Scientific Research Publishing

Open Journal of Genetics, 2026, 16(2), 59-74
https://www.scirp.org/journal/ojgen
ISSN Online: 2162-4461
ISSN Print: 2162-4453

A Study of Variants in the *MTCYB* Gene in Atrial Fibrillation: A Case Series of 30 Patients at FANN Hospital in Dakar

Marie Khadidiatou Sarr, Mbacké Sembène

Genomics Laboratory, Department of Animal Biology, Faculty of Science and Technology, Cheikh Anta Diop University (UCAD), Dakar, Senegal
Email: mbacke.sembene@ucad.edu.sn

How to cite this paper: Sarr, M.K. and Sembène, M. (2026) A Study of Variants in the *MTCYB* Gene in Atrial Fibrillation: A Case Series of 30 Patients at FANN Hospital in Dakar. *Open Journal of Genetics*, **16**, 59-74.
https://doi.org/10.4236/ojgen.2026.162006

Received: May 14, 2026
Accepted: June 23, 2026
Published: June 26, 2026

Abstract

This case series was conducted to help better characterize the *MTCYB* gene variants identified in patients with atrial fibrillation. It involved 30 hospitalized patients with atrial fibrillation from the cardiology centre at FANN Hospital in Dakar. Tissue samples taken from each patient were sent to the laboratory and subjected to genetic analysis, which enabled the determination of the gene's polymorphism, diversity, differentiation and genetic structure. The *MTCYB* gene amplification products obtained from the DNA extracts revealed a migration profile of a size estimated at approximately 1246 bp. All the non-synonymous mutations identified are implicated either in the structure and function of the protein, in its pathogenicity, or in its stability, and 15 of these were predicted by in silico tools to potentially affect all these aspects simultaneously. One of the synonymous mutations, m.15301G>A, which is thought to have no effect on protein stability, was found to be quite common in the population (14.54%) and is reported as a risk factor for high-altitude sickness. The functional and structural impact of the variants has been demonstrated, and four of them have effects that could alter the protein's secondary structure through steric hindrance, the replacement of a proline in a cis position, and the replacement of a tyrosine (TYR > 6.3%) by a lysine (RSA < 20%) in a position deep within the hydrophobic core of cytochrome b. The *MTCYB* gene encodes cytochrome b, a central subunit of mitochondrial complex III involved in electron transport and proton translocation across the inner mitochondrial membrane. The variants identified in this case series may warrant further functional studies to determine whether they could contribute to mitochondrial dysfunction in the context of atrial fibrillation.

Keywords

Atrial Fibrillation, Genetics, *MTCYB* Gene, Mutations

1. Introduction

Cardiovascular diseases account for one third of deaths worldwide [1]. Atrial fibrillation is one such arrhythmic disorder that is fairly common in adults and whose prevalence is constantly changing. Indeed, it occurs in elderly patients who often suffer from a number of underlying conditions that act as predisposing factors for cardiovascular disease [2]. It would not be straightforward to conclude that genetics plays a significant role in this heart condition. However, in cases where the condition arises suddenly in elderly individuals without comorbidities, the question arises and a genetic influence becomes the most likely explanation. Genetic causes may be linked to mitochondrial dysfunction and lead to oxidative damage to proteins and DNA [3]. As a gene that encodes a protein, the mitochondrially encoded cytochrome b (*MTCYB*) gene is involved in several processes, including proton transport coupled with electron transport [4]. The aim of this study was to analyze this gene in order to determine its polymorphism and its role in the onset of the disease. Genetics has a major role in determining the basis of diseases, based on the candidate gene approach, which is a strategy for identifying genetic causes.

2. Method

The study involved 30 patients hospitalized for atrial fibrillation at the Cardiology Centre of FANN Hospital, comprising 73% women and 27% men, with a sex ratio of 0.36. All patients included had rheumatic mitral valve disease complicated by atrial fibrillation, requiring surgical mitral valve replacement. The link between rheumatic mitral stenosis and atrial fibrillation has been well established, as chronic mitral valve disease leads to progressive dilation of the left atrium, followed by atrial fibrillation. The diagnosis was confirmed on the basis of clinical, echocardiographic and surgical findings. All types of atrial fibrillation (permanent, paroxysmal, persistent) were included without distinction. For each patient undergoing surgery, a fresh tissue sample was taken from the surgical site (the junction of the left atrium and the right superior pulmonary vein), placed in a dry tube and stored at 20˚C. Clinical records containing information on the demographic and clinical-pathological characteristics of each patient were made available to us. After collection, the cardiac tissue samples were sent directly to the genomics laboratory of the Department of Animal Biology at the Faculty of Science and Technology of Cheikh Anta Diop University in Dakar, where the tissues were preserved in 96% alcohol for various molecular analyses.

The 20 samples consisted of biological specimens already available at the laboratory of the Faculty of Science and Technology at Cheikh Anta Diop University

in Dakar at the time of the study. These were stored blood samples taken from apparently healthy subjects with no known history of heart disease or atrial fibrillation.

Using the standard method described in the Zymo Research Kit [5], total DNA was extracted from the tissue, followed by polymerase chain reaction (PCR) amplification of the *MTCYB* gene using the following primers: 5'-CGGACTACAACCACGACCAA-3' and R: 5'-TCCGGTTTACAAGACTGGTGT-3'. Once the PCR was complete, the products were kept at 10˚C in the thermal cycler. The size of the DNA fragments was approximately assessed using 6 µl of a 100-base-pair Purple Ladder size marker. After electrophoresis, the gel was placed on a transilluminator to visualize the results. Subsequently, the PCR products to which the primers had bound were sequenced using the method of F. Sanger (1977), which is based on a specific PCR reaction, utilizing, in addition to the usual components (template DNA, polymerase, primers, dNTPs, Mg^{2+}), modified nucleotides: deoxyribonucleotides (ddNTPs) [6]. Sequencing reactions were carried out in an MJ Research PTC-225 Peltier thermal cycler using the ABIPRISM BigDye™ Terminator Cycle kits. Each sample was sequenced using the forward primer, with 10 µl of PCR product. The fluorescent fragments were purified using the BigDye Xterminator purification protocol. The samples were suspended in distilled water and subjected to electrophoresis in an ABI 3730xl sequencer (Applied Biosystems).

- **Analysis of conserved regions identified by the primers**

The *MTCYB* gene, the only subunit of complex III (ubiquinol-cytochrome c reductase) encoded by mitochondrial DNA, spans 1140 bp between nucleotide positions 14,747 and 15,887 of the revised Cambridge reference sequence (rCRS), and encodes a protein of 380 amino acids organized into eight transmembrane helices [7]. The *MTCYB* protein is predominantly hydrophobic and contains two haem groups as well as two functional ubiquinol-binding sites, designated the Qo (oxidation) site and the Qi (reduction) site, which are central to the catalytic mechanism of complex III. The design of the primers used in this study is based on targeting conserved regions of the gene, a strategy founded on the high degree of evolutionary conservation of *MTCYB* across eukaryotic species [8]. The effectiveness of universal primers targeting cytochrome b lies precisely in their binding to conserved regions flanking variable sequences, which allows for efficient amplification even on degraded samples or those from distantly related species. Functionally, *MTCYB* is a protein that has been highly conserved throughout evolution, containing 8 to 9 transmembrane domains and two haem groups with distinct redox potentials—b-562 (high-energy haem) and b-566 (low-energy haem)—which are essential for electron transfer within the mitochondrial respiratory chain [9]. Recent studies have also demonstrated that the C-terminal region of cytochrome b plays an indispensable structural role in the assembly of the bc1 complex, and that its integrity determines the functional biogenesis of complex III [10]. Precise mapping of mutations in the conserved regions targeted by the primers. It is therefore essential to distinguish pathogenic variants from silent polymorphisms, as

these variants are frequently found in the healthy population but may also be associated with a spectrum of mitochondrial or common diseases depending on their location within the functional domains of *MTCYB* [8].

In this study, all 15 mutations were checked against the canonical UniProt reference sequence P00156 (rCRS, NC_012920.1, positions 14747–15887) (Table 1).

The W337/P342/P346/Y358/F359/P367 cluster is located in the transmembrane helix G (Qo domain—ubiquinol binding site), a functionally critical region effectively described as a hotspot in cytochrome b.

- W337 (Trp337) is a highly conserved residue involved in the Qo site of complex III.
- P367 is located in a region of significant conformational constraint important for the stability of the helix.

All numbering is consistent with the rCRS and the canonical UniProt sequence P00156.

Table 1. Localization of mutations in mitochondrial "hotspots".

Position rCRS	Mutations	Positions	Reference residue	Residue found	Status
m.14905G>C	M53I	53	M	M	Correct
m.14954C>A	P261T	261	P	P	Correct
m.15176C>CT	L335Y	335	L	L	Correct
m.15179C>CT	T336H	336	T	T	Correct
m.15182T>TG	W337P	337	W	W	Correct
m.15182T>TG	W337Q	337	W	W	Correct
m.15182A>AG	W337N	337	W	W	Correct
m.15197A>AG	P342Q	342	P	P	Correct
m.15209T>C	P346I	346	W	W	Correct
m.15209T>TA	Y358K	358	Y	Y	Correct
m.15248C>CT	F359L	359	F	F	Correct
m.15248C>CT	F359S	359	F	F	Correct
m.15272C>A	P367N	367	P	P	Correct
m.15272C>A	P367Q	367	P	P	Correct
m.15272C>A	P367T	367	P	P	Correct

- **Presence of mutations in pathogen databases (MITOMAP, ClinVar, GnomAD)**

All variants of the *MTCYB* gene identified by Mutation Surveyor were systematically cross-referenced against three major databases to assess their novelty and potential pathogenic significance. Firstly, the variants were checked against MITOMAP, which provided information on the evolutionary conservation of the

relevant residues and on previously reported pathogenic mutations. Next, the variants were searched for in ClinVar to obtain their classification based on clinical significance. Finally, the variants were searched for in the Genome Aggregation Database (gnomAD) to determine whether the identified variants correspond to rare or common polymorphisms within the population. All database queries were performed using absolute rCRS positions (NC_012920.1) and HGVS nomenclature, at both the nucleotide and protein levels.

A comparison of the fifteen *MTCYB* variants with the MITOMAP, ClinVar and gnomAD databases revealed distinct patterns in terms of novelty and conservation. Thirteen variants (M53I, P261T, L335Y, T336H, W337P, W337Q, P342Q, P346I, Y358K, F359L, F359S, P367N, P367Q, P367T) were absent from all three databases, suggesting that they represent novel *MTCYB* variants not previously reported in the literature. Two variants (p.Trp337Asn and p.Pro346Ile) were recorded as variants of uncertain significance (VUS) in gnomAD, indicating their presence in the general population without established clinical significance (Table 2).

An evolutionary conservation analysis carried out using MITOMAP revealed that the majority of the residues in question are highly conserved across species, with ten variants occurring at positions that are 95% - 100% conserved (M53I, L335Y, T336H, W337P, W337Q, W337N, P342Q, P346I, Y358K, P367N, P367Q, P367T), two variants at positions conserved at 97% (P342Q, Y358K) and three variants at positions exhibiting lower levels of conservation (P261T at 22%, F359L and F359S at 17%). The high degree of evolutionary conservation observed for the majority of the affected residues confirms the potential functional importance of these positions within the cytochrome b protein and suggests that the identified variants could impair the structure and function of mitochondrial complex III.

Variants spanning residues 335 to 359 were associated with insertions and deletions resulting in frameshift mutations relative to rCRS, as confirmed by analysis using Mutation Surveyor. These structural variants were absent from all the databases consulted, further confirming their novel nature. The other substitution variants were caused by specific nucleotide changes within individual codons, which differ from any previously reported mutations at the same residues, confirming their status as novel variants.

Table 2. *MTCYB* variant annotation across MITOMAP, ClinVar, and gnomAD databases.

Variants	Conclusions
M53I, P261T, P346I, P367N/Q/T	These variants didn't appear in the MITOMAP, ClinVar and gnomAD databases; they are therefore new variants of the *MTCYB* gene. The nucleotide mutation (C > A) affects a different position within the codon compared with previously reported variants at the same residue.
L335Y→F359S	Variants present in the region spanning residues 335 to 359 have been associated with insertions and deletions confirmed by Mutation Surveyor, resulting in frameshift mutations relative to the rCRS. These indels were absent from the MITOMAP, ClinVar and gnomAD databases.

3. Results

- **General characteristics of the population**

The results showed that, out of a total of 30 cases, 73% were women and 27% were men, with a sex ratio of 0.36. The mean age of the patients was 33.93 ± 17.28 years, ranging from 14 to 63 years. The most common age group was under 30 years. The medical history was dominated by cases of rheumatism (+90%) and abnormal sinus rhythm (40%). Other underlying conditions were detected in the patients: 3.33% of cases of stroke, 43.33% of recurrent angina, 40% of polyarthralgia, and 10% of endocarditis. Permanent atrial fibrillation was the most common form, accounting for 46.6% of cases. It was more frequent in subjects over 30 years of age. In the majority of younger subjects, the type of AF was not specified. One case of paroxysmal AF, 7% of persistent cases and 10% of newly diagnosed AF cases were detected. Other comorbidities were noted among 40% of patients, with the majority found in those over 30 years of age.

- **In silico identification and prediction of mutations in the *MTCYB* gene**

The resulting migration profile had an estimated size of approximately 1246 bp. The results, presented as an electropherogram, showed the sequence of bases making up the DNA fragment. During sequencing, three sequences were excluded due to their poor quality, as assessed by the sequencing service provider. The 27 sequences obtained were 1143 base pairs long. The various mutations were identified using the Mutation Surveyor software [11]. For the in-silico prediction of the impact of the mutations, several software programmers were used to determine the impact of an amino acid substitution on the structure and function of the human protein, the effect of these mutations on protein stability, and the probability that non-synonymous mutations are the cause of the disease. Non-synonymous *MTCYB* mutations were considered pathogenic if at least two of the prediction tools indicated this. Analysis of the subjects' chromatograms against the NC_012920 reference revealed a total of 46 variants, comprising 28 heterozygous variants (60.87%) and 18 homozygous variants (39.13%). The results also revealed 9 deletions and 2 duplications. Among these mutations, there are more synonymous mutations than non-synonymous mutations, spanning 47 different positions. One of the synonymous mutations, m.15301G>A (L185L), is reported as a disease-causing factor. This homoplastic transition mutation affects the stability of the protein. Analyses carried out using prediction software showed that these non-synonymous substitutions are predicted to be likely harmful, pathogenic, benign or deleterious. We identified 15 that were classified as pathogenic by all prediction software Table 3.

- **Structural impact tests using MISSENSE3D software**

Analysis of Table 4 showed that four of the pathogenic mutations in the *MTCYB* gene have secondary structural consequences for the cytochrome b protein. The m.15176C>CT and m.15197A>AG mutations result in structural clashes between the wild-type and variant structures. The m.15209T>C mutation results in a change from proline to cysteine in the wild-type position. With m.15209T>TA, the

Table 3. Presentation of mutations identified as pathogenic by all prediction software.

Position rCRS	P.AA	Statute	POLYPHEN-2	PANTHER-PSEP	SNAP	SNP-PHD	SNP&GO	I-MUTANT
m.14905G>C	M53I	*Homoplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.14954C>A	P261T	*Homoplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15176C>CT	L335Y	*Heteroplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15179C>CT	T336H	*Heteroplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15182T>TG	W337P	*Heteroplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15182T>TG	W337Q	*Heteroplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15182A>AG	W337N	*Heteroplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15197A>AG	P342Q	*Heteroplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15209T>C	P346I	*Homoplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15209T>TA	Y358K	*Heteroplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15248C>CT	F359L	*Heteroplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15248C>CT	F359S	*Heteroplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15272C>A	P367N	*Homoplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15272C>A	P367Q	*Homoplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease
m.15272C>A	P367T	*Homoplasmic*	Probably harmful	Probably harmful	Deleterious	Pathogenic	Pathogenic	Decrease

Table 4. Functional impacts of mutations on protein secondary structure by Missene3D.

Position rCRS	P.AA	Predictions
m.14905G>C	M53I	Benign
m.14954C>A	P261T	Benign
m.15176C>CT	L335Y	Harmful: This substitution triggers a collision alert. The local collision score for the wild-type structure is 13.68 and that for the variant structure is 39.24. The increase in the collision score is ≥18 compared to the wild type

Continued

m.15179C>CT	T336H	Benign
m.15182T>TG	W337P	Benign
m.15182T>TG	W337Q	Benign
m.15182A>AG	W337N	Benign
m.15197A>AG	P342Q	Harmful: This substitution triggers a collision warning. The local collision score for the wild-type structure is 18.33 and that for the variant structure is 37.65. The increase in the collision score is ≥18 compared to the wild type
m.15209T>C	P346I	Detrimental: In this variant, a proline at the cis position of the wild-type is replaced, which affects the three-dimensional structure of the protein
m.15209T>TA	Y358K	Detrimental: This substitution replaces a buried uncharged residue (TYR, RSA 6.3%) with a charged residue (LYS, RSA 3.4%)
m.15248C>CT	F359L	Benign
m.15248C>CT	F359S	Benign
m.15272C>A	P367N	Benign
m.15272C>A	P367Q	Benign
m.15272C>A	P367T	Benign

substitution replaces a buried uncharged residue with a charged residue. All other mutations have no effect on the secondary structure of the protein.

- **Polymorphism and genetic diversity**

The parameters of our population were defined according to age and sex, with age groups ranging from 0 to 30 years and another group aged 30 years and above. Genetic variability was defined according to several parameters (**Table 5**). These parameters were estimated using DnaSP version 5.10 software [12], as were genetic diversity (haplotype diversity (hd) and nucleotide diversity (Pi) indices) and mismatch distribution graphs. This corresponds to the graphical representation of the distribution of genetic distances between individuals in a population. However, nucleotide frequencies, the nature and rate of mutations, and intra- and inter-population genetic distances were determined using the MEGA version 7.0.14 program [13], using the Tamura Nei parameter model, which proved to be the most suitable. Using the same software, the Z-test for codon selection was performed using the Kumar model, based on the hypothesis that the *MTCYB* gene follows the neutral evolutionary model. **Table 5** shows the results for the polymorphism and genetic diversity of the *MTCYB* gene.

- **Genetic differentiation and structuring**

Estimating genetic differentiation between populations in phylogenetics requires the degree of genetic differentiation (FST) and genetic distance (D). The FST values between populations, the intra-specific neutrality test (Tajima's D, Fu's Fs), and the SSD and Rag indices were calculated using the Arlequin software version 3.5.1.3 [14]. These allow us to predict the genetic evolution of the studied

Table 5. Parameters of genetic variability of the *MTCYB* gene.

Parameters genetic		PT	Ages		Sex	
			<30 ans	≥30 ans	Men	Women
Sample size n		27	14	12	8	19
Number of sites N		1018	1019	1018	1021	1018
Number of no variable sites		995	1004	1007	1013	998
Number of variable sites		23	15	11	8	20
Non-informative variables		17	11	7	4	16
Number of mutations Eta		24	15	11	8	21
Informative variables		6	4	4	4	4
Nucleotide frequencies	A	27.8%	27.8%	27.8%	27.8%	27.8%
	T/U	25.1%	25.1%	25.1%	25.1%	25.2%
	C	35%	35%	35%	35%	35%
	G	12%	12%	12%	12.1%	12%
Mutation rate R		1,125	1,125	1,125	1,123	1,127
Average number of nucleotide differences k		3,234	3,385	2,818	2,964	3,386
Number of haplotypes h		18	11	9	7	16
Haplotypic Diversity Hd		0.954	0.956	0.955	0.964	0.965
Nucleotide diversity Pi		0.00319	0.00332	0.00277	0.00290	0.00333
synonym substitution		0.008 ± 0.003	0.007 ± 0.003	0.009 ± 0.003	0.010 ± 0.004	0.007 ± 0.003
No synonym substitution		0.002 ± 0.001	0.002 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	0.002 ± 0.001

population over time and test its goodness of fit. A p-value of <0.05 was considered statistically significant. With regard to the genetic structure of the populations, analysis of molecular variance (AMOVA) implemented in the Arlequin software (version 3.5.1.3) was used. AMOVA estimates indices of genetic structure using information on the allelic content of haplotypes (allele frequencies). Analysis of genetic structure revealed a low genetic distance within populations. For intergroup distances, the values are broadly similar across age groups and between subjects of different sexes. They also show higher molecular variance within populations (98.27; 100.87) than between populations (1.73; −0.87). Similarly, the differentiation coefficient (FST) is statistically insignificant for both groups, with p-values greater than 0.05. The results are presented in Table 6. Within each population group, demographic tests show negative and insignificant values for Tajima's D and negative but significant values for Fu's Fs (Table 6).

The Mistmatch test shows that the SSD and Rag values are positive but not significant. The non-significant p-values from the Z-test indicate that the *MTCYB* gene is undergoing neutral evolution under the assumption of neutrality and under the assumption of positive selection at the population level (Table 7). How-

ever, under the assumption of negative selection, the p-values are significant; this shows that the *MTCYB* gene is under negative selection in the population.

Table 6. Distance and genetic structuring.

Populations	Parameters	Nei's genetic distance		Percentage of variations		FST
		Intra ± SE	Inter ± SE			
Age	<30 ans	0.002 ± 0.001	0.002 ± 0.001	Enter pop.	1.73	0.01729 (0.23340)
	≥30 ans	0.001 ± 0.001		Enter ind.	98.27	
Sex	Men	0.001 ± 0.001	0.001 ± 0.001	Enter pop.	−0.87	0 (0.53711)
	Women	0.002 ± 0.001		Enter ind.	100.87	

Table 7. Genetic evolution.

Tests		Z-test		
Tajima	Fu's Fs	N	dS < dN	dS > dN
−1.6517 ± 0.02	−8.1345 ± 0	0.067	1	0.033
−1.1715 ± 0.09	−6.1396 ± 0	0.072	1	0.034
−0.9416 ± 0.19	−7.01941 ± 0			
−0.1898 ± 0.42	−3.2839 ± 0.01	0.076	1	0.046

The graphical representation of the genetic distances between pairs of individuals shows a unimodal curve. This is illustrated in the following Figure 1.

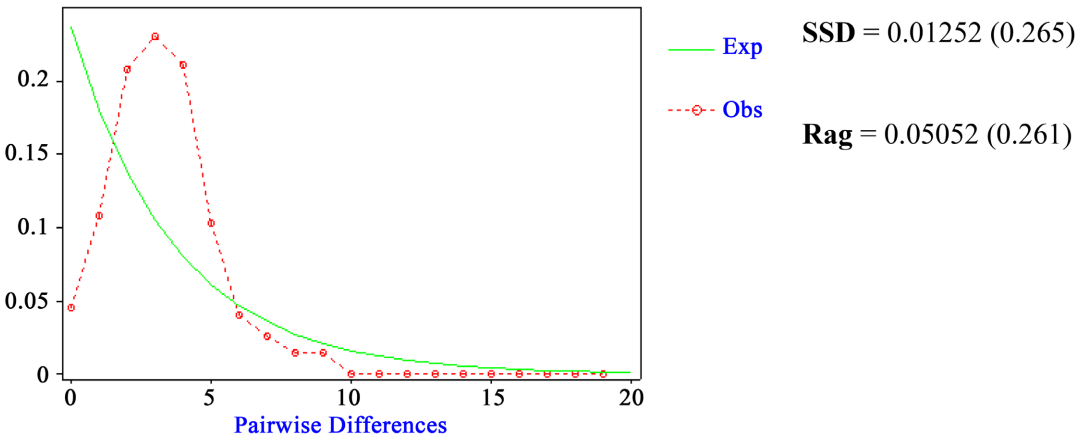

Figure 1. Graph of the Mistmatch distribution curve.

4. Discussion

- **Impact on protein stability and function**

The novel variants of the *MTCYB* gene identified in this case series, located at highly conserved residues and predicted to affect the structure and function of the cytochrome b protein, provide a preliminary basis for exploring the potential role of mitochondrial Complex III dysfunction in atrial fibrillation. Certain mutations

(including 15) are involved in all these different aspects simultaneously and appear to be novel. They could therefore serve as markers for the early clinical diagnosis of atrial fibrillation in the future. However, among the synonymous mutations, the silent mutation m.15301G>A is a common mutation in most of our patients (14.54%) and has been reported as a possible factor in high-altitude diseases [15]. Synonymous variants in protein-coding genes are generally considered to be silent and to have no effect on protein function. However, previous studies have shown that silent variants can significantly alter gene expression by affecting mRNA stability and folding [16] [17], and may also influence the rate of protein translation and post-translational modification [18]. This suggests a link to the disease. Similarly, there are nine deletions, most of which involve the pyrimidine T. These results strongly suggest that mtDNA mutations, particularly small deletions, may play a crucial role in atrial dysfunction in patients with AF. Several studies have confirmed that deletions in the mitochondrial DNA of atrial tissue lead to chronic atrial fibrillation, and that deletions detected in peripheral blood are associated with age-related atrial structural remodeling. It has been reported that deletion mutations in mitochondrial DNA (mtDNA) in human atrial tissue are associated with chronic atrial fibrillation, suggesting that mitochondrial genetic alterations may play a role in the pathophysiology of this condition [19]-[21].

- **Secondary structural implications of mutated genes**

Analysis of mutations in the secondary structure of proteins reveals clash effects between two substitutions. A clash occurs when two unbound atoms come within a distance less than the sum of their van der Waals radii, creating steric repulsion. The clash score, calculated by the MolProbity tool, measures the number of unfavorable steric contacts between unbound atoms that come within a distance less than the sum of their van der Waals radii, with an overlap threshold set at 0.4 Å [22]. A local clash scores greater than 30, combined with an increase in the clash score relative to the wild-type structure of ≥18, is considered a structural criterion for pathogenicity according to the thresholds defined by the Missense3D tool, which evaluates structural alerts within a 20 Å radius of the Cα of the variant residue [23]. The energy penalty of such a steric overlap can range from 0 to 10 kcal/mol depending on the types of atoms involved, with severe clashes being extremely rare in high-resolution structures [24]. The use of protein structural analysis in the clinical classification of variants has proven effective: in a series of 99 variants initially classified as variants of uncertain significance (VUS), 47 were reclassified as pathogenic or likely pathogenic based on structural evidence, notably using the ACMG's PM1 criterion [25].

Another variant analyzed involves the replacement of a proline residue that adopts a cis conformation in the wild-type protein. Proline is the only amino acid whose side chain forms a ring with the peptide nitrogen, giving it unique conformational properties. Unlike other amino acids, for which the cis conformation is extremely rare, proline can adopt this conformation with a frequency of 5% - 10%, thereby playing an irreplaceable structural role in the local architecture of pro-

teins. In cytochrome b, cis-prolines help maintain the geometry of the loops between the transmembrane helices and the precise positioning of the bL and bH hem groups, which are essential for electron transfer via the Q cycle. Structural analysis using MISSENSE3D confirmed that replacing a conserved cis-proline with a trans residue systematically results in multiple atomic collisions, indicating a major structural disruption [22]. Comparative structural data have shown that the substitution of a cis-proline with a non-proline residue in type I DNA polymerases systematically generates a structure incompatible with enzymatic function, highlighting the universal nature of this conformational constraint [22]. The loss of this cis-proline in *MTCYB* is therefore predicted to be highly detrimental to the assembly and function of Complex III.

The m.15209T>TA variant involves the replacement of a tyrosine (TYR) with a lysine (LYS) at a position deep within the hydrophobic core of cytochrome b. The solvent-accessible surface area (RSA) values are 6.3% for the wild-type TYR and 3.4% for the variant LYS, confirming that both residues are located within the hydrophobic core of the protein (RSA < 20%). This type of substitution combines three distinct destabilization mechanisms. Firstly, the introduction of a positively charged residue (LYS, +1) into a deeply buried hydrophobic environment demonstrates that it is thermodynamically highly unfavorable, as no water molecules or counterions are available to stabilize the charge. Systematic mutagenesis studies have shown that the introduction of charged residues at buried positions typically leads to misfolding of the protein and an inactive phenotype, with lysine being highly destabilizing in this context [26]. Secondly, large-scale analyses have confirmed that variations occurring in buried positions are more likely to be associated with diseases, and that this is particularly true for charged residues [27]. Thirdly, computational studies of 20 proteins associated with hereditary diseases have demonstrated that substitutions at known pathogenic mutation sites are significantly more likely to destabilize the native protein folding, supporting the idea that structural destabilization underlies the pathogenicity of a large set of missense mutations [28]. The loss of wild-type TYR also leads to the disruption of internal hydrogen bonds via its hydroxyl group, further exacerbating local destabilization.

- **Functional implications for complex III and FA**

The combination of these structural arguments (replacement of cis-proline, introduction of a charge at a buried position, and a high steric collision score) predicts severe dysfunction of cytochrome b and, by extension, of mitochondrial Complex III. Pathogenic mutations in *MTCYB* have been associated with severe deficiencies in the activity of Complexes I and III, due to the structural interdependence between these two complexes [29]. This dysfunction of Complex III leads to a reduction in electron transfer from ubiquinol to cytochrome c, a decrease in ATP production and an overproduction of ROS. In the cardiac context, these disturbances are directly arrhythmogenic: they alter calcium homeostasis, sarcoplasmic ionic currents and the membrane potential of atrial cardiomyocytes, creating an electrophysiological substrate conducive to AF [30].

The data set presented thus supports the classification of these two variants as likely pathogenic according to the ACMG/AMP (2015) criteria, based on criteria PS3 (functional/structural data), PM1 (location within a critical domain), and PM2 (absent or very rare in population databases) [31].

- **Genetic diversity of the population**

Variability in the *MTCYB* gene was studied according to sex and age in order to determine the parameters of genetic diversity in the nucleotide sequences of individuals with atrial fibrillation. The results revealed very little difference between the different sex groups and the different age groups. They show a high rate of polymorphism within our population, with a total of 24 Eta mutations. Analysis of the nucleotide composition revealed an unstable *MTCYB* gene with a higher A+T content compared to C+G, suggesting that this gene is prone to mutations in atrial fibrillation. Indeed, as DNA is double-stranded, the higher the A+T content, the more susceptible the gene is to mutations. *MTCYB* is also characterized by high haplotype diversity and low nucleotide diversity within each population. This indicates a signal of rapid population growth from an ancestral population of small effective size over a period long enough for haplotype diversity to be restored through mutation, but too short for significant differences between sequences to accumulate.

The nature of mutations suggests a higher rate of transitions than of transversions. Indeed, these transversions would cause a change in the conformation of the protein structure and would be more likely to induce missense mutations. Generally speaking, synonymous substitutions outnumber non-synonymous substitutions, suggesting that mutations in the *MTCYB* gene do not have a significant impact on protein stability. When the number of synonymous substitutions exceeds the number of non-synonymous substitutions, negative selection occurs, eliminating the non-synonymous substitutions. These results are confirmed by the Z-test, where the p-values are significant under the assumption of negative selection.

Similarly, genetic distance is higher among men than among women. For intergroup distances, the values are broadly similar across age groups and between individuals of different sexes. The populations are therefore genetically identical. The results also revealed a lack of genetic differentiation between men and women, as well as across age groups. The values for Tajima's D and fu's Fs tests indicate that the population is undergoing moderate expansion. These results are consistent with the Mistmatch curve, whose indices (SSD and Rag) show that there is no significant discrepancy between the observed and expected values. Large-scale epidemiological studies have established a direct link between mtDNA integrity and the risk of AF.

In a prospective analysis involving 19,709 participants from three cohorts (ARIC, MESA and CHS), a mtDNA copy number in the lowest quintile was associated with a 13% increased risk of developing incident AF, independent of traditional cardiovascular risk factors [32]. These findings suggest that the number of mtDNA copies could be a potential area of interest in the investigation of atrial fibrillation;

further studies are needed to determine its significance as a risk factor.

5. Conclusions

This case series presents, for the first time, a catalogue of novel variants in the *MTCYB* gene identified in Senegalese patients with atrial fibrillation at FANN Hospital in Dakar. The variants identified in this series are mostly previously unreported, as they do not appear in the major databases, notably MITOMAP, ClinVar and gnomAD. In silico structural and functional prediction analyses suggest that several of these variants could potentially affect the structure and stability of the cytochrome b protein, particularly those located at residues that are highly conserved across species. These observations highlight the importance of studying mitochondrial genetic variation in the context of atrial fibrillation.

Although the design of this study, which is limited to isolated cases, does not allow conclusions to be drawn regarding pathogenicity or disease risk, the novelty of the identified variants and their predicted structural impact provide a strong rationale for further research. Future studies involving larger cohorts, including appropriate control groups and employing functional validation methods, would be necessary to determine whether these *MTCYB* gene variants contribute to mitochondrial dysfunction and a predisposition to atrial fibrillation within the Senegalese population, and to open up new avenues for genetic counselling and personalized medicine within this population.

Statement

The study has been approved by the UCAD Ethics and Research Committee (reference number: Protocol 0274/2018/CER/UCAD).

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Joseph, P., Leong, D., McKee, M., Anand, S.S., Schwalm, J., Teo, K., *et al.* (2017) Reducing the Global Burden of Cardiovascular Disease, Part 1. *Circulation Research*, **121**, 677-694. https://doi.org/10.1161/circresaha.117.308903

[2] Brundel, B.J.J.M., Ai, X., Hills, M.T., Kuipers, M.F., Lip, G.Y.H. and de Groot, N.M.S. (2022) Atrial Fibrillation. *Nature Reviews Disease Primers*, **8**, Article No. 21. https://doi.org/10.1038/s41572-022-00347-9

[3] Schumacher, B., Pothof, J., Vijg, J. and Hoeijmakers, J.H.J. (2021) The Central Role of DNA Damage in the Ageing Process. *Nature*, **592**, 695-703. https://doi.org/10.1038/s41586-021-03307-7

[4] Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., *et al.* (2016) The Genecards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Current Protocols in Bioinformatics*, **54**, 1.30.1-1.30.33. https://doi.org/10.1002/cpbi.5

[5] (2026) Zymo Research International. https://zymoresearch.eu/

[6] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA Sequencing with Chain-Terminating Inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, **74**, 5463-5467. https://doi.org/10.1073/pnas.74.12.5463

[7] Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., *et al.* (1981) Sequence and Organization of the Human Mitochondrial Genome. *Nature*, **290**, 457-465. https://doi.org/10.1038/290457a0

[8] Song, Z., Laleve, A., Vallières, C., McGeehan, J.E., Lloyd, R.E. and Meunier, B. (2016) Human Mitochondrial Cytochrome *b* Variants Studied in Yeast: Not All Are Silent Polymorphisms. *Human Mutation*, **37**, 933-941. https://doi.org/10.1002/humu.23024

[9] Lopez-Oceja, A., Gamarra, D., Borragan, S., Jiménez-Moreno, S. and de Pancorbo, M.M. (2016) New cyt b Gene Universal Primer Set for Forensic Analysis. *Forensic Science International: Genetics*, **23**, 159-165. https://doi.org/10.1016/j.fsigen.2016.05.001

[10] Flores-Mireles, D., Camacho-Villasana, Y., Lutikurti, M., García-Guerrero, A.E., Lozano-Rosas, G., Chagoya, V., *et al.* (2023) The Cytochrome *b* Carboxyl Terminal Region Is Necessary for Mitochondrial Complex III Assembly. *Life Science Alliance*, **6**, e202201858. https://doi.org/10.26508/lsa.202201858

[11] Dong, C. and Yu, B. (2011) Mutation Surveyor: An *in Silico* Tool for Sequencing Analysis. In: Yu, B. and Hinchcliffe, M., Eds., *In Silico Tools for Gene Discovery*, Humana Press, 223-237. https://doi.org/10.1007/978-1-61779-176-5_14

[12] Librado, P. and Rozas, J. (2009) DnaSP V5: A Software for Comprehensive Analysis of DNA Polymorphism Data. *Bioinformatics*, **25**, 1451-1452. https://doi.org/10.1093/bioinformatics/btp187

[13] Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, **33**, 1870-1874. https://doi.org/10.1093/molbev/msw054

[14] Excoffier, L. and Lischer, H.E.L. (2010) Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

[15] Mallet, R.T., Burtscher, J., Pialoux, V., Pasha, Q., Ahmad, Y., Millet, G.P., *et al.* (2023) Molecular Mechanisms of High-Altitude Acclimatization. *International Journal of Molecular Sciences*, **24**, Article 1698. https://doi.org/10.3390/ijms24021698

[16] Komar, A.A. (2007) SNPs, Silent but Not Invisible. *Science*, **315**, 466-467. https://doi.org/10.1126/science.1138239

[17] Shabalina, S.A., Spiridonov, N.A. and Kashina, A. (2013) Sounds of Silence: Synonymous Nucleotides as a Key to Biological Regulation and Complexity. *Nucleic Acids Research*, **41**, 2073-2094. https://doi.org/10.1093/nar/gks1205

[18] Kimchi-Sarfaty, C., Oh, J.M., Kim, I., Sauna, Z.E., Calcagno, A.M., Ambudkar, S.V., *et al.* (2007) A "Silent" Polymorphism in the *MDR* 1 Gene Changes Substrate Specificity. *Science*, **315**, 525-528. https://doi.org/10.1126/science.1135308

[19] Wei, Y., Lu, C., Lee, H., Pang, C. and Ma, Y. (1998) Oxidative Damage and Mutation to Mitochondrial DNA and Age-Dependent Decline of Mitochondrial Respiratory Function. *Annals of the New York Academy of Sciences*, **854**, 155-170. https://doi.org/10.1111/j.1749-6632.1998.tb09899.x

[20] Michikawa, Y., Mazzucchelli, F., Bresolin, N., Scarlato, G. and Attardi, G. (1999) Aging-Dependent Large Accumulation of Point Mutations in the Human mtDNA Con-

trol Region for Replication. *Science*, **286**, 774-779. https://doi.org/10.1126/science.286.5440.774

[21] Park, H., Ahn, Y., Jeong, M., Cho, J., Park, J., Kang, J., *et al.* (2007) Chronic Atrial Fibrillation Associated with Somatic Mitochondrial DNA Mutations in Human Atrial Tissue. *Journal of Clinical Pathology*, **60**, 948-950. https://doi.org/10.1136/jcp.2007.047134

[22] Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., *et al.* (2007) MolProbity: All-Atom Contacts and Structure Validation for Proteins and Nucleic Acids. *Nucleic Acids Research*, **35**, W375-W383. https://doi.org/10.1093/nar/gkm216

[23] Ittisoponpisan, S., Islam, S.A., Khanna, T., Alhuzimi, E., David, A. and Sternberg, M.J.E. (2019) Can Predicted Protein 3D Structures Provide Reliable Insights into Whether Missense Variants Are Disease Associated? *Journal of Molecular Biology*, **431**, 2197-2212. https://doi.org/10.1016/j.jmb.2019.04.009

[24] Ramachandran, S., Kota, P., Ding, F. and Dokholyan, N.V. (2011) Automated Minimization of Steric Clashes in Protein Structures. *Proteins*: *Structure*, *Function*, *and Bioinformatics*, **79**, 261-270. https://doi.org/10.1002/prot.22879

[25] Caswell, R.C., Gunning, A.C., Owens, M.M., Ellard, S. and Wright, C.F. (2022) Assessing the Clinical Utility of Protein Structural Analysis in Genomic Variant Classification: Experiences from a Diagnostic Laboratory. *Genome Medicine*, **14**, Article No. 77. https://doi.org/10.1186/s13073-022-01082-2

[26] Bajaj, K., Chakrabarti, P. and Varadarajan, R. (2005) Mutagenesis-based Definitions and Probes of Residue Burial in Proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 16221-16226. https://doi.org/10.1073/pnas.0505089102

[27] Savojardo, C., Manfredi, M., Martelli, P.L. and Casadio, R. (2021) Solvent Accessibility of Residues Undergoing Pathogenic Variations in Humans: From Protein Structures to Protein Sequences. *Frontiers in Molecular Biosciences*, **7**, Article 626363. https://doi.org/10.3389/fmolb.2020.626363

[28] Redler, R.L., Das, J., Diaz, J.R. and Dokholyan, N.V. (2016) Protein Destabilization as a Common Factor in Diverse Inherited Disorders. *Journal of Molecular Evolution*, **82**, 11-16. https://doi.org/10.1007/s00239-015-9717-5

[29] Blakely, E.L., Mitchell, A.L., Fisher, N., Meunier, B., Nijtmans, L.G., Schaefer, A.M., *et al.* (2005) A Mitochondrial Cytochrome *b* Mutation Causing Severe Respiratory Chain Enzyme Deficiency in Humans and Yeast. *The FEBS Journal*, **272**, 3583-3592. https://doi.org/10.1111/j.1742-4658.2005.04779.x

[30] Mauriello, A., Correra, A., Molinari, R., Del Vecchio, G.E., Tessitore, V., D'Andrea, A., *et al.* (2024) Mitochondrial Dysfunction in Atrial Fibrillation: The Need for a Strong Pharmacological Approach. *Biomedicines*, **12**, Article 2720. https://doi.org/10.3390/biomedicines12122720

[31] Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., *et al.* (2015) 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. *Genetics in Medicine*, **17**, 405-424. https://doi.org/10.1038/gim.2015.30

[32] Zhao, D., Bartz, T.M., Sotoodehnia, N., Post, W.S., Heckbert, S.R., Alonso, A., *et al.* (2020) Mitochondrial DNA Copy Number and Incident Atrial Fibrillation. *BMC Medicine*, **18**, Article No. 246. https://doi.org/10.1186/s12916-020-01715-6