



# Estimating the prevalence of allelic variants in the transthyretin gene by analysing large-scale sequencing data

Carmen Lahuerta Pueyo<sup>1</sup> · Miguel Ángel Aibar Arregui<sup>2</sup> · Anyuli Gracia Gutierrez<sup>2</sup> · Esperanza Bueno Juana<sup>2</sup> · Sebastián Menao Guillén<sup>1</sup>

Received: 10 May 2018 / Revised: 12 December 2018 / Accepted: 18 December 2018 / Published online: 25 January 2019  
© European Society of Human Genetics 2019

## Abstract

Transthyretin amyloidosis (ATTR amyloidosis) is a rare disease characterised by extracellular deposition of amyloid fibrils composed by transthyretin. ATTR amyloidosis can be sub-classified as wild-type ATTR (ATTR-wt) or as hereditary amyloidosis (ATTR-m); the prevalence of both types are likely underestimated. There are tools that can help us to study ATTR-m, as gnomAD database. Our primary aim was to estimate prevalence of variants, especially amyloidogenic variants, in the TTR gene using gnomAD database. We analysed TTR missense variants found in gnomAD. The variables studied were classified according to their clinical significance and according to the different populations. We found 71 missense variants in the TTR gene. Eleven variants were described as affects function variants (prevalence 1:230). The most frequently detected variant were c.424G>A (p.(Val142Ile)) (prevalence 1:332, MAF 0.00151) and c.148G>A (p.(Val50Met)) (prevalence 1:4924, MAF 0.000102), which represented 88% and 5%, respectively, of all affects function variants detected. Seventeen variants were classified as probably affects function, 29 as unknown variants, 4 as probably does not affect function and 10 as does not affect function variants. In terms of different populations, c.424G>A (p.(Val142Ile)) was especially prevalent in African population (MAF 0.01602; prevalence of 1:31) and c.148G>A (p.(Val50Met)) in European population (MAF 0.000179; prevalence of 1:2792). Prevalence of amyloidogenic variants in the general population was higher than prevalence heretofore described. This difference could be explained by incomplete penetrance of the disease, but other factors contributing to this fact, fundamentally the underdiagnosis of the disease.

## Introduction

Systemic amyloidoses are a group of rare diseases characterised by extracellular deposition of amyloid fibrils. The most common subset of amyloidoses includes primary systemic amyloidosis (AL), reactive systemic amyloidosis (AA) and transthyretin (TTR) amyloidosis (ATTR amyloidosis). However, more than 30 different proteins have been described as causing these disease [1, 2].

Focusing on ATTR amyloidosis, amyloid deposition material is produced by the destabilisation of the

transthyretin protein (TTR). This protein is predominantly produced in the liver (>95%), the choroid plexus and the retinal pigment epithelia, and it is involved in the transport of thyroxine and retinol [3].

ATTR amyloidosis can be sub-classified as one of two types. One type is associated with aging, and it is known as senile systemic amyloidosis or “wild-type ATTR” (ATTR-wt). The other type is known as hereditary amyloidosis (ATTR-m) due to point changes in the TTR allele that cause a deposition of mutant TTR amyloid [4].

ATTR-m has an autosomal dominant inheritance. It is caused by single amino-acid changes in the transthyretin gene, which is located on chromosome 18 (18q12.1). This gene is divided into four exons and, despite being small, over 130 allelic variants have been described, most of which are amyloidogenic [5, 6].

ATTR amyloidosis was first described in 1952 by Andrade in Portugal. In the following decades, other endemic areas in Japan and Sweden were also described [7]. Although the

✉ Carmen Lahuerta Pueyo  
carmenpueyo@gmail.com

<sup>1</sup> Service of Clinical Biochemistry, Hospital Clínico Universitario de Zaragoza (HCUZ)-IIS Aragón, Zaragoza, Spain

<sup>2</sup> Service of Internal Medicine, HCUZ-IIS Aragón, Zaragoza, Spain

disease first appeared to be endemic, advances in diagnostic techniques have led to diagnoses worldwide [5].

Clinical manifestations depend mainly on the existing variant. In the c.148G>A (p.(Val50Met)) located in exon 2 (most common in Europe), neurological manifestations are predominant (familial amyloid polyneuropathy [FAP]); in other variants, such as c.424G>A (p.(Val142Ile)) located in exon 4 (most common in the United States), a cardiac phenotype is predominant. In any case, there are described mixed phenotypes and other organ involvement [4, 8].

FAP has an estimated prevalence of 1/100,000 in Europe and the United States. In endemic regions of Portugal, Japan, Brazil or Sweden, the prevalence ranges from 1/1000 to 1/10,000 people. In Majorca, 5/100,000 people have been observed with FAP, and in Cyprus, 3.72/100,000 people [9–11]. ATTR-m with a cardiac phenotype has a higher frequency than FAP; the c.424G>A (p.(Val142Ile)) in the African population is 3–4%, while in Caucasian and Latin populations are <0.5% [5, 12].

The age of onset ranges from the second to ninth decade of life [13]. There is great heterogeneity in penetrance data depending on the phenotype, genotype and environmental factors. There are few references in the literature regarding the penetrance of the disease, the available data focus on the variant c.148G>A (p.(Val50Met)), the penetrance of the disease in carriers of these variant is different in different areas, increasing the risk of suffering the disease with age. In Portugal, the penetrance at 50 years is 80% compared with an 18% in French population. These percentages increase considerably with age, so at 70 years penetrance in Portugal increases to 91 and 50% in France. Swedish carriers have an even lower risk of 11% by age 50 and 36% by age 70 years [10, 14]. Anticipation phenomenon has been observed and furthermore, there is a risk of earlier onset when the mutation is inherited from the mother [12, 15].

In addition to the difficulties of knowing the true prevalence and penetrance of the disease, there are many references that suggest that the disease is underdiagnosed due to the clinical variability and the lack of specific symptoms or biomarkers [16–18]. For those reasons, we find the study of population data to be relevant to the analysis of both the TTR gene variants and the frequency of each of those variants.

There are tools that can help us to study rare diseases, such as ATTR amyloidosis. Also, there are large genomic databases—for example, the gnomAD database—that can support the study of genetic variation frequencies [19]. Our primary aim was to estimate the allele frequencies of variants in the TTR gene using the gnomAD database, especially those variants that are known to play a role in pathological processes.

## Materials and methods

GnomAD is an update of the ExAC project. The gnomAD database currently contains 123,136 exomes and 15,496 genomes from unrelated individuals and from different populations, sequenced as part of various disease-specific and population genetic studies. Individuals known to be affected by severe paediatric disease, as well as their first-degree relatives, have been removed, so this data set should serve as a useful reference of allele frequencies for severe disease studies—however, individuals with severe disease may still be included in the data set, albeit likely at a frequency equivalent to or lower than that seen in the general population. No statistical methods were used to predetermine sample size. The experiments were not randomised [19].

The analysis of GnomAD data was carried out from November 2017 to January 2018. All TTR (canonical transcript ENST00000237014) variants were selected from the gnomAD database (<http://gnomad.broadinstitute.org/>). The variants studied were missense variants for which a change of one amino acid for another occurred; variants of disease are those that affect protein folding and that induce the accumulation of misfolded proteins, not those leading to a non-functional truncated protein. Thus, splicing variants, stop codons or frameshift changes were eliminated. Intron variants, variants in non-coding regions, or synonymous variants were not taken into account.

The variables studied were classified according to their clinical significance and according to the different populations studied in gnomAD.

- In terms of the clinical significance, variants of the TTR gene were identified based on descriptions in other databases and in the literature. The variants found in the database, previously described in the literature, have been classified mainly according to their clinical significance, placing special interest on those described in the literature as disease producers.

The following sources were used: ClinVar database (<https://www.ncbi.nlm.nih.gov/clinvar/>), the Mutations in Hereditary Amyloidosis database (<http://www.amyloidosismutations.com/>), the Human Genetic Mutation database ([www.hgmd.cf.ac.uk/](http://www.hgmd.cf.ac.uk/)) and available literature from PubMed Central ([www.PubMed.com](http://www.PubMed.com)).

Following the recommendations of Human Genome Variation Society, the variants were classified as affects function, probably affects function, unknown, probably does not affect function and does not affect function [20].

It has been considered if these variants were described or not before or, also if there were or not changes described in that protein's position. In this study we have been mainly

**Table 1** Affects function variants of the TTR gene from the gnomAD database

	Variant	Allele count	Exome/genome number	Number of homozygotes	Allele frequency
Affects function variants	c.88T>C (p.(Cys30Arg))	1	15492	0	0.0000323
	c.122G>A (p.(Arg41Gln))	2	123121	0	0.0000081
	c.148G>A (p.(Val50Met))	25	123118	0	0.000102
	c.190T>C (p.(Phe64Leu))	14	138600	0	0.0000505
	c.227A>G (p.(His76Arg))	1	15491	0	0.0000323
	c.238A>G (p.(Thr80Ala))	1	123108	0	0.0000041
	c.250T>C (p.(Phe84Leu))	1	123106	0	0.0000041
	c.262A>T (p.(Ile88Leu))	6	138591	0	0.0000217
	c.301G>A (p.(Ala101Thr))	3	123095	0	0.0000122
	c.349G>T (p.(Ala117Ser))	2	138367	0	0.0000072
	c.424G>A (p.(Val142Ile))	417	138570	3	<b>0.00151</b>

Bold values indicate only that it is the highest allele frequency value in the group

**Table 2** Population descriptions of c.424G>A (p.(Val142Ile)) and c.148G>A (p.(Val50Met)) of the TTR gene

	c.424G>A (p.(Val142Ile))			c.148G>A (p.(Val50Met))		
	Allele count	%	Allele frequency	Allele count	%	Allele frequency
African	385	92.33	0.016	0	0	0
Latino	23	5.52	0.000668	2	8	0.0000596
European (Non-Finnish)	5	1.2	0.0000395	20	80	0.000179
European (Finnish)	0	0	0	1	4	0.0000448
South Asian	2	0.48	0.000065	0	0	0
Other	2	0.48	0.000309	2	8	0.000365
Total	417			25		

focused on the group of variants that produce disease, these variants are associated to the formation of misfolded proteins. For those variants which are probably pathological or those with an uncertain significance, most of them not described in the literature yet, these studies are not available.

The pathogenic potential of each unreported variant was classified as recommended by the American College of Medical Genetics and Genomics [21].

- In terms of populations, all variants analysed were classified according to the diverse ethnicities defined in the gnomAD database. The population classifications were as follows: non-Finnish European (45.71%), African/African American (8.67%), Latino (12.41%), Finnish (9.30%), Ashkenazi Jewish (3.66%), East Asian (6.81%), South Asian (11.10 %) and other (2.33%).

## Results

We found 71 missense variants in the TTR gene, with an average of 115,780 exomes studied. The variants were divided into the following categories: affects

function, probably affects function, unknown, probably does not affect function and does not affect function variants.

### Affects function variants

From the database search results, 11 variants were described in this group (Table 1). The frequency of variants that affect function was 1:230 people. The most frequently detected were c.424G>A (p.(Val142Ile)) and c.148G>A (p.(Val50Met)) which represented 88% and 5%, respectively, of all affect function variants detected.

The c.424G>A (p.(Val142Ile)) variant had a minor allelic frequency (MAF) of 0.00151 and a prevalence of 1:332. The most commonly represented population was African, with a prevalence of 1:31 people (MAF 0.01602). The c.148G>A (p.(Val50Met)) variant represented 5% of this group, resulting in an estimated prevalence of 1:4924 (MAF 0.000102). The distribution by subpopulations of these variants is presented in Table 2. By dividing this group by ethnicity, we found that in Europe, the prevalence was 1:1269 people, and the most frequent was c.148G>A (p.(Val50Met)), with one carrier per 2792 inhabitants (MAF 0.000179) (Table 3).

**Table 3** Variants of the TTR gene classified as affects function variants from the gnomAD database in the European population

Variant	Allele count	Exome/genome number	Number of homozygotes	Allele frequency
c.88T>C (p.(Cys30Arg))	1	1747	0	0.000067
c.148G>A (p.(Val50Met))	20	55847	0	0.000179
c.190T>C (p.(Phe64Leu))	1	63352	0	0.000008
c.227A>G (p.(His76Arg))	1	7507	0	0.000067
c.238A>G (p.(Thr80Ala))	1	55842	0	0.000009
c.250T>C (p.(Phe84Leu))	1	55838	0	0.000009
c.262A>T (p.(Ile88Leu))	5	63342	0	0.00004
c.301G>A (p.(Ala101Thr))	2	55824	0	0.000018
c.424G>A (p.(Val142Ile))	5	63331	0	0.00004

### Probably affects function variants

Seventeen variants were classified as probably affects function variants, and the allelic frequencies of those variants ranged from 0.0000041 to 0.0000646. The most frequently identified variants were c.235A>G (p.(Thr79Ala)), with an estimated prevalence of 1:11,191 people (0.0000447 MAF), and c.406T>C (p.(Tyr136His)), with an estimated prevalence of 1:17,588 people (MAF 0.0000284) (Table 4).

### Unknown variants

There were 29 variants that were classified as effect on function not known. The highest frequencies were for c.14G>A (p.(Arg5His)) and c.368G>A (p.(Arg123His)), with prevalence rates of 1:3959 people (MAF 0.000126) and 1:6927 people (MAF 0.0000722), respectively (Table 4).

### Probably does not affect function variants

Four variants were classified as probably does not affect function variants, with a prevalence of 1:15,371 (MAF 0.0000325). The c.355G>A (p.(Asp119Asn)) variant was the most common in this group. Other variants included c.98T>C (p.(Met33Thr)), c.362G>A (p.(Gly121Asp)) and c.383C>T (p.(Ala128Val)) (Table 4).

### Does not affect function variants

The most frequent variant in the group of no functional effect variants was c.76G>A (p.(Gly26Ser)), representing 94.5% of this group (MAF 0.0515) (Table 4).

## Discussion

We used the gnomAD database to estimate the population prevalence of ATTR-m. In most papers published about this

disease, the prevalence was estimated from the recorded data of patients already diagnosed. In our work, we started from a general population sample, in which we analysed allele frequency [9, 22]. Although ATTR-m is considered to be a rare disease, there is sufficient evidence to conclude that it is underdiagnosed [17, 18], and until now data of the true prevalence in the general population are not conclusive.

After analysing all exonic variants present in the gnomAD database, 71 missense variants were detected. Only 28 had been previously reported, of which 11 were considered to be amyloidogenic [6]. We classified 17 of the variants as variants that probably affects function; these had not been described before, but pathological changes in the same position had already been reported. Only 6 of the 29 variants that we determined to be of unknown significance had been described as such in the bibliography (c.14G>A (p.(Arg5His)), c.25C>T (p.(Leu9Phe)), c.52T>G (p.(Ser18-Ala)), c.68C>T (p.(Thr23Met)), c.140A>G (p.(Asn47Ser)) and c.368G>A (p.(Arg123His))) [6], the others 23 are not described variants or with no changes reported in that position. Within the probably does not affect function group, the c.355G>A (p.(Asp119Asn)) variant was included, all variants with similar amino acid modifications described previously as does not affect function (c.98T>C (p.(Met33Thr)), c.362G>A (p.(Gly121Asp)) and c.383C>T (p.(Ala128Val)) [6]. Finally, in the group of does not affect function variants, all variants have been previously described, among which the most common was c.76G>A (p.(Gly26Ser)), which represented 94.5% of the group (MAF 0.0515).

If only the 11 recognised affects function variants were considered, the prevalence of amyloidogenic variants in the general population was 1:230 cases. If data from the most represented group in the gnomAD database (that is, participants with European ancestry) were considered, the prevalence decreased to 1:1269 people, probably due to the influence of the African population on the whole population. If we included variants classified as probably affects function, the overall prevalence was 1:207.

**Table 4** Variants of the TTR gene from the gnomAD database. Classified as probably affects function, unknown, probably does not affect function and does not affect function variants

	Variant	Allele count	Exome/genome number	Number of homozygotes	Allele frequency	
Probably affects function variants	c.118G>C (p.(Val40Leu))	1	123121	0	0.00000406	
	c.119T>C (p.(Val40Ala))	1	123127	0	0.00000406	
	c.127A>C (p.(Ser43Arg))	1	15485	0	0.0000323	
	c.208A>G (p.(Ser70Gly))	1	123081	0	0.00000406	
	c.209G>A (p.(Ser70Asn))	6	123083	0	0.0000244	
	c.212A>C (p.(Glu71Ala))	1	15481	0	0.0000323	
	c.215C>A (p.(Ser72Tyr))	1	123086	0	0.00000406	
	c.226C>T (p.(His76Tyr))	1	123090	0	0.00000406	
	c.235A>G (p.(Thr79Ala))	11	123104	0	0.0000447	
	c.226A>G (p.(Tyr89Cys))	1	123103	0	0.00000406	
	c.268A>C (p.(Lys90Gln))	2	138593	0	0.00000722	
	c.301G>T (p.(Ala101Ser))	1	123095	0	0.00000406	
	c.301A>C (p.(Ile104Leu))	1	123089	0	0.00000406	
	c.367C>G (p.(Arg123Gly))	2	15483	0	0.0000646	
	c.367C>T (p.(Arg123Cys))	3	123074	0	0.0000122	
	c.380T>C (p.(Ile127Thr))	1	123105	0	0.00000406	
	c.406T>C (p.(Tyr136His))	7	123118	0	0.0000284	
	Unknown variants	c.11A>T (p.(His4Leu))	1	123081	0	0.00000406
		c.13C>T (p.(Arg5Cys))	3	138566	0	0.0000108
		c.14G>A (p.(Arg5His))	35	138570	1	0.000126
c.23T>A (p.(Leu8His))		1	123101	0	0.00000406	
c.25C>T (p.(Leu9Phe))		2	123098	0	0.00000812	
c.37G>A (p.(Gly13Arg))		1	123095	0	0.00000406	
c.40C>G (p.(Leu14Val))		1	123099	0	0.00000406	
c.50T>C (p.(Val17Ala))		1	123090	0	0.00000406	
c.52T>G (p.(Ser18Ala))		1	123091	0	0.00000406	
c.65C>T (p.(Pro22Leu))		1	15490	0	0.0000323	
c.68C>T (p.(Thr23Met))		7	123046	0	0.0000284	
c.70G>A (p.(Gly24Ser))		1	123117	0	0.00000406	
c.71G>A (p.(Gly24Asp))		1	123116	0	0.00000406	
c.125G>A (p.(Gly42Asp))		1	123117	0	0.00000406	
c.136A>G (p.(Ile46Val))		7	123120	0	0.0000284	
c.140A>G (p.(Asn47Ser))		7	138599	0	0.0000253	
c.151C>A (p.(His51Asn))		4	123118	0	0.0000162	
c.152A>C (p.(His51Pro))		1	123121	0	0.00000406	
c.169G>A (p.(Ala57Thr))		1	123123	0	0.00000406	
c.170C>A (p.(Ala57Asp))		2	123122	0	0.00000812	
c.246G>C (p.(Glu82Asp))		3	123111	0	0.0000122	
c.305T>C (p.(Leu102Pro))		1	123102	0	0.00000406	
c.358T>G (p.(Ser120Ala))		1	122574	0	0.00000408	
c.368G>A (p.(Arg123His))		20	138547	0	0.0000722	
c.377C>A (p.(Thr126Asn))		1	15487	0	0.0000323	
c.409T>A (p.(Ser137Thr))		1	123115	0	0.00000406	
c.413C>T (p.(Thr138Ile))		1	123112	0	0.00000406	
c.431A>G (p.(Asn144Ser))		2	123105	0	0.00000812	

**Table 4** (continued)

	Variant	Allele count	Exome/genome number	Number of homozygotes	Allele frequency
Probably does not affect function variants	c.441A>T (p.(Glu147Asp))	2	123093	0	0.00000812
	c.98T>C (p.(Met33Thr))	1	123126	0	0.0000041
	c.362G>A (p.(Gly121Asp))	2	122012	1	0.0000082
Does not affect function variants	c.383C>T (p.(Ala128Val))	1	15487	0	0.0000323
	c.355G>A (p.(Asp119Asn))	9	138343	0	0.0000325
	c.76G>A (p.(Gly26Ser))	14283	138587	520	<b>0.0515</b>
	c.280G>C (p.(Asp94His))	13	138590	0	0.0000469
	c.328C>A (p.(His110Asn))	152	138574	0	0.000548
	c.361G>A (p.(Gly121Ser))	10	122896	0	0.0000407
	c.370C>T (p.(Arg124Cys))	15	138560	1	0.0000541
	c.371G>A (p.(Arg124His))	205	138537	2	0.00074
	c.385G>A (p.(Ala129Thr))	17	138582	0	0.0000613
	c.386C>T (p.(Ala129Val))	5	138590	0	0.000018
	c.416C>T (p.(Thr139Met))	406	138589	0	0.00147
	c.437A>G (p.(Lys146Arg))	88	138569	0	0.000318

Bold values indicate only that it is the highest allele frequency value in the group

The results show a much higher frequency of amyloidogenic variants of the expected prevalence of disease according to the heretofore described 1:100,000 inhabitants or less [9, 22]. Part of this difference could be explained by incomplete penetrance of the disease, but we believe that there are likely other factors contributing to this fact, such as a fundamentally the underdiagnosis of the disease, as previously discussed [17].

In terms of the amyloidogenic variants described, the most common in the gnomAD database was the c.424G>A (p.(Val142Ile)), with a carrier frequency of 1:332 (MAF 0.00151), followed by c.148G>A (p.(Val50Met)) with a frequency of 1:4924 people (MAF 0.000102). For both variants, ethnic differences were observed—c.424G>A (p.(Val142Ile)) was especially prevalent in the African population (MAF 0.01602; prevalence of 1:31) and c.148G>A (p.(Val50Met)) was especially prevalent in the European population (MAF 0.000179; prevalence of 1:2,792).

The c.424G>A (p.(Val142Ile)) variant is currently recognised as the most common cause of hereditary amyloid heart disease worldwide, which is consistent with the results of our work in which it represents 88% of all affect function variants. Our data show that 92.33% of the carriers of this variants detected in the gnomAD database were of African descent (385 of 417 registered c.424G>A (p.(Val142Ile)) variant), 5.52% were of Latino descent and 1.2% were of European descent. According to the THAOS registry, data from the United States demonstrated similar proportions; of the 91 American patients registered as

c.424G>A (p.(Val142Ile)) carriers, 79 (86.8%) were of African descent compared to 6 (6.6%) of Caucasian descent [23].

There are limited data on the overall prevalence of the c.424G>A (p.(Val142Ile)) variant, but there are several studies that focus on the African-American population. In 1996, Jacobson et al. concluded that the allelic frequency of this variant in people of African descent was 0.02, which was equivalent to 4% of that population [24]. In a new study by Jacobson et al. in 2015, a study of 1688 African-Americans showed an allelic prevalence of 0.0193 [25]. In data released by Buxbaum et al. in 2017, the allele frequency in an African-American population was estimated to be 0.017 [26]. An analysis of gnomAD data resulted in a similar pattern, with a prevalence of 1:31, which indicated that 3.2% of the African population carried the c.424G>A (p.(Val142Ile)) variant (MAF 0.01602).

A recent study of Jacobson et al in 2016, establishes the prevalence of c.424G>A (p.(Val142Ile)) in the African population [27]. The results of the countries studied show a higher prevalence in West African countries (allele frequency 0.0253) and a mean allele frequency of 0.011 in the rest of the countries. These data support the idea of a small number of founder carriers of the amyloidogenic allele in southern West Africa

Revised data showed that the single amyloidogenic variant that appears in homozygosity is c.424G>A (p.(Val142Ile)), probably due to the low number of total cases of other variants. According to the present study,

1.56% of c.424G>A (p.(Val142Ile)) carriers are homozygous; this ratio is slightly below ratios that have been described so far (3.0–9.2%) [28–31].

Accepting that c.424G>A (p.(Val142Ile)) is much more prevalent in African populations, it is possible to find patients with the c.424G>A (p.(Val142Ile)) variant worldwide. According to data from gnomAD, 1:12,666 Europeans carry this variant (MAF 0.00003948), 1:748 people in a Latin population (MAF 0.0006683) and 1:7695 people in a South Asian population carry this amyloidogenic variant (MAF 0.00006497). In the literature there are few cases of patients with no African descent, which were presented as sporadic cases, and there was a small series of cases described in Italy by Capelli et al., in which 5 of the 33 Caucasian patients diagnosed at its centre carried the variant c.424G>A (p.(Val142Ile)) [32–34].

The second most common amyloidogenic variant identified in the gnomAD database was c.148G>A (p.(Val50Met)) (MAF 0.000102; prevalence of 1:4924), which represented 5% of all affect function variants. To date, c.148G>A (p.(Val50Met)) has been described in many patients globally, and it represents more than 50% of the cases of familial amyloid polyneuropathy that have been described in the literature [9]. This discrepancy in the percentage is probably due to the increased aggressiveness of neurological manifestations; a greater awareness by being the first described variant; and, above all, that is present in areas considered endemic, for example, Portugal, Sweden or Japan, where this variant is responsible for almost all cases [7, 9, 12].

As for the association between c.148G>A (p.(Val50Met)) and an ethnic group, our data are consistent with previously reported data in terms of the association with European and Latin-American populations [35]. Of the 25 cases identified in gnomAD, 20 people belong to European populations, with a prevalence of 1:2792 (MAF 0.000179), and 2 people were of Latin descent, with a prevalence of 1:8395 (MAF 0.00005956).

The remaining variants that were classified as affects function occurred less frequently, including c.262A>T (p.(Ile88Leu)), with a prevalence of 1:23,098 (MAF 0.00002165). This variant is primarily associated with people of European ancestry, in which the prevalence is 1:12,668 (MAF 0.00003947). The clinical manifestations are almost exclusively cardiac; in fact, in a number of cases of Caucasian patients, c.262A>T (p.(Ile88Leu)) was the variant that was found most frequently [36]. This is in contrast to the data in African populations, in which the most frequent cardiac variant was c.424G>A (p.(Val142Ile)).

The c.190T>C (p.(Phe64Leu)) variant, described in the literature as producing a mixed phenotype [37, 38], was identified in 1:9900 persons studied in the gnomAD

database (MAF 0.00005051). It was mainly present in people of African descent, with a prevalence of 1:924 (MAF 0.0005411), and only one case was identified in a person in the European population (MAF 0.000007892).

The c.238A>G (p.(Thr80Ala)) variant that only appeared in the European population in our study, with a prevalence of 1:7507 (MAF 0.00006660), has been described previously in a cluster in Northern Ireland, where it was estimated that 1.1% of the population carries these variant, and in isolated cases in England and Scotland [22, 39].

Cases of c.349G>T (p.(Ala117Ser)) that appeared in the gnomAD database were all people of an Eastern Asia population, where the prevalence was estimated at 1:9,431 (MAF 0.000106). This amyloidogenic variant has been described in the literature as one of the most common in data from Chinese populations [40], which is consistent with our findings.

In the group of variants that probably affects function there is predominance according to ethnicity in some variants. The variants c.367C>T (p.(Arg123Cys)) and c.367C>G (p.(Arg123Gly)) appear in gnomAD only in European population, other frequent variants as c.235A>G (p.(Thr79Ala)) or c.406T>C (p.(Thr136His)) are represented, almost exclusively, for Latin population. Therefore, in this group there is also an association between the variants and the ethnic groups.

The data derived from the present study supports the idea of an underdiagnosis of ATTR-m. Some current studies indicate this trend, concluding that the overall prevalence figure is higher than what has been reported to date [7, 17]. The prevalence of ATTR-m polyneuropathy has been estimated to be 5000 to 10,000 cases of disease worldwide, and some authors have concluded that it could be as high as 40,000 cases [17]. Other studies indicated that 32% of patients are undiagnosed [35]. For ATTR-m with cardiomyopathy, there are an estimated 40,000 cases [17]. This is thought to be an underestimation because, for example, 5% patients with hypertrophic cardiomyopathy have ATTR-m [18].

According to the data obtained in this work and comparing it with the current literature, the prevalence of the c.424G>A (p.(Val142Ile)) variant is well established in the African population, but not in the rest of the populations, with lower frequency such as in the European population. The problem of underdiagnosis of this variant is frequent in all population. For other variants, including c.148G>A (p.(Val50Met)), there is an underestimation of prevalence and underdiagnosis more notable in non-endemic areas than in those regions recognised as endemic. This may be caused in part due to the low clinical suspicion and the lack of family history of the patients in the non-endemic areas.

Because of large population studies like gnomAD, the prevalence of variants involved in various diseases can be

studied; [41–43] however, one of the apparent limitations of the use of these databases in amyloidosis ATTR is that the prevalence of amyloidogenic variants in the population is not the same as disease prevalence, so the data should be taken with caution. In addition, phenotype data are not available for gnomAD samples. That is, this study describes the frequency of people carriers, presymptomatic carriers should be identified and receive adequate follow-up so that in the event of the occurrence of symptoms, pharmacological treatment should be instituted without delay [12].

Another fact to take into account is that affects function variants are associated to the formation of misfolded proteins. But for new variants included in other groups, these studies are not available, it would be an interesting study to better understand of the new variants.

In addition, it should be considered that not all populations are equally represented, with the European population being the majority. Populations like Asian or Latino should be more present to have more appropriate conclusions to these groups.

In conclusion, although recent data suggest that the actual prevalence of amyloidosis ATTR-m could be much higher than what has been reported to date, data provided by an analysis of the gnomAD database are even more compelling, opening the possibility that might be hundreds of thousands of people having potentially disease-causing variants.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Wechalekar AD, Gillmore JD, Hawkins PN. Systemic amyloidosis. *Lancet*. 2016;387:2641–54.
- Sipe JD, Benson MD, Buxbaum JN, Ikeda S, Merlini G, Saraiva MJM, et al. Amyloid fibril proteins and amyloidosis: chemical identification and clinical classification International Society of Amyloidosis 2016 Nomenclature Guidelines. *Amyloid*. 2016;23:209–13.
- Gertz MA, Benson MD, Dyck PJ, Grogan M, Coelho T, Cruz M, et al. Diagnosis, prognosis, and therapy of transthyretin amyloidosis. *J Am Coll Cardiol*. 2015;66:2451–66.
- Sekijima Y. Transthyretin (ATTR) amyloidosis: clinical spectrum, molecular pathogenesis and disease-modifying treatments. *J Neurol Neurosurg Psychiatry*. 2015;86:1036–43.
- Ando Y, Coelho T, Berk JL, Cruz MW, Ericzon B, Ikeda S, et al. Guideline of transthyretin-related hereditary amyloidosis for clinicians. *Orphanet J Rare Dis*. 2013;8:31.
- Mutations in Hereditary Amyloidosis [Internet]. Available from <http://www.amyloidosismutations.com/>
- Galant NJ, Westermark P, Higaki JN, Chakrabarty A. Transthyretin amyloidosis: an under-recognized neuropathy and cardiomyopathy. *Clin Sci*. 2017;131:395–409.
- Coelho T, Maurer MS, Suhr OB. THAOS – The Transthyretin Amyloidosis Outcomes Survey: initial report on clinical manifestations in patients with hereditary and wild-type transthyretin amyloidosis. *Curr Med Res Opin*. 2013;29:63–76.
- Hawkins PN, Ando Y, Dispenzeri A, Gonzalez-Duarte A, Adams D, Suhr OB. Evolving landscape in the management of transthyretin amyloidosis. *Ann Med*. 2015;47:625–38.
- Parman Y, Adams D, Obici L, Galán L, Guergueltcheva V, Suhr OB, et al. Sixty years of transthyretin familial amyloid polyneuropathy (TTR-FAP) in Europe: where are we now? A European network approach to defining the epidemiology and management patterns for TTR-FAP. *Curr Opin Neurol*. 2016;29(Suppl 1):S3–S13.
- Reinés JB, Vera TR, Martín UM, Serra HA, Millán JMD, Lezaun CG, et al. Epidemiology of transthyretin-associated familial amyloid polyneuropathy in the Majorcan area: Son Llätzer Hospital descriptive study. *Orphanet J Rare Dis*. 2014;9:29.
- Obici L, Kuks JB, Buades J, Adams D, Suhr OB, Coelho T, et al. Recommendations for presymptomatic genetic testing and management of individuals at risk for hereditary transthyretin amyloidosis. *Curr Opin Neurol*. 2016;29(Suppl 1):S27.
- Conceição I, González-Duarte A, Obici L, Schmidt HHJ, Simoneau D, Ong ML, et al. Red-flag symptom clusters in transthyretin familial amyloid polyneuropathy. *J Peripher Nerv Syst*. 2016;21:5–9.
- Hellman U, Alarcon F, Lundgren H-E, Suhr OB, Bonaiti-Pellié C, Planté-Bordeneuve V. Heterogeneity of penetrance in familial amyloid polyneuropathy, ATTR Val30Met, in the Swedish population. *Amyloid*. 2008;15:181–6.
- Lemos C, Coelho T, Alves-Ferreira M, Martins-da-Silva, A, Sequeiros, J, Mendonça, D, et al. Overcoming artefact: anticipation in 284 Portuguese kindreds with familial amyloid polyneuropathy (FAP) ATTRV30M. *J Neurol Neurosurg Psychiatry*. 2014;85:326–30.
- Molina OG, Judge D, Campbell W, Chahal H, Mugmon M. Transthyretin cardiac amyloidosis: an under-diagnosed cause of heart failure. *J Community Hosp Intern Med Perspect*. 2014;4:25500.
- Schmidt HH, Waddington-Cruz M, Böttner MF, Carter JA, Chopra AS, Hopps M, et al. Estimating the global prevalence of transthyretin familial amyloid polyneuropathy. *Muscle Nerve*. 2018;57:829–37.
- Damy T, Costes B, Hagège AA, Donal E, Eicher JC, Slama M, et al. Prevalence and clinical phenotype of hereditary transthyretin amyloid cardiomyopathy in patients with increased left ventricular wall thickness. *Eur Heart J*. 2015;37:1826–34.
- Lek, M, Karczewski, K, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536:285–91.
- Human Genome Variation Society. [Internet]. Available from <http://varnomen.hgvs.org>
- Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405–24.
- Lachmann HJ, Booth DR, Booth SE, Bybee A, Gilbertson JA, Gillmore JD, et al. Misdiagnosis of hereditary amyloidosis as AL (primary) amyloidosis. *N Engl J Med*. 2002;346:1786–91.
- Maurer MS, Hanna M, Grogan M, Dispenzieri A, Wittes R, Drachman B, et al. Genotype and phenotype of transthyretin cardiac amyloidosis in the United States: THAOS (The Transthyretin Amyloid Outcome Survey). *J Am Coll Cardiol*. 2016;68:161–72.

24. Jacobson DR, Pastore R, Pool S, Malendowicz S, Kane I, Shivji A, et al. Revised transthyretin Ile 122 allele frequency in African-Americans. *Hum Genet.* 1996;98:236–8.
25. Jacobson DR, Alexander AA, Tagoe C, Buxbaum JN. Prevalence of the amyloidogenic transthyretin (TTR) V122I allele in 14 333 African-Americans. *Amyloid.* 2015;22:171–4.
26. Buxbaum J, Alexander A, Koziol J, Tagoe C, Fox E, Kitzman D. Significance of the amyloidogenic transthyretin Val 122 Ile allele in African Americans in the Arteriosclerosis Risk in Communities (ARIC) and Cardiovascular Health (CHS) Studies. *Am Heart J.* 2010;159:864–70.
27. Jacobson DR, Alexander AA, Tagoe C, Garvey WT, Williams SM, Tishkoff S, et al. The prevalence and distribution of the amyloidogenic transthyretin (TTR) V122I allele in Africa. *Mol Genet Genom Med.* 2016;4:548–56.
28. Dungu JN, Papadopoulou SA, Wykes K, Mahmood I, Marshall J, Valencia O, et al. Afro-Caribbean Heart Failure in the United Kingdom: cause, outcomes, and ATTR V122I cardiac amyloidosis. *Circ Hear Fail.* 2016;9:e003352.
29. Yamashita T, Asl KH, Yazaki M, Benson MD. A prospective evaluation of the transthyretin Ile122 allele frequency in an African-American population. *Amyloid.* 2005;12:127–30.
30. Reddi H V., Jenkins S, Theis J, Thomas BC, Connors LH, Van Rhee F, et al. Homozygosity for the V122I mutation in transthyretin is associated with earlier onset of cardiac amyloidosis in the African American population in the seventh decade of life. *J Mol Diagn.* 2014;16:68–74.
31. Gillmore JD, Booth DR, Pepys MB, Hawkins PN. Hereditary cardiac amyloidosis associated with the transthyretin Ile122 mutation in a white man. *Heart.* 1999;82:e2–e2.
32. Cappelli F, Frusconi S, Bergesio F, Grifoni E, Fabbri A, Giuliani C, et al. The Val142Ile transthyretin cardiac amyloidosis: Not only an Afro-American pathogenic variant? A single-centre Italian experience. *J Cardiovasc Med.* 2016;17:122–5.
33. Asl KH, Nakamura M, Yamashita T, Benson MD. Cardiac amyloidosis associated with the transthyretin Ile122 mutation in a Caucasian family. *Amyloid.* 2001;8:263–9.
34. Nienhuis HL, Bijzet J, Hazenberg BP. The prevalence and management of systemic amyloidosis in Western countries. *Kidney Dis.* 2016;2:10–19.
35. Plante-Bordeneuve V. Transthyretin familial amyloid polyneuropathy: an update. *J Neurol.* 2018;265:976–83.
36. Rapezzi C, Quarta CC, Obici L, Perfetto F, Longhi S, Salvi F, et al. Disease profile and differential diagnosis of hereditary transthyretin-related amyloidosis with exclusively cardiac phenotype: an Italian perspective. *Eur Heart J.* 2012;34:520–8.
37. Damy T, Judge DP, Kristen AV, Berthet K, Li H, Aarts J. Cardiac findings and events observed in an open-label clinical trial of tafamidis in patients with non-Val30Met and non-Val122Ile hereditary transthyretin amyloidosis. *J Cardiovasc Transl Res.* 2015;8:117–27.
38. Wixner J, Mundayat R, Karayal ON, Anan I, Karling P, Suhr OB. THAOS: gastrointestinal manifestations of transthyretin amyloidosis - common complications of a rare disease. *Orphanet J Rare Dis.* 2014;9:61
39. Reilly MM, Staunton H, Harding AE. Familial amyloid polyneuropathy (TTRala60) in north west Ireland: a clinical, genetic, and epidemiological study. *J Neurol Neurosurg Psychiatry.* 1995;59:45–9.
40. Meng LC, Lyu H, Zhang W, Liu J, Wang ZX, Yuan Y. Hereditary transthyretin amyloidosis in eight Chinese families. *Chin Med J.* 2015;128:2902.
41. Hughes BG, Harrison PM, Hekimi S. Estimating the occurrence of primary ubiquinone deficiency by analysis of large-scale sequencing data. *Sci Rep.* 2017;7:17744.
42. Song W, Gardner SA, Hovhannisyan H, Natalizio A, Weymouth KS, Chen W, et al. Exploring the landscape of pathogenic genetic variation in the ExAC population database: insights of relevance to variant classification. *Genet Med.* 2016;18:850–4.
43. Nicolas G, Charbonnier C, Campion D, Veltman JA. Estimation of minimal disease prevalence from population genomic data: application to primary familial brain calcification. *Am J Med Genet Part B.* 2018;177:68–74.