Pathogenic mutations and non-pathogenic DNA polymorphisms in the most common neurodegenerative disorders

Andrzej Kochański
Neuromuscular Unit, Mossakowski Medical Research Centre, Polish Academy of Sciences, Warsaw, Poland

Folia Neuropathol 2007; 45 (4): 164-169

Abstract
In the past, each nucleotide change causing amino acid substitution in a gene in which other mutations responsible for a neurodegenerative disorder had been found was considered as a causative mutation. However, in recent years, mainly due to the progress of the Human Genome Project (HGP), numerous DNA variants have been identified in many neurodegenerative disorders. Some of them likely belong to the class of pathogenic (causative) mutations, whereas others, which may occasionally coexist with the disease phenotype, should be classified as non-pathogenic DNA polymorphisms. How to differentiate between a pathogenic mutation and a harmless DNA polymorphism nowadays, i.e. in the post-genomic era? The question still remains open. Erroneously assumed pathogenicity of a mutation may result in a misdiagnosis of the disease and in consequence lead to inappropriate genetic counselling. The aim of this short review is to present a set of mutations with no clear pathogenic effect that have been identified in some neurodegenerative disorders.

Key words: neurodegenerative disorders, pathogenic mutation, harmless polymorphism, genetic counselling

Introduction
It is worth noting that genetic polymorphisms were defined by evolutionists in the pre-genomic era, in the early sixties. Given that advantageous or neutral variability of the organisms should be preferred by the basic process of natural evolution, i.e. natural selection, their frequency within a given population is expected to be high. In contrast, negative features should be gradually eliminated throughout the generations, so their frequency should be low. Although quite arbitrarily, variability occurring in the population with a minimal frequency of 1% was classified as DNA polymorphism. Similarly, rare variants with a frequency of <1% should be considered pathogenic due to their low frequency [13].

In fact, 54% of DNA variants dispersed in the human genome occur with a frequency below 1%, but they are not deleterious mutations. Only 23% of neutral DNA variants occur with a frequency higher than 10% [29]. The probability of finding a neutral variant in a patient which does not occur in 50 healthy controls (100 chromosomes) is about 15%. Thus, the probability of a false assumption that a harmless DNA variant is a pathogenic mutation reaches 15% [29].

In the pre-genomic era, any DNA variant not resulting in an amino acid change was considered...
a harmless polymorphism. Nowadays, however, numerous DNA variants that are not associated with an amino acid change have been shown to have deleterious effects [5].

A DNA variant localized in the conserved protein domain with a significant biochemical function (regulatory domains, ATP-binding domains) is thought to have a deleterious effect.

A deletion of a non-conserved hydrophilic loop domain VI (HLVI) of presenilin (Cys263-Leu383) does not alter Aβ42 production; thus mutations in the presenilin 1 and 2 genes located within the HL domain may be expected to have no deleterious effect. Indeed, the Glu318Gly and Thr354Ile substitutions have been shown to be non-pathogenic mutations [20, 24, 42].

Finally, a question whether a certain mutation segregates with a phenotype is important in the assessment of its pathogenic effect. In fact, over 97% of cases observed by clinical geneticists are sporadic ones, and thus there is no possibility to determine segregation with the phenotype [27]. Even in large pedigrees the lack of segregation of the mutation with the phenotype may be caused by a low penetrance.

Pathogenic mutations in Parkinson’s disease

Parkinson’s disease (PD) is the second most frequent neurodegenerative disorder, diagnosed in 4% of the population. Although the vast majority of patients affected with PD represent sporadic cases, multigenational families enabled 6 genes located within the HL domain may be expected to have no deleterious effect. Indeed, the Glu318Gly and Thr354Ile substitutions have been shown to be non-pathogenic mutations [20, 24, 42].

At least 3% of patients with PD harbour mutations in the LRRK2 and PINK1 genes. Of 50 known variants of the LRRK2 gene, at least 16 sequence changes seem to be pathogenic. The most common and best known LRRK2 gene mutation is Gly2019Ser, so far detected in hundreds of PD patients originating from different ethnic groups [25]. It is of interest whether the G2019S mutation may be associated with a peculiar type of pathology. In a study encompassing 1179 LB-negative brains, the G2019S variant was also identified in a healthy control and in the brain of a patient with Alzheimer’s disease (AD). In LB-positive brains (n=405), the G2019S variant was identified in only 8 cases [3].

The penetrance of G2019S mutation is age-dependent, but in fact this mutation was even identified in two healthy octogenarians [4,18]. Due to an extremely various age of onset and low penetration (30%), the segregation of the G2019S mutation with PD may be hard to prove.

In the vast majority of cases, PD with a recessive mode of inheritance is associated with mutations in the Parkin and PINK1 genes. Similarly to LRRK2 mutations, some PD patients who are compound heterozygotes for Parkin gene mutations remain healthy until the 6th decade of life [7].

On the other hand, heterozygous PINK1 and Parkin mutations have been identified in PD patients, suggesting the possibility of an autosomal dominant trait of inheritance. To estimate the contribution of the heterozygous PINK1 and Parkin gene mutations to the PD pathogenesis, their presence in the control groups must be excluded.

It is important to note that healthy individuals do not undergo as rigorous neurological examination as patients affected with PD. Moreover, there are no follow-up studies of the subject included in the control groups. Thus, some of the “healthy controls” may belong to the group of carriers of PINK1 and Parkin pathogenic mutations with low penetrance.

Neurofilament pathology is found in PD, AD, Down’s syndrome, infantile neuroaxonal dystrophy, Halleworden-Spatz syndrome, and subacute panencephalitis [9].

In 2002, a Gly336Ser mutation in the NEFM gene coding for medium neurofilament subunits was identified for a first time in a French-Canadian patient affected by early onset PD (at 16 years). Surprisingly, three siblings of the patient in their early 30s and 40s were healthy in spite of the presence of the Gly336Ser mutation. The Gly336Ser variant was not identified in a large control group consisting of 648 chromosomes and was located in the coil 2B domain of the NEFM protein essential for the assembly of neurofilaments. Additionally, the Gly336Ser substitution has been shown to be conserved within 8 species [22].

To address the question of a causative role of the G336S variant in PD pathogenesis, Perez-Olle and colleagues performed a functional analysis of the
neurofilament assembly in cell cultures transfected with the G336S variant. Since no obvious deleterious effect of the G336S mutation has been observed on the neurofilament assembly, this variant was classified as a harmless polymorphism [33]. However, the possibility of some other, unknown pathogenic effect of the G336S variant that is not associated with neurofilament assembly could not be excluded.

Mutations in the presenilin I (PSEN1) gene and early onset Alzheimer’s disease (EOAD)

Familial aggregation of AD had been noted for a long time in the “pre-genomic era”. A first AD gene coding for amyloid precursor protein (APP) was mapped to chromosome 21 [14]. To date, mutations in the APP gene have been identified in less than 1% of familial dominant EOAD cases. A second gene (mostly mutated in the EOAD patients) coding for presenilin 1 (PSEN1) was mapped to chromosome 14 q24.3 in 1992 [38,39]. So far over 150 mutations in the PSEN1 gene have been identified [10].

A third EOAD locus for the presenilin 2 (PSEN 2) gene was mapped to chromosome 1 [26,36]. Some mutations in the PSEN1 gene are thought to be associated with very early onset of cognitive decline (L85P, P117L, S169L, L424R), others were postulated to segregate with EOAD with aphasia (E120D, H163R, P264L), and spastic paraparesis was observed in EOAD patients harbouring L166P, F237I, V261F and other mutations [21].

Over 10% of PSEN1 mutations have been found in sporadic cases of EOAD [21].

The E318G variant in the PSEN1 gene was reported in 1998 in a small, two generational family. Although both the mother and her son were affected with EOAD, an analysis of segregation of this mutation with the phenotype was not possible, since DNA from the mother was not available [6].

In the Polish population, the E318G mutation was identified in two unrelated patients with the sporadic form of EOAD and was suggested to be a possible genetic risk factor contributing to the pathogenesis of familial AD [20].

In a study encompassing 256 AD patients, 210 healthy age-matched control subjects and 100 PD patients and centenarians, there was no statistically significant difference in the frequency of the E318G substitution between AD affected patients and healthy controls [42].

In other studies, the E318G mutation was identified both in EOAD affected patients and in the control group consisting of healthy subjects, with frequencies of 4.5% and 6.8%, respectively [1].

Mutations in genes coding for neurofilaments in amyotrophic lateral sclerosis and Charcot-Marie-Tooth disease

A role of neurofilaments in the pathogenesis of amyotrophic lateral sclerosis (ALS) was postulated because neurofilaments have been shown to accumulate in perikarya and proximal axons in patients with sporadic and familial ALS [15]. Additionally, transgenic mice with a mutant gene coding for neurofilaments were shown to develop motor neuron disease [23].

Accumulation of the neurofilaments observed in the ALS-affected patients may result from structural defects caused by mutations in three genes coding for neurofilaments or may be an additional effect of accumulation of the other “primarily” defected proteins. Given that mutations in the genes coding for neurofilaments may be causative for some forms of ALS, Figlewicz and colleagues screened a large group of ALS patients consisting of 306 individuals for mutations in the gene NEFH coding for the human neurofilament heavy subunit. In four unrelated ALS patients, a 3 bp deletion in the NEFH gene was identified, and a 102 bp long deletion was found in the fifth patient. Since these mutations were not found in healthy individuals, and additionally they have been located in the multiple repeat motifs important in phosphorylation of neurofilaments, their status has been considered pathogenic [11]. In the further studies of the NEFH gene in ALS affected patients, neither 3bp nor 102 bp deletions were identified. Although molecular genetic analysis of the gene coding for the human neurofilament light subunit revealed D469N change, its status was reported as polymorphism due to the lack of segregation with the ALS phenotype [40]. To conclude, to date no deleterious mutations have been identified in the NEFL, NEFM and NEFH genes in patients affected with ALS. Thus, it seems probable that accumulation of neurofilaments observed in ALS-affected patients may reflect the effects of some other, as yet unknown mecha-
nisms in which neurofilaments are involved. In contrast to ALS, a primary causative role of mutations in the NEFL gene has been shown in Charcot-Marie-Tooth affected patients. Similarly to ALS, accumulation of neurofilaments was observed in giant axonal neuropathy (GAN) in the pre-genomic era. The first mutation in the NEFL gene in CMT disease, E333K, was reported in 2005. To date, over 20 mutations in the NEFL gene have been identified [16]. A pathogenic effect was proved for the majority of them. The NEFL gene mutations affect mitochondrial distribution and axonal transport of NEFL protein and result in formation of NEFL protein aggregates in the cells [32]. Some of the NEFL mutations initially reported as pathogenic mutations, e.g. del528Glu, were identified in healthy individuals [41]. One of the mutations, I214M, which was identified in two unrelated patients, was shown not to segregate with the CMT phenotype. Functional studies of the I214M mutation revealed no obvious pathogenic effect [17].

Due to a high ratio of harmless polymorphisms in the NEFL gene, functional studies are necessary to distinguish between polymorphic variants and harmless polymorphisms in CMT disease.

**Mutations in the mitochondrial DNA**

Evaluation of pathogenic effects of variants occurring in the mitochondrial genome is a challenge for investigators for a number of reasons. First, in contrast to the nuclear genome, mitochondrial DNA mutates at a high rate. In addition, mitochondrial DNA errors are repaired about tenfold less efficiently than those of nuclear DNA. The vast majority of mitochondrial mutations are naturally mosaic due to the phenomenon of heteroplasmy.

Clinical variability of mitochondrial diseases is a challenge to draw reasonable phenotype-genotype correlations. In recent years, the number of mitochondrial DNA variants has increased exponentially. Thus, the issue of pathogenicity of mitochondrial mutations is nowadays widely and extensively discussed.

One of the first ribosomal RNA mutations (1555A to G) was reported in three families with antibiotic-induced and non-syndromic deafness. Among 22 variants of mitochondrial DNA detected in the deaf patients, the 1555A to G substitution is the most likely one to be pathogenic. In contrast to other DNA variants, the 1555A to G substitution has been shown to segregate with deafness in three unrelated families. This mutation was absent in a large control group, whereas other DNA variants occurred also in the healthy controls. Additionally, the 1555A to G substitution changes a nucleotide in a highly conserved region of the 12S rRNA which is known to bind aminoglycosides. Finally, this nucleotide was mutated in the aminoglycoside-resistant forms of different species of bacteria [35].

In contrast to the 1555A to G mutation located in the 12S rRNA gene, pathogenic status of the variants coding for mitochondrial tRNAs is much more difficult to assess. In a study encompassing 68 known pathogenic mutations and 64 harmless variants in mitochondrial tRNA genes, random distribution of variants within tRNA has been shown. Surprisingly, both polymorphisms and mutations were distributed with a similar frequency in anticodon, stem domains and tRNA loops.

In fact, the reaction between a certain aminacyl-tRNA synthetase and tRNA depends not solely on particular loops or stem domains but has a multifactorial mechanism [31].

In conserved tRNA residues, mutations occurred more frequently than polymorphisms. The majority of pathogenic mutations were transversions and deletions [12].

There is a possibility that mutated tRNA gene will be used as natural nonsense or missense suppressor tRNA in the process of protein translation [2].

A homoplasmic A5814G mutation in the tRNA\(^{\text{Cys}}\) gene has been shown to segregate with encephalomyopathy in mitochondrial mode of inheritance in three generations. In addition, this mutation is located in the D-stem of tRNA\(^{\text{Cys}}\), highly conserved within numerous species. Moreover, this mutation was previously identified in three other unrelated patients presenting with different phenotypes [37].

In contrast to the A5814G mutation, the G3283A transition in the tRNA\(^{\text{Aau}}\) gene was identified only in one 80-year-old woman with late-onset ocular myopathy [28]. It is unclear whether the G3283A mutation was a causative or harmless tRNA\(^{\text{Aau}}\) variant associated with acquired age-dependent features in that 80-year-old patient.

Some authors have recently proposed a scoring system for pathogenic effects of mitochondrial mutations. To estimate the likelihood that a given sequence variant identified in one of the 7 genes coding complex I (MTND) is pathogenic, 50 mutations in the
MTND genes were compared for their potential pathogenicity. The scoring system included biochemical defect, functional studies, heteroplasmia of mutation, number of reported families, segregation and conservation between species. Using this scoring system, only 32% of the DNA variants identified in the MTND genes were described as pathogenic [30].

Although the scoring system of pathogenicity of MTND mutations represents one of the first attempts to classify DNA variants as pathogenic mutations or harmless polymorphisms, its application seems to be limited. First, this scoring system may be addressed to MTND mutations only. Secondly, the scoring criteria were established arbitrarily. Finally, the boundaries between unclear, polymorphic, provisional and confirmed status of mutations seem to be artificial.

Conclusions and outlook

Due to the complexity of pathogenesis of neurodegenerative disorders and clinical variability of their phenotypes, genetic counselling for patients and their family members who are at risk of these disorders is challenging.

Medical management and genetic counselling in neurodegenerative disorders require a rigorous estimation of the pathogenic effect of the particular DNA variant identified in the patient.

An explosion in research into the genetic background of neurodegenerative disorders has resulted in identification of new genes and hundreds of DNA variants. Only recurrent mutations occurring in different populations and segregating with the disease phenotype in multigenerational pedigrees may be undoubtably classified as pathogenic mutations. In contrast to the previous genetic studies, defining the DNA variants as pathogenic mutations or harmless polymorphisms is an extremely difficult issue in the sporadic cases associated with novel DNA variants. Two examples, i.e. the E318G mutation in the PSEN1 gene in Alzheimer’s disease and the G336S variant in the NEFM gene in Parkinson’s disease, seem to be illustrative for the problem of a harmless polymorphism erroneously classified as a pathogenic mutation.

The most common recurrent mutation associated with Parkinson’s disease, G2019S in the LRRK2 gene, was identified in healthy octogenarians, which provokes the question of the frequency of other pathogenic mutations with low penetrance in the control groups [4,18].

Classical criteria may be still used for recurrent common mutations.

The scoring system proposed to estimate the pathogenic status of the mutations in the MTND genes may serve as a “prototype” approach in other neurodegenerative disorders [30]. A separate scoring systems for PSEN1 or LRRK2 gene mutations elaborated by international multidisciplinary neurogenetic teams (including basic scientists, clinical geneticists and neurologists) may be helpful to categorize new DNA variants.

There is no doubt that the number of DNA variants identified in neurodegenerative disorders will rise in the next few years. Thus, the introduction and subsequent improvement of the scoring systems is a sine qua non condition in optimal patient care.

References
Pathogenic mutations and non-pathogenic DNA polymorphisms in the most common neurodegenerative disorders


