

Research Article

Genetic Variability of *Pitx2* Exon 4 in Senegalese Patients with Rheumatic Mitral Valve Disease with and Without Atrial Fibrillation: A Prospective Study with in Silico Analysis

Diagne Papa Amath^{1,*} , Mbaye Fatimata² , Diop Momar Sokhna¹ ,
Ba Papa Ousmane¹ , Diop Moussa Seck¹ , Faye Cheikh Abdou Khadre¹ ,
Ba Papa Salmane¹ , Ciss Amadou Gabriel¹ , Sembene Mbacke² 

¹Department of Thoracic and Cardiovascular Surgery, Fann Hospital, Dakar, Senegal

²Faculty of Science and Technology, Cheikh Anta Diop University, Dakar, Senegal

Abstract

Introduction: Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia and is associated with significant morbidity and mortality. In sub-Saharan Africa, AF frequently occurs at a younger age, most often in the context of rheumatic mitral valve disease (RMVD), and is characterized by advanced atrial remodeling. Genetic studies have identified the Paired-like Homeodomain Transcription Factor 2 (*PITX2*) gene at the 4q25 locus as a major susceptibility factor for AF. **Objective:** to investigate genetic variability within exon 4 of the *PITX2* gene in African patients with RMVD and to explore its potential role in AF susceptibility. **Materials and Methods:** this prospective observational study was conducted at the Cuomo Cardiac Surgery Center, Aristide Le Dantec University Hospital (Dakar, Senegal). Forty-five patients undergoing surgery for RMVD were included, along with 15 control subjects without documented AF. Intraoperative left atrial endocardial biopsies were obtained in operated patients. Genomic Deoxyribonucleic Acid (DNA) extracted from atrial tissue and control blood samples was used for targeted amplification and Sanger sequencing of exon 4 of *PITX2*. Identified variants were annotated using public databases and evaluated with in silico prediction tools. Variants were classified according to the American College of Medical Genetics and Genomics (ACMG) guidelines. **Results:** the study population was young (mean age 31 years) and predominantly female. AF was present in 56% of patients and was associated with marked left atrial dilatation and preserved left ventricular systolic function. Sequencing identified four rare missense variants (c.59C>T, c.97A>G, c.121C>G, and c.149G>T) located within the conserved homeobox domain of *PITX2*. All variants were heterozygous and showed extremely low allele frequencies in population databases. In silico predictions were heterogeneous, and all variants were classified as variants of uncertain significance. Genetic diversity analyses revealed slightly higher haplotype and nucleotide diversity in AF patients, without significant genetic differentiation between groups. **Conclusion:** rare variants in exon 4 of *PITX2* were identified in African patients with RMVD. Although their functional significance remains uncertain, these findings support a potential modulatory role of *PITX2* in susceptibility to atrial fibrillation and highlight the importance of genetic studies in underrepresented African populations.

*Correspondence: Diagne Papa Amath (Papaamath.diagne@ucad.edu.sn)

Received: 10 February 2026; **Accepted:** 16 March 2026; **Published:** 28 March 2026



Copyright: © The Author(s), 2026. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Keywords

Atrial Fibrillation, Rheumatic Mitral Valve Disease, *PITX2*, Chromosome 4q25, Africa, Cardiovascular Genetics

1. Introduction

Atrial fibrillation (AF) is the most common sustained cardiac arrhythmia worldwide, with an estimated prevalence of 1–2% in adults and a rapidly increasing global burden. AF is associated with a substantial risk of stroke, heart failure, and all-cause mortality. In sub-Saharan Africa, AF frequently occurs at a younger age in the context of rheumatic mitral valve disease, which remains highly prevalent, and is often characterized by advanced atrial remodeling and persistent or permanent forms requiring surgical management. AF is a multifactorial arrhythmia resulting from the interaction between an abnormal atrial substrate, triggering foci, and modulating factors. Beyond structural and electrical remodeling, recent advances in molecular genetics have demonstrated an inherited component of AF susceptibility. Genome-wide association studies have consistently identified the chromosome 4q25 locus, near the *PITX2* gene, as the strongest genetic signal associated with AF across populations. The *PITX2* gene encodes a homeobox transcription factor essential for left atrial development and electrophysiological stability through regulation of ion channel expression. Reduced *PITX2* expression has been linked to atrial electrical remodeling and increased vulnerability to AF. However, data on *PITX2* genetic variation in African populations remain limited. The present study aims to investigate genetic variability of exon 4 of *PITX2* in African patients with rheumatic mitral valve disease, with or without atrial fibrillation, and to explore its potential role in AF susceptibility in this underrepresented population.

2. Materials and Methods

This prospective, observational, and analytical study was conducted over a one-year period, from August 2023 to July 2024, at the Department of Thoracic and Cardiovascular Surgery, Cuomo Cardiac Surgery Center, Aristide Le Dantec University Hospital (Dakar, Senegal). Patients with acquired rheumatic mitral valve disease scheduled for mitral valve replacement surgery were consecutively enrolled, regardless of the presence or absence of documented atrial fibrillation (AF). The study population comprised 44 operated patients and a control group of 15 subjects without a documented history of AF.

Clinical data collected included age, sex, medical history, and functional symptoms. Rhythm status was assessed using standard electrocardiography and classified as sinus rhythm or atrial fibrillation (paroxysmal, persistent, permanent, or post-operative). Transthoracic echocardiography was performed to

evaluate mitral valve pathology, left atrial size, and right and left ventricular function.

After written informed consent was obtained, atrial tissue samples were collected intraoperatively. All surgical procedures were performed under cardiopulmonary bypass. Following median sternotomy, cardioplegic arrest, and left atriotomy, an approximately 1 cm² endocardial biopsy was obtained from the junction between the left atrium and the right superior pulmonary vein. Samples were processed according to downstream analyses: fixation for histological and ultrastructural studies, or preservation for genetic analysis.

Genetic analysis focused on the *PITX2* gene and was performed on atrial tissue samples using standardized protocols for DNA extraction, amplification, and molecular analysis. Data entry and descriptive analyses were performed using Microsoft Excel, and results were analyzed according to patients' rhythm status.

Genetic analysis focused on the extraction, amplification, and targeted sequencing of exon 4 of the *PITX2* gene from atrial tissue samples and control blood samples. Genomic DNA was extracted from approximately 25 mg of left atrial endocardial tissue collected at the left atrium–right superior pulmonary vein junction using a commercial silica membrane–based kit (Zymo Research), according to the manufacturer's standard protocol. Control DNA was extracted from peripheral whole blood using the QIAamp DNA Blood Mini Kit (Qiagen).

DNA quality was assessed by agarose gel electrophoresis. Exon 4 of *PITX2* (approximately 300 bp), located on chromosome 4q25, was amplified by polymerase chain reaction (PCR) using specific primers designed from the reference sequence NM_000325.5. PCR products were visualized on 2% agarose gels, purified, and subjected to Sanger sequencing (ABI Prism 3130xl, Applied Biosystems).

Sequence chromatograms were analyzed using ChromasPro and aligned to the reference sequence NM_000325.6 with BioEdit software to identify nucleotide variants. Variant annotation was performed using public databases (dbSNP, ClinVar, and 1000 Genomes), and predicted functional impact was assessed using in silico tools (PolyPhen-2, Sorting Intolerant From Tolerant (SIFT), and MutationTaster).

Statistical analyses were conducted using IBM SPSS Statistics version 25.0 and GraphPad Prism version 8.0. Continuous variables were compared using Student's *t* test or the Mann–Whitney *U* test, as appropriate, while categorical vari-

ables were analyzed using the χ^2 test or Fisher’s exact test. Associations between *PITX2* variants and atrial fibrillation were evaluated by logistic regression analysis, with statistical significance set at $p < 0.05$. Statistical results are presented with 95% confidence intervals (95% CI) where applicable in order to provide an estimate of the precision of the calculated parameters.

3. Results

3.1. Epidemiological, Clinical, and Paraclinical Characteristics

The study population included 45 patients undergoing surgery for rheumatic mitral valve disease. The majority were female (36/45; female-to-male ratio 4:1), with a mean age of 31 years (range, 12–63), including 13 patients younger than 18 years. All patients originated from West Africa, predominantly Senegal (86%).

A clinical history suggestive of acute rheumatic fever was frequent, with recurrent pharyngitis reported in 42% of cases and polyarthralgia in 44%. Mitral valve involvement consisted of isolated mitral regurgitation in 49% of patients, isolated mitral stenosis in 17%, and mixed mitral disease in 34%.

Atrial fibrillation (AF) was present in 25 patients (56%), predominantly in the permanent form (35%). Other AF patterns included newly diagnosed AF (13%), persistent AF (6%), and paroxysmal AF (2%). AF occurred preferentially in older patients within the cohort. The mean cardiothoracic index was 0.61, and cardiomegaly was observed in 73% of cases.

Transthoracic echocardiography demonstrated marked left

atrial dilatation, with a mean anteroposterior diameter of 52 mm (range, 29–78 mm) and a mean left atrial surface area of 40 cm²(range, 17–69 cm²), observed in both patients with AF and those in sinus rhythm. Left ventricular systolic function was globally preserved, with a mean left ventricular ejection fraction of 65%, suggesting a predominantly atrial mechanism for AF. Mean tricuspid annular plane systolic excursion (TAPSE) was 22 mm.

Mitral valve disease was associated with tricuspid regurgitation in 69% of patients, with a mean systolic pulmonary artery pressure of 52 mmHg, and with aortic regurgitation in 28%. No patient exhibited biological evidence of active inflammation. Infectious markers were negative, and antistreptolysin O titers were positive in only 6.7% of cases, without clinical correlation.

3.2. Genetic Findings

DNA extracted from cardiac tissue samples and control blood samples showed concentrations ranging from 80 to 120 ng/ μ L, with A260/A280 ratios between 1.8 and 2.0, indicating satisfactory purity. PCR amplification of exon 4 of the *PITX2* gene consistently yielded a single amplicon of approximately 300 bp, confirming reaction specificity. In silico primer validation using BLASTn demonstrated exclusive alignment to exon 4 of *PITX2* (reference transcript NM_000325.6, locus 4q25).

Targeted Sanger sequencing identified four rare missense variants: c.59C>T (p.Ala20Val), c.97A>G (p.Lys33Glu), c.121C>G (p.Pro41Ala), and c.149G>T (p.Ser50Ile), all located within the Homeobox domain critical for *PITX2* transcriptional activity (Table 1).

Table 1. In Silico Pathogenicity Predictions of Variants Identified in Exon 4 of the *PITX2* Gene.

Mutation	p.AA	dbSNP	CADD	FATHMM	PolyPhen-2	MutationTaster	SIFT
c.59C>T	p.Ala20Val	rs150684621	24.7	Damaging	Benign	Benign	Tolerated
c.149G>T	p.Ser50Ile	—	24.2	Damaging	Benign	Benign	Deleterious
c.97A>G	p.Lys33Glu	rs1578464704	23.4	Damaging	Probably damaging	Benign	Tolerated
c.121C>G	p.Pro41Ala	rs758414076	17.35	Damaging	Benign	Benign	Tolerated

Software Used: CADD, FATHMM, PolyPhen-2, MutationTaster, SIFT.

Each variant was detected in a heterozygous state, with no homozygous carriers observed. Allele frequencies in public databases (gnomAD, dbSNP, ClinVar) were extremely low, particularly for c.59C>T, c.97A>G, and c.121C>G ($\leq 2 \times 10^{-5}$), confirming their rarity.

In silico pathogenicity prediction revealed heterogeneous results. High Combined Annotation Dependent Depletion (CADD) scores (≥ 23) were observed for variants c.59C>T, c.97A>G, and c.149G>T, suggesting a potential functional

impact, whereas only c.149G>T was predicted to be deleterious by SIFT. According to the American College of Medical Genetics and Genomics (ACMG) criteria, all identified variants were classified as variants of uncertain significance (VUS), due to insufficient functional evidence.

In silico assessment of the functional impact of *PITX2* exon 4 variants revealed variable physicochemical effects. The p.Ala20Val variant (c.59C>T) represents a moderate amino acid substitution (Grantham distance = 64) and was predicted

to be tolerated by SIFT and benign by PolyPhen-2. The p.Lys33Glu variant (c.97A>G) involves a charge change from a basic lysine to an acidic glutamate (Grantham distance = 56) and was classified as probably damaging by PolyPhen-2. The p.Pro41Ala variant (c.121C>G), corresponding to the substitution of a structurally rigid proline by a more flexible alanine (Grantham distance = 27), was globally predicted as benign. In contrast, the p.Ser50Ile variant (c.149G>T) induces a major change in polarity and side-chain size (Grantham distance = 142), affects a highly conserved position, and was predicted to be deleterious by SIFT.

Analysis of genetic diversity parameters for *PITX2* exon 4

(Table 2) demonstrated a globally low level of polymorphism in both study groups. However, haplotype diversity was higher in patients with atrial fibrillation (Haplotype Diversity $H_d = 0.847$) compared with controls ($H_d = 0.657$). Similarly, nucleotide diversity was slightly increased in the atrial fibrillation group (Nucleotide Diversity $\pi = 0.008$) relative to controls ($\pi = 0.005$), indicating greater intragroup heterogeneity among affected patients.

Analysis of molecular variance (AMOVA) based on *PITX2* exon 4 sequences did not reveal significant genetic differentiation between control subjects and patients with atrial fibrillation.

Table 2. Genetic Diversity Parameters of *PITX2* Exon 4 in Control and Atrial Fibrillation Groups.

Genetic Parameters	Témoins	Fibrillation atriale	Software Used
Sample size (n)	15	44	DnaSP / MEGA
Total sites analyzed (N)	359	359	DnaSP
Polymorphic sites	9	15	DnaSP
Haplotypes (h)	7	16	DnaSP / MEGA
Haplotype diversity (H_d)	0,657	0,847	DnaSP
Nucleotide diversity (π)	0,005	0,008	DnaSP / Arlequin
Mean nucleotide differences (k)	1,924	2,929	Arlequin
Transitions / Transversions (%)	57,46 / 42,52	53,34 / 46,68	MEGA
Mutation rate (R)	1,67	0,68	DnaSP

Software Used: BioEdit 7.2.5, DnaSP v6.12.03, MEGA 11.0.13, Arlequin v3.5.

Nearly all genetic variability was observed within groups (99.96%), whereas the proportion attributable to intergroup differences was negligible (0.04%), with a very low and non-significant fixation index (F_{ST} ; $p > 0.05$) (Table 3).

Table 3. Analysis of Molecular Variance (AMOVA).

Source of variation	Degrees of freedom	Sum of squares	Variance component	% of variation
Among populations	1	1,353	0,00052 (V_a)	0,04%
Within populations	57	76,444	1,34112 (V_b)	99,96%
Total	58	77,797	1,34164	—

Arlequin v3.5.

Fixation index (F_{ST}) = 0.00038; p-value (1,023 permutations) = 0.383 ± 0.017

Nei's genetic distances were consistent with this homogeneity, as intergroup distance (0.0083) was comparable to intragroup distances observed in both controls and atrial fibrillation patients (0.01) (Table 4).

Table 4. *Nei's genetic distances.*

Group	Within-group	Between-group
Control group	0.01	0.0083
Atrial fibrillation group	0.01	

Bioedit, version 7.2.5, DNASP, version 6.12.03 et Mega 11, version 11.0.13

Neutrality tests showed negative values of Tajima's D and Fu's Fs in both groups, indicating an excess of rare variants. These deviations were not statistically significant in controls,

whereas Fu's Fs was significantly negative in patients with atrial fibrillation (-5.88 ; $p = 0.010$), suggesting a deviation from the neutral model (Table 5).

Table 5. *Neutrality Tests (Tajima's D and Fu's FS).*

Parameters	Control group	Fibrillation	Mean	Standard deviation
Sample size (n)	15	44	29,5	20,5
Number of polymorphic sites (S)	9	15	12,0	4,2
Mean number of nucleotide differences (k)	1,92381	2,92918	2,42649	0,7109
Tajima's D	-1,1435	-0,4718	-0,8077	0,4749
p-value (Tajima's D)	0,112	0,364	0,238	0,178
Number of observed haplotypes (h)	7	16	11,5	6,36
Fu's FS	-2,00292	-5,88345	-3,94318	2,74
p-value (Fu's FS)	0,069	0,010	0,0395	0,0417

Arlequin v3.5.

Demogenetic analyses of *PITX2* exon 4 were primarily qualitative, as formal mismatch distribution analyses could not be performed due to the lack of complete individual sequence data. Nevertheless, all identified variants displayed extremely low allele frequencies in reference databases (all $< 4 \times 10^{-4}$ in gnomAD v4.1), and no homozygous carriers were reported, supporting the presence of strong evolutionary constraint. Population-level distribution showed marked geographic heterogeneity, with variant c.121C>G observed almost exclusively in African populations, whereas c.59C>T and c.97A>G were restricted to non-Finnish European populations. This pattern is compatible with localized genetic drift and/or recent selective pressures acting on rare variants within the Homeobox domain of *PITX2*.

4. Discussion

4.1. Epidemiological, Clinical, and Paraclinical Findings

The epidemiological characteristics of this cohort are con-

sistent with the classical profile of rheumatic mitral valve disease complicated by atrial fibrillation (AF) observed in low- and middle-income countries. The marked female predominance (female-to-male ratio = 0.25) is in line with recent reports indicating that women account for 60–75% of patients with rheumatic heart disease, particularly in sub-Saharan Africa [1]. This female predominance has been attributed to socioeconomic factors, delayed access to healthcare, and prolonged progression of valvular lesions among women [2].

The young mean age of the study population (31 years), with a substantial proportion of children and adolescents, mirrors findings from contemporary African series and contrasts sharply with data from high-income countries, where AF predominantly affects individuals older than 65 years and is commonly associated with degenerative or hypertensive heart disease [3, 4]. According to Global Burden of Disease estimates, rheumatic heart disease remains highly prevalent in low-income regions, particularly in sub-Saharan Africa, contributing to a high burden of valvular AF and an earlier onset of arrhythmia in these populations [5]. A history suggestive of acute rheumatic fever, including recurrent pharyngitis and poly-

arthralgia, was frequently reported, consistent with observations in chronic rheumatic heart disease cohorts, even in the absence of recent biological confirmation [6]. Clinically, the predominance of isolated mitral regurgitation or mitral regurgitation associated with mitral stenosis aligns with recent African studies, in which mitral regurgitation represents the most common presentation of rheumatic heart disease [7]. Mixed mitral lesions, observed in approximately one-third of patients, are recognized as a major contributor to combined pressure and volume overload of the left atrium, thereby promoting AF onset and persistence [5].

Radiographic findings confirmed chronic hemodynamic burden, with an elevated mean cardiothoracic index and cardiomegaly present in nearly three-quarters of patients. Similar rates have been reported in African cohorts of rheumatic mitral valve disease complicated by AF [8]. The more pronounced cardiothoracic enlargement observed in patients with AF suggests a close relationship between global cardiac dilation and arrhythmia chronicity, as previously described [9].

The high prevalence of AF in this series (56%), predominantly in its permanent form, is consistent with contemporary data showing that AF complicates 40–70% of hospitalized cases of rheumatic mitral valve disease in Africa [10]. The preferential occurrence of AF among older patients within the cohort likely reflects cumulative disease duration rather than age-related degenerative mechanisms, as highlighted in recent epidemiological studies [11].

Echocardiographic assessment revealed marked left atrial dilation, a parameter now recognized as the principal structural determinant of valvular AF [12]. Recent studies have shown that left atrial enlargement often precedes AF onset and constitutes a robust marker of advanced atrial remodeling [13]. The preservation of left ventricular systolic function observed in this cohort supports a predominantly atrial mechanism of AF, secondary to fibrosis and electrical heterogeneity induced by chronic hemodynamic overload [14]. The frequent association with tricuspid regurgitation and pulmonary hypertension further exacerbates biatrial remodeling and contributes to arrhythmia maintenance, as reported in recent series [15].

Finally, the low rate of antistreptolysin O positivity is consistent with evidence indicating that these markers lose diagnostic value long after the acute phase, particularly in adults, and do not exclude a remote rheumatic etiology [16]. The absence of biological inflammatory syndrome and negative infectious markers support the concept that AF in this setting is primarily driven by chronic structural remodeling rather than ongoing inflammatory activity, as previously demonstrated [17].

4.2. Genetic Findings

4.2.1. DNA Quality and PCR Amplification

DNA quality represents a critical determinant of the reliability of molecular analyses based on polymerase chain reac-

tion (PCR) and sequencing. In the present study, DNA concentrations ranged from 80 to 120 ng/ μ L, with A260/A280 absorbance ratios between 1.8 and 2.0, indicating adequate purity of the extracted genomic DNA. These values fall within the accepted standards for molecular biology applications and are compatible with efficient PCR amplification while minimizing the risk of enzymatic inhibition. Similar DNA quality parameters have been reported in molecular genetic studies investigating atrial fibrillation, including large-scale genomic analyses by Roselli et al [18], in which high-quality DNA was considered essential for reliable variant detection. Agarose gel electrophoresis revealed a single band of approximately 300 bp, confirming the specificity of the PCR amplification protocol. The absence of secondary amplicons or non-specific products suggests that the annealing and extension conditions were optimal. Such specificity is crucial to avoid amplification of paralogous sequences or pseudogenes that could compromise sequencing accuracy. In silico validation of the primers using BLASTn confirmed their unique alignment with exon 4 of the *PITX2* gene (transcript NM_000325.6) located at the 4q25 locus. The absence of significant hybridization with other genomic regions further supports the specificity of the amplification strategy. This validation step is widely recommended in molecular genetics protocols to minimize co-amplification of homologous sequences and ensure reliable variant identification [19]. These methodological findings are consistent with contemporary studies investigating the genetic determinants of Atrial fibrillation, particularly those highlighting the importance of precise targeting of the *PITX2* locus [20].

4.2.2. Identification of *PITX2* Variants

Sequencing of exon 4 of *PITX2* identified four rare missense variants: c.59C>T (p.Ala20Val), c.97A>G (p.Lys33Glu), c.121C>G (p.Pro41Ala), and c.149G>T (p.Ser50Ile). All substitutions are located within the highly conserved Homeobox domain, which plays a central role in DNA binding and transcriptional regulation. The structural integrity of this domain is essential for *PITX2* function in cardiac development and for maintaining electrophysiological stability of the left atrium [14, 20]. All variants were detected in the heterozygous state and occurred independently among individuals, with no homozygous carriers identified. This distribution suggests that these variants are rare within the studied population and may reflect purifying selection acting on genes involved in cardiac morphogenesis and transcriptional regulation [18]. Comparison with international databases (gnomAD, dbSNP, and ClinVar) confirmed the extremely low global allele frequencies of these variants. Notably, the c.149G>T variant is absent from ClinVar and dbSNP but shows a slightly higher frequency in African populations according to gnomAD (3.6×10^{-4}), suggesting a potential population-specific variant or regional founder effect. This observation aligns with previous reports highlighting the un-

derrepresentation of African genomic diversity in global databases [21].

4.2.3. Evolutionary Conservation and Functional Context

Evolutionary conservation analysis further supports the potential biological relevance of these variants. Positions c.59 and c.149 show high phyloP conservation scores (3.95 and 4.06), indicating strong evolutionary constraint, whereas c.97 and c.121 display lower conservation scores, suggesting greater tolerance to sequence variation. All substitutions map to a functionally critical region of *PITX2* on chromosome 4q25 (GRCh38: g.110,632,468–g.110,632,378). Experimental studies have demonstrated that alterations in *PITX2* expression or activity can disrupt atrial electrophysiological homeostasis by modulating ion channel expression and atrial conduction pathways [20]. Although none of these variants is currently classified as pathogenic in clinical databases, their rarity, localization within a key functional domain, and evolutionary conservation suggest potential biological relevance. These findings support the hypothesis that rare variants in regulatory genes such as *PITX2* may contribute to susceptibility to atrial fibrillation through complex interactions with structural and environmental [14, 19].

4.2.4. In Silico Pathogenicity Prediction

Pathogenicity assessment using several in silico prediction algorithms (CADD, SIFT, PolyPhen-2, MutationTaster, and FATHMM) revealed heterogeneous prediction profiles. High CADD scores (≥ 23) for variants c.59C>T, c.97A>G, and c.149G>T suggest potential deleterious effects; however, predictions varied across algorithms. For example, c.149G>T was classified as deleterious by SIFT but benign by PolyPhen-2 and MutationTaster, whereas c.97A>G was predicted as probably damaging by PolyPhen-2 but tolerated by SIFT. Such discrepancies are common in the interpretation of rare variants because different algorithms integrate evolutionary conservation, physicochemical changes, and structural context using distinct computational frameworks. According to the guidelines of the American College of Medical Genetics and Genomics (ACMG), all identified variants remain classified as variants of uncertain significance (VUS) due to insufficient functional evidence. Physicochemical analysis suggests variable functional effects among the substitutions. The p.Ala20Val and p.Pro41Ala variants represent relatively conservative amino acid changes, whereas p.Lys33Glu involves a charge shift that may affect electrostatic interactions. The p.Ser50Ile substitution introduces a larger change in polarity and side-chain volume, potentially affecting the structural stability of the Homeobox domain or its interactions with DNA and protein partners.

4.2.5. Genetic Diversity and Population Structure

Genetic diversity analysis revealed an overall low level of

polymorphism within exon 4, reflecting strong evolutionary conservation of this transcriptionally active region [14]. However, slightly higher diversity was observed among AF patients, with greater haplotype diversity ($H_d = 0.847$ vs 0.657) and nucleotide diversity ($\pi = 0.008$ vs 0.005). These findings suggest increased intragroup genetic heterogeneity among affected individuals, consistent with studies reporting an accumulation of rare variants in genes involved in cardiac development and atrial conduction [18, 22]. Genetic differentiation analyses did not reveal significant molecular structuring between controls and AF patients. Analysis of molecular variance (AMOVA) showed that 99.96% of genetic variability occurred within populations, whereas only 0.04% was attributable to differences between groups. The very low and non-significant fixation index ($F_{ST} = 0.00038$; $p > 0.05$) further confirms the absence of notable genetic differentiation. Neutrality tests nevertheless indicated an excess of rare alleles. While Tajima's D and Fu's F_s values were not significant in controls, the strongly negative and statistically significant Fu's F_s value observed in AF patients (-5.88 ; $p = 0.010$) may reflect recent demographic expansion or purifying selection favoring the persistence of rare haplotypes.

4.2.6. Demogenetic Interpretation

Population database comparisons revealed moderate inter-population differences. Variants c.121C>G and c.149G>T appear slightly more frequent in African populations, whereas c.59C>T and c.97A>G remain extremely rare and mainly reported in non-Finnish European populations. These observations suggest geographically differentiated distributions of rare variants and highlight the importance of expanding genomic studies to include African cohorts [21]. Overall, the evolutionary patterns observed in exon 4 of *PITX2* suggest a non-strictly neutral model resulting from the combined effects of genetic drift, functional constraints, and selective pressures. Such evolutionary dynamics may contribute to shaping the distribution of rare variants influencing individual susceptibility to atrial fibrillation. More comprehensive analyses based on individual sequence data and encompassing the entire 4q25 locus will be necessary to better understand the evolutionary mechanisms influencing *PITX2* genetic variability.

The distribution of rare variants observed in this study may also reflect the particular genetic architecture of African populations. Human populations in Africa harbor the highest levels of genetic diversity due to their longer evolutionary history and larger effective population sizes. As a consequence, African populations tend to exhibit a greater number of rare variants compared with non-African populations. In addition, the demographic history of African populations, including ancient population expansions and local population structure, may contribute to the persistence and diversification of rare genetic variants. Another important factor is the persistent underrepresentation of African populations in global genomic databases, which can lead to an apparent enrichment of population-specific variants when studies are conducted in African cohorts.

These factors highlight the importance of including diverse populations in genetic studies in order to better characterize the full spectrum of human genetic variation and its potential role in complex diseases such as atrial fibrillation.

Taken together, these findings indicate that exon 4 of the *PITX2* gene is highly conserved, reflecting the strong functional constraints associated with this key cardiac transcription factor. The identification of rare missense variants within the Homeobox domain suggests that subtle genetic alterations may contribute to the molecular mechanisms underlying Atrial fibrillation. Although none of the variants are currently classified as pathogenic, their rarity, evolutionary conservation, and localization within a critical functional domain suggest a potential influence on *PITX2* transcriptional activity or downstream regulatory pathways involved in atrial electrophysiology [14, 20]. The excess of rare alleles observed among patients further supports the hypothesis that rare variants in regulatory genes may contribute to individual susceptibility to atrial fibrillation through polygenic mechanisms [18, 22].

However, the relatively small sample size, the restriction of the analysis to exon 4, and the absence of functional validation represent important limitations. Larger studies integrating extended genomic analyses and experimental validation will be necessary to clarify the role of *PITX2* variants in atrial fibrillation pathophysiology. Another limitation concerns the composition of the control group. Controls consisted of individuals without rheumatic mitral valve disease or atrial fibrillation, introducing structural heterogeneity between groups. Future studies including patients with rheumatic mitral valve disease but without atrial fibrillation would help better isolate the specific genetic contribution to atrial fibrillation susceptibility.

5. Conclusion

The present study provides new insights into the genetic variability of the *PITX2* gene in the context of atrial fibrillation. Sequencing of exon 4 identified several rare missense variants located within the highly conserved Homeobox domain, a region essential for the transcriptional regulatory activity of *PITX2* and for the maintenance of atrial electrophysiological stability. Although these variants are currently classified as variants of uncertain significance, their rarity, evolutionary conservation, and functional localization suggest potential biological relevance. The observed excess of rare alleles and the slightly higher genetic diversity among patients further support the hypothesis that rare coding variants in key regulatory genes may contribute to individual susceptibility to atrial fibrillation. These findings are consistent with current models proposing that atrial fibrillation results from the combined effects of common susceptibility loci and rare functional variants affecting cardiac developmental and electrophysiological pathways. However, given the exploratory nature of

this study and the limited sample size, these observations should be interpreted with caution and are consistent with, but do not establish, a potential modulatory role of *PITX2* variants in atrial fibrillation susceptibility. Overall, this work highlights the importance of *PITX2* as a central molecular regulator of atrial electrophysiology and underscores the need for larger studies integrating extended genomic analyses and functional investigations to clarify the contribution of rare variants within the 4q25 locus to atrial fibrillation pathophysiology and risk stratification.

Abbreviations

AF	Atrial Fibrillation
RMVD	Rheumatic Mitral Valve Disease
<i>PITX2</i>	Paired-like Homeodomain Transcription Factor 2
PCR	Polymerase Chain Reaction
DNA	Deoxyribonucleic Acid
GWAS	Genome-Wide Association Studies
VUS	Variant of Uncertain Significance
ACMG	American College of Medical Genetics and Genomics
CADD	Combined Annotation Dependent Depletion
SIFT	Sorting Intolerant from Tolerant
AMOVA	Analysis of Molecular Variance
Hd	Haplotype Diversity
π	Nucleotide Diversity

Author Contributions

Diagne Papa Amath: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration

Mbaye Fatimata: Conceptualization, Formal analysis, Methodology, Project administration

Diop Momar Sokhna: Data curation

Ba Papa Ousmane: Data curation

Diop Moussa Seck: Data curation

Faye Cheikh Abdou Khadre: Data curation

Ba Papa Salmane: Data curation

Ciss Amadou Gabriel: Conceptualization, Methodology

Sembene Mbacke: Conceptualization, Methodology, Project administration

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Mocumbi AO, Sliwa K, Zühlke L. Rheumatic heart disease in Africa: epidemiology and challenges. *Lancet Glob Health*. 2018; 6(2): e130–e131. [https://doi.org/10.1016/S2214-109X\(17\)30484-9](https://doi.org/10.1016/S2214-109X(17)30484-9)
- [2] Zühlke L, Engel ME, Karthikeyan G, et al. Characteristics, complications, and gaps in evidence-based interventions in rheumatic heart disease: the Global Rheumatic Heart Disease Registry (REMEDY). *Eur Heart J*. 2015; 36(18): 1115–1122. <https://doi.org/10.1093/eurheartj/ehu449>
- [3] Chugh SS, Havmoeller R, Narayanan K, et al. Worldwide epidemiology of atrial fibrillation. *Circulation*. 2014; 129(8): 837–847. <https://doi.org/10.1161/CIRCULATIONAHA.113.005119>
- [4] Hindricks G, Potpara T, Dagres N, et al. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation. *Eur Heart J*. 2021; 42(5): 373–498. <https://doi.org/10.1093/eurheartj/ehaa612>
- [5] Watkins DA, Johnson CO, Colquhoun SM, et al. Global, regional, and national burden of rheumatic heart disease. *N Engl J Med*. 2017; 377(8): 713–722. <https://doi.org/10.1056/NEJMoa1603693>
- [6] Carapetis JR, Beaton A, Cunningham MW, et al. Acute rheumatic fever and rheumatic heart disease. *Nat Rev Dis Primers*. 2016; 2: 15084. <https://doi.org/10.1038/nrdp.2015.84>
- [7] Marijon E, Mirabel M, Celermajer DS, Jouven X. Rheumatic heart disease. *Lancet*. 2012; 379(9819): 953–964. [https://doi.org/10.1016/S0140-6736\(11\)61171-9](https://doi.org/10.1016/S0140-6736(11)61171-9)
- [8] Diao M, Kane A, Ndiaye MB, et al. Echocardiographic features of rheumatic heart disease in sub-Saharan Africa. *Cardiovasc J Afr*. 2018; 29(4): 219–224. <https://doi.org/10.5830/CVJA-2018-017>
- [9] Karthikeyan G, Teo KK, Islam S, et al. Clinical course of rheumatic heart disease in low-income countries. *Circulation*. 2018; 138(7): 709–721. <https://doi.org/10.1161/CIRCULATIONAHA.117.032791>
- [10] Sliwa K, Mocumbi AO. Forgotten cardiovascular diseases in Africa. *Nat Rev Cardiol*. 2020; 17(3): 145–156. <https://doi.org/10.1038/s41569-019-0284-1>
- [11] Zulkifly H, Lip GYH, Lane DA. Epidemiology of atrial fibrillation. *Int J Clin Pract*. 2018; 72(3): e13070. <https://doi.org/10.1111/ijcp.13070>
- [12] Thomas L, Abhayaratna WP. Left atrial structure and function in atrial fibrillation. *J Am Coll Cardiol*. 2017; 69(8): 1023–1034. <https://doi.org/10.1016/j.jacc.2016.11.064>
- [13] Tops LF, Delgado V, Bax JJ. The role of atrial remodeling in atrial fibrillation. *J Am Coll Cardiol*. 2020; 75(6): 632–645. <https://doi.org/10.1016/j.jacc.2019.12.018>
- [14] Nattel S, Heijman J, Zhou L, et al. Molecular basis of atrial fibrillation pathophysiology and therapy. *Circ Res*. 2020; 127(1): 51–72. <https://doi.org/10.1161/CIRCRESAHA.120.316363>
- [15] Gertz ZM, Raina A, Saghy L, et al. Atrial remodeling and arrhythmogenesis in valvular heart disease. *JACC Clin Electrophysiol*. 2022; 8(3): 350–362. <https://doi.org/10.1016/j.jacep.2021.10.016>
- [16] Gewitz MH, Baltimore RS, Tani LY, et al. Revision of the Jones criteria for the diagnosis of acute rheumatic fever. *Circulation*. 2015; 131(20): 1806–1818. <https://doi.org/10.1161/CIR.0000000000000205>
- [17] Beaton A, Okello E, Engelman D, et al. Rheumatic heart disease: pathogenesis and global burden. *Lancet*. 2021; 397(10286): 2087–2097. [https://doi.org/10.1016/S0140-6736\(21\)00503-5](https://doi.org/10.1016/S0140-6736(21)00503-5)
- [18] Roselli C, Chaffin MD, Weng LC, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet*. 2018; 50(9): 1225–1233. <https://doi.org/10.1038/s41588-018-0133-9>
- [19] Nielsen JB, Thorolfsdottir RB, Fritsche LG, et al. Biobank-driven genomic discovery of atrial fibrillation risk variants. *Nat Genet*. 2018; 50(9): 1234–1239. <https://doi.org/10.1038/s41588-018-0171-3>
- [20] Syeda S, Kirchhof P, Fabritz L. PITX2-dependent gene regulation in atrial fibrillation. *Basic Res Cardiol*. 2016; 111(6): 66. <https://doi.org/10.1007/s00395-016-0580-3>
- [21] Fatumo S, Chikowore T, Choudhury A, et al. A roadmap to increase diversity in genomic studies. *Nat Med*. 2022; 28(2): 243–250. <https://doi.org/10.1038/s41591-021-01672-4>
- [22] Christophersen IE, Rienstra M, Roselli C, et al. Large-scale analyses of atrial fibrillation genetics. *Nat Genet*. 2016; 48(7): 733–741. <https://doi.org/10.1038/ng.3593>