the relatively lower frequency of the haplogroup A in the general population. Furthermore, notable differences in the instability of intergenerational CAG repeats depending on parental haplotype and sex were noted. CAG repeats in haplogroup A increased by 3.18 CAG units in paternal transmissions but only by 0.11 units in maternal transmissions ($p = 0.008$), while haplogroup C repetition length decreased in both paternal and maternal transmissions²⁹. The frequency of A1-A2 haplotypes is null in HD-affected and unaffected black South African individuals²⁵. Baine et al.²⁵ reported that the most frequent haplotype in the African population is B2. This finding is correlated with the prevalence of HD in African ethnic groups, since previous reports indicate that the prevalence of HD is higher in Caucasians (22 per 1,000,000) than in blacks (0.1 per 1,000,000) on this continent³⁷.

Concerning HD prevalence in Latin America, a large cluster of HD is noted in Lake Maracaibo, Venezuela¹⁸. The overall prevalence of the disease is still unknown in Brazil, a country of continental dimensions. However, the minimum prevalence has been determined in some localities, such as Ervália-MG (7.2 per 10,000 inhabitants)³⁸, Feira Grande - AL (1 per 1000 inhabitants)³⁹, three regions in the state of Ceará (with prevalences of 12 per 370 inhabitants, 1 per 83,750 inhabitants and 1 per 1740 inhabitants)⁴⁰ and the state of Rio Grande do Sul (1.85 per 100,000 inhabitants)⁴¹. In this systematic review, the only included study investigating the Peruvian population (Latin America) indicates that, in addition to haplotypes of European origin, the population presents a subtype of the A1 haplotype, the Amerindian A1 haplotype³¹. Evidence suggests that HD originated from a mixture of founding populations from Europe and indigenous populations in Latin America³¹.

As stated previously, the A1 and A2 haplotypes are usually related to larger CAG expansions²⁰. Among normal chromosomes, 53% of normal alleles (mean of 19 CAGs) belong to these haplotypes in the Canadian population²⁰. Two studies evaluating populations of Asian origin^{23,24} and one assessing an African population²⁵ observed average numbers of CAG replicates that are considered low [mean 16.49 in Thailand and 16.91 in South Africa (blacks)]. In addition, the A1 and A2 haplotype frequencies were also low in the following populations: Thailand²⁴ 0.2% A1 and 1.4% A2 and South Africa²⁵ (black) 0% A1 and A2.

Among HD-affected individuals, the correlation between the mean CAG numbers and the A1 and A2 haplotypes differs according to the ethnic origin of the participants. It is noteworthy that in one study, individuals from certain European and Asian countries showed similar average CAG numbers (45 and 45.1, respectively)²³, whereas the A1 and A2 haplotypes were found in 50% and 29% of European chromosomes, respectively, and were absent in chromosomes from the Asian population²³. A similar result was observed by Baine et al.²⁵,

where South African Caucasian and black populations showed average numbers of CAG repeats of 44.77 and 44.51, respectively. Table 1 indicates that the A1 and A2 haplotypes were present at higher frequencies in populations of European origin^{20, 23, 25, 26, 28, 32}. On the other hand, no A1 or A2 haplotypes were found in the black population of East Asia²³ or the black population of South Africa²⁵.

The fact that a higher number of studies (37.5%) has been conducted in Europe may suggest a bias in indicating higher A1 and A2 frequencies on the European continent compared to other continents. Notably, very few studies involving African (n = 1), Asian (n = 2) and Latin American (n = 1) populations met the inclusion criteria of this review. It is important to mention that the low prevalence of HD in Asia, Africa and Latin America could also be influenced by underestimation. For example, mental health and mental health care have not become a high priority in most Asian countries⁴², despite the recent emphasis placed on mental health at the international level. However, Kay et al.³¹ demonstrated that the frequency of IAs in the healthy general population was the highest in Hispanic Americans and Northern Europeans and the lowest in black Africans and East Asians. The prevalence of HD correlates with the IA frequency by population and with the proportion of IAs found in the HD-associated A1 haplotype.

5. CONCLUSION

The frequencies of A1 and A2 haplotypes are higher in populations of European origin in both HD and non-HD groups than in South Africa and Asia, according to the articles included in this review. Haplotype investigation in a population is paramount, as HD prevalence and incidence are related to haplotypes observed more frequently in the population. The articles included in this review demonstrate that populations with a high HD prevalence exhibit

higher frequencies of the A1 or A2 haplotypes than populations displaying low HD prevalence, even when the populations exhibit similar average CAG numbers.

Based on the articles presented in Table 1, we suggest that the haplotypic profile is more closely related to ancestral origin than to the size of the CAG tract. Identifying populations presenting a higher frequency of high-risk genotypes for HD can contribute to more accurate genetic counseling, while also providing knowledge on HD epidemiology.

According to continued progress²⁰ in the development of specific genetic silencing therapies by different groups and pharmaceutical companies, such as haplotype-targeting strategies for allele-specific *HTT* suppression, we can conclude that the definition of the haplotypes in phase with *HTT* mutants will contribute to the design of gene-silencing drugs that are specific for different populations worldwide. It will be a great scientific achievement if a new specific drug (or drugs) that slows, halts, or even reverses HD symptoms reaches the market within 10 years. According to Kay et al.⁴³, HD is the most common genetic condition to be addressed with antisense oligonucleotides (ASOs). Targeting a transcript (with A1 or A2 haplotypes in phase with expanded *HTT*) with ASOs would selectively mark the expanded transcript for degradation.

REFERENCES

1. Gusella JF, Wexler NS, Conneally MP, et al. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Macmillan Journals*, 83. Retrieved from http://books.google.com/books?hl=en&lr=&id=A1hChCwPefsC&oi=fnd&pg=PA153&dq=A+polymorphic+DNA+marker+genetically+linked+to+Huntington's+disease&ots=_wu7mFhVsc&sig=qLwmfsaLLMnp1FEh1HbJPsQoBks%5Cnhttp://books.google.com/books?hl=en&lr=&id=A1hChCwPefsC&oi=fnd

2. Gusella JF, MacDonald ME. Trinucleotide instability: a repeating theme in human inherited disorders. *Annu Rev Med.* 1996;47:201-209.
[doi:10.1146/annurev.med.47.1.201](https://doi.org/10.1146/annurev.med.47.1.201). [Medline](#)
3. The 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(6):971-983. [doi:10.1016/0092-8674\(93\)90585-E](https://doi.org/10.1016/0092-8674(93)90585-E).
[Medline](#)
4. Nalavade R, Griesche N, Ryan DP, Hildebrand S, Krauß S. Mechanisms of RNA-induced toxicity in CAG repeat disorders. *Cell Death Dis.* 2013;4(8):e752-e11.
[doi:10.1038/cddis.2013.276](https://doi.org/10.1038/cddis.2013.276). [Medline](#)
5. Martí E. RNA toxicity induced by expanded CAG repeats in Huntington's disease. *Brain Pathol.* 2016;26(6):779-786. [doi:10.1111/bpa.12427](https://doi.org/10.1111/bpa.12427). [Medline](#)
6. Vuono R, Winder-Rhodes S, de Silva R, et al; REGISTRY Investigators of the European Huntington's Disease Network. The role of tau in the pathological process and clinical expression of Huntington's disease. *Brain.* 2015;138(7):1907-1918.
[doi:10.1093/brain/awv107](https://doi.org/10.1093/brain/awv107). [Medline](#)
- 7- Bassi, S., Tripathi, T., Monziani, A., Di Leva, F., & Biagioli, M. (2017). Epigenetics of Huntington's Disease. *Neuroepigenomics in Aging and Disease*, 277–299.
[doi:10.1007/978-3-319-53889-1_15](https://doi.org/10.1007/978-3-319-53889-1_15)
8. Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. Identification of Genetic Factors that Modify Clinical Onset of Huntington's Disease. *Cell.* 2015;162(3):516-526. [doi:10.1016/j.cell.2015.07.003](https://doi.org/10.1016/j.cell.2015.07.003). [Medline](#)
9. The American College of Medical Genetics/American Society of Human Genetics Huntington Disease Genetic Testing Working Group. ACMG/ASHG statement. Laboratory guidelines for Huntington disease genetic testing. *Am J Hum Genet.*

- 1998;62(5):1243-1247. [doi:10.1086/301846](https://doi.org/10.1086/301846). [Medline](#)
10. 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(4):398-403. [doi:10.1038/ng0893-398](https://doi.org/10.1038/ng0893-398). [Medline](#)
11. Harper P. The epidemiology of Huntington's disease. *Hum Genet.* 1992;89(4):365-376. [doi:10.1007/BF00194305](https://doi.org/10.1007/BF00194305). <http://www.ncbi.nlm.nih.gov/pubmed/1535611>. [Medline](#)
12. Rawlins MD, Wexler NS, Wexler AR, et al. The Prevalence of Huntington's Disease. *Neuroepidemiology.* 2016;46(2):144-153. [doi:10.1159/000443738](https://doi.org/10.1159/000443738). [Medline](#)
13. Evans SJW, Douglas I, Rawlins MD, Wexler NS, Tabrizi SJ, Smeeth L. Prevalence of adult Huntington's disease in the UK based on diagnoses recorded in general practice records. *J Neurol Neurosurg Psychiatry.* 2013;84(10):1156-1160. [doi:10.1136/jnnp-2012-304636](https://doi.org/10.1136/jnnp-2012-304636). [Medline](#)
14. Squitieri F, Griguoli A, Capelli G, Porcellini A, D'Alessio B. Epidemiology of Huntington disease: first post- *HTT* gene analysis of prevalence in Italy. *Clin Genet.* 2016;89(3):367-370. [doi:10.1111/cge.12574](https://doi.org/10.1111/cge.12574). [Medline](#)
15. Pringsheim T, Wiltshire K, Day L, Dykeman J, Steeves T, Jette N. The incidence and prevalence of Huntington's disease: A systematic review and meta-analysis. *Mov Disord.* 2012;27(9):1083-1091. [doi:10.1002/mds.25075](https://doi.org/10.1002/mds.25075). [Medline](#)
16. Richards RI. Dynamic mutations: a decade of unstable expanded repeats in human genetic disease. *Hum Mol Genet.* 2001;10(20):2187-2194. [doi:10.1093/hmg/10.20.2187](https://doi.org/10.1093/hmg/10.20.2187). [Medline](#)
17. Rolfsmeier ML, Dixon MJ, Pessoa-Brandão L, Pelletier R, Miret JJ, Lahue RS. Cis-elements governing trinucleotide repeat instability in *Saccharomyces cerevisiae*. *Genetics.* 2001;157(4):1569-1579. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1461582/pdf/11290713.pdf> [Medline](#).

18. Wexler, N. and, & The U.S.–Venezuela Collaborative Research Project. (2004). Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. *PNAS*, *101*(10). Retrieved from www.pnas.org/cgi/doi/10.1073/pnas.0308679101
19. Nguyen GH, Bouchard J, Boselli MG, et al. DNA stability and schizophrenia in twins. *Am J Med Genet*. 2003;120B(1):1-10. [doi:10.1002/ajmg.b.20010](https://doi.org/10.1002/ajmg.b.20010). [Medline](#)
20. Warby SC, Montpetit A, Hayden AR, 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):351-366. [doi:10.1016/j.ajhg.2009.02.003](https://doi.org/10.1016/j.ajhg.2009.02.003). [Medline](#)
21. Squitieri F, Andrew SE, Goldberg YP, et al. DNA haplotype analysis of Huntington disease reveals clues to the origins and mechanisms of CAG expansion and reasons for geographic variations of prevalence. *Hum Mol Genet*. 1994;3(12):2103-2114. [doi:10.1093/hmg/3.12.2103](https://doi.org/10.1093/hmg/3.12.2103). [Medline](#)
22. Falush D. Haplotype background, repeat length evolution, and Huntington's disease. *Am J Hum Genet*. 2009;85(6):939-942. [doi:10.1016/j.ajhg.2009.11.002](https://doi.org/10.1016/j.ajhg.2009.11.002). [Medline](#)
23. Warby SC, Visscher H, Collins JA, et al. HTT haplotypes contribute to differences in Huntington disease prevalence between Europe and East Asia. *Eur J Hum Genet*. 2011;19(5):561-566. [doi:10.1038/ejhg.2010.229](https://doi.org/10.1038/ejhg.2010.229). [Medline](#)
24. Pulkes T, Papsing C, Wattanapokayakit S, Mahasirimongkol S. Cag-expansion haplotype analysis in a population with a low prevalence of huntington's disease. *J Clin Neurol*. 2014;10(1):32-36. [doi:10.3988/jcn.2014.10.1.32](https://doi.org/10.3988/jcn.2014.10.1.32). [Medline](#)
25. Baine FK, Kay C, Ketelaar ME, et al. Huntington disease in the South African population occurs on diverse and ethnically distinct genetic haplotypes. *Eur J Hum Genet*. 2013;21(10):1120-1127. [doi:10.1038/ejhg.2013.2](https://doi.org/10.1038/ejhg.2013.2). [Medline](#)
26. Kay C, Collins JA, Skotte NH, et al. Huntingtin haplotypes provide prioritized target Haplotypes A1-A2 HTT gene

- panels for allele-specific silencing in huntington disease patients of european ancestry. *Mol Ther.* 2015;23(11):1759-1771. [doi:10.1038/mt.2015.128](https://doi.org/10.1038/mt.2015.128). [Medline](#)
27. Lee JM, Gillis T, Mysore JS, et al. Common SNP-based haplotype analysis of the 4p16.3 Huntington disease gene region. *Am J Hum Genet.* 2012;90(3):434-444. [doi:10.1016/j.ajhg.2012.01.005](https://doi.org/10.1016/j.ajhg.2012.01.005). [Medline](#)
28. Semaka A, Kay C, Doty CN, Collins JA, Tam N, Hayden MR. High frequency of intermediate alleles on huntington disease-associated haplotypes in British Columbia's general population. *Am J Med Genet B Neuropsychiatr Genet.* 2013;162(8):864-871. [doi:10.1002/ajmg.b.32193](https://doi.org/10.1002/ajmg.b.32193). [Medline](#)
29. Ylönen S, Sipilä JOT, Hietala M, Majamaa K. HTT haplogroups in Finnish patients with Huntington disease. *Neurol Genet.* 2019;5(3):e334. [doi:10.1212/NXG.0000000000000334](https://doi.org/10.1212/NXG.0000000000000334). [Medline](#)
30. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6(7):e1000097. [doi:10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097). [Medline](#)
31. Kay C, Tirado-Hurtado I, Cornejo-Olivas M, et al. The targetable A1 Huntington disease haplotype has distinct Amerindian and European origins in Latin America. *Eur J Hum Genet.* 2017;25(3):332-340. [doi:10.1038/ejhg.2016.169](https://doi.org/10.1038/ejhg.2016.169). [Medline](#)
32. Chao MJ, Gillis T, Atwal RS, et al. Haplotype-based stratification of Huntington's disease. *Eur J Hum Genet.* 2017;25(11):1202-1209. [doi:10.1038/ejhg.2017.125](https://doi.org/10.1038/ejhg.2017.125). [Medline](#)
33. Conneally PM. Huntington disease: genetics and epidemiology. *Am J Hum Genet.* 1984;36(3):506-526 [Medline](#).
34. Falush D, Almquist EW, Brinkmann RR, Iwasa Y, Hayden MR. Measurement of mutational flow implies both a high new-mutation rate for Huntington disease and

substantial underascertainment of late-onset cases. *Am J Hum Genet.* 2001;68(2):373-385. [doi:10.1086/318193](https://doi.org/10.1086/318193).

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1235271/pdf/AJHGv68p373.pdf>.

[Medline](#)

35. Kay C, Collins JA, Wright GEB, et al. The molecular epidemiology of Huntington disease is related to intermediate allele frequency and haplotype in the general population. *Am J Med Genet B Neuropsychiatr Genet.* 2018;177(3):346-357. [doi:10.1002/ajmg.b.32618](https://doi.org/10.1002/ajmg.b.32618). [Medline](#)
36. Wexler NS, Collett L, Wexler AR, et al. Incidence of adult Huntington's disease in the UK: a UK-based primary care study and a systematic review. *BMJ Open.* 2016;6(2):e009070. [doi:10.1136/bmjopen-2015-009070](https://doi.org/10.1136/bmjopen-2015-009070). [Medline](#)
37. Hayden MR, MacGregor JM, Beighton PH. (1980). The prevalence of Huntington's chorea in South Africa. *South African Medical Journal = Suid-Afrikaanse Tydskrif Vir Geneeskunde*, 58(5), 193–196. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6447364>.
38. Agostinho LA, da Silva IS, Maia LA, et al. A Study of a Geographical Cluster of Huntington's Disease in a Brazilian Town of Zona da Mata, Minas Gerais State. *Eur Neurol.* 2015;74(1-2):62-68. [doi:10.1159/000434630](https://doi.org/10.1159/000434630). [Medline](#)
39. Alencar MA, Lopez AM, Figueiredo E, Porciúncula CG, Monlleó I. E05 Prevalence of Huntington's disease in Feira Grande, a small city in Northeastern Brazil. *Journal of Neurology, Neurosurgery & Psychiatry* 2010;81:A22. [doi:10.1136/jnnp.2010.22638.5](https://doi.org/10.1136/jnnp.2010.22638.5).
40. Cavalcante Filho JRM, Furtado LETA. (2018). Epidemiologia da Doença de Huntington no Ceará. *Primeiro Encontro Latino-Americano Da Doença de Huntington*. Barranquilla-Colômbia.
41. Castilhos RM, Santos JA, Augustin MC, et al; RedeNeurogenética. Minimal prevalence Haplotypes A1-A2 HTT gene

of Huntington's disease in the South of Brazil and instability of the expanded CAG tract during intergenerational transmissions. *Genet Mol Biol.* 2019;42(2):329-336.

[doi:10.1590/1678-4685-gmb-2018-0032](https://doi.org/10.1590/1678-4685-gmb-2018-0032). [Medline](#)

42. Meshvara D. Mental health and mental health care in Asia. *World Psychiatry.*

2002;1(2):118-120 [Medline](#).

43. Kay C, Collins JA, Caron NS, et al. A Comprehensive Haplotype-Targeting Strategy for

Allele-Specific HTT Suppression in Huntington Disease. *Am J Hum Genet.*

2019;105(6):1112-1125. [doi:10.1016/j.ajhg.2019.10.011](https://doi.org/10.1016/j.ajhg.2019.10.011). [Medline](#)

AUTHOR AFFILIATIONS

Thays Andrade Apolinário, PhD*; Dionatan Costa Rodrigues, MSc*; Mayra Braga Lemos, BSc‡; Carmen Lúcia Antão Paiva, PhD*,†,‡,§; *Luciana Andrade Agostinho, PhD*,§,¶

*Graduate Program in Neurology, Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Rio de Janeiro, RJ, Brazil

†Graduate Program in Molecular and Cell Biology, UNIRIO, RJ, Brazil

‡Department of Genetics and Molecular Biology, Instituto Bimédico, UNIRIO, RJ, Brazil

§University Center UNIFAMINAS - UNIFAMINAS, Muriaé, MG, Brazil

¶Authors who contributed equally to this work