Pharmacogenomics and Personalized Medicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Whole Exome Sequencing Reveals Novel Genetic Variants Associated with Atrial Septal Defect in a Tibetan Patient Cohort

Hongwei Li[1,*], Yongjun He[2,*], Zhengyao Cai[1], Qianqiu Che[1], Yong Wu[1], Mingshuang Zhou[1], Zeng He[3], Liming Zhao[1]

[1]Department of Cardiology, Hospital of Chengdu Office of People's Government of Xizang Autonomous Region (Hospital.C.X.), Chengdu, Sichuan, 610041, People's Republic of China; [2]School of Medicine, Xizang Minzu University, Xianyang, Shaanxi, 712082, People's Republic of China; [3]Biobank, Hospital of Chengdu Office of People's Government of Xizang Autonomous Region (Hospital.C.X.), Chengdu, Sichuan, 610041, People's Republic of China

*These authors contributed equally to this work

Correspondence: Liming Zhao, Department of Cardiology, Hospital of Chengdu Office of People's Government of Xizang Autonomous Region (Hospital.C.X.), 20 Heng Street, Ximianqiao, Chengdu, Sichuan, 610041, People's Republic of China, Email ermine1048@163.com

Objective: Atrial septal defect (ASD) is a common congenital heart defect with incompletely understood genetic underpinnings, particularly in specific ethnic groups. This study aimed to identify novel genetic variants related to ASD within the Tibetan population using whole exome sequencing (WES).

Methods: Genomic DNA was extracted from blood samples of 17 Tibetan ASD patients. WES was performed using the Illumina HiSeq platform. After rigorous filtering, detection, and annotation of single nucleotide variations (SNVs) and insertion-deletions (InDels), potentially pathogenic variants were prioritized. Functional impact predictions were conducted using SIFT, PolyPhen V2, MutationTaster, and CADD databases to identify variants likely contributing to ASD etiology.

Results: We identified nine high-confidence candidate variants in Tibetan ASD patients, including rs145116532 (*ALKAL1*, c.287G >A: p. R96Q), rs374798430 (*AVL9*, c.1267G >A: p.D423N), rs138933092 (*C5*, c.4432C >T: p.R1478W), rs141638421 (*CRYAB*, c.470G>A: p. R157H), rs147287319 (*DOCK8*, c.989G>A: c.1193G>A: p.R330Q, p.R398Q), rs141616597 (*NTN3*, c.1243C>T: p.R415C), rs117506395 (*PIWIL1*, c.2207C>T: p.T736M), rs142533677 (*PLEKHG4*, c.2246G>A: p.R749Q), and rs118203532 (*TSC1*, c.1460C>G: p.S487C). Function annotation further suggested potential associations of *C5, CRYAB, PIWIL1*, and *TSC1* with congenital heart diseases.

Conclusion: This first WES-based study of Tibetan ASD patients reveals population-specific genetic determinants. The nine novel candidate variants, particularly in *C5, CRYAB, PIWIL1*, and *TSC1*, provide preliminary insights into ASD etiology in high-altitude populations and highlight potential targets for future diagnostic biomarker development.

Keywords: genetic variants, atrial septal defect, Tibetan population, whole exome sequencing

Introduction

Atrial septal defect (ASD) is a common congenital heart malformation arising from abnormal development of the atrial septum during embryogenesis. Characterized by an abnormal communication between the left and right atria, ASD accounts for 10%–15% of congenital heart diseases (CHD).[1] ASD can cause a variety of clinical complications, such as pulmonary hypertension and right heart failure.[2] In addition, patients with ASD may also experience arrhythmias and strokes.[3] Epidemiological studies indicate a female predominance in ASD, and familial aggregation patterns, particularly in subtypes like secundum ASD, strongly implicate genetic factors in its pathogenesis.[4,5] Genetic investigations have revealed that ASD exhibits substantial heterogeneity, involving mutations in numerous genes critical foe cardiac development and function. For example, variations in the NKX2-5 gene were identified in 8.3% of Indonesian ASD patients, underscoring its key regulatory role.[6] The GATA4 M310T mutation disrupts cardiac progenitor cell differentiation, leading to abnormal development of the atrial septum.[7] Besides, mutations in the

Received: 28 February 2025
Accepted: 25 August 2025
Published: 10 September 2025

MYH3 tail domain impair myocardial contractility transmission, increasing ASD risk.[8] Despite these advances, the genetic architecture of ASD remains incompletely characterized, with many contributing loci likely undiscovered.[9]

Whole exome sequencing (WES) has proven highly effective in identifying pathogenic variants within protein-coding regions across various diseases, including CHD.[10–13] Its application to ASD genetics has predominantly focused on specific populations, such as the Han Chinese, where Liu et al identified novel variants associated with sporadic ASD.[14] A critical gap exists in understanding the genetic basis of ASD within unique populations, particularly those residing in high-altitude regions. The Tibetan Plateau population represents a vital cohort for such investigation due to its distinct genetic adaptations to chronic hypobaric hypoxia and the documented elevation in CHD incidence associated with increasing altitude.[15,16] Remarkably, research on ASD-associated gene mutations specifically within the Tibetan population is currently absent. Given the potential interplay between hypoxia-adaptive genetic variants (eg, those related to the HIF pathway) and genes critical for cardiac development, genetic studies in this population are not only necessary but may reveal novel, altitude-relevant pathogenic mechanisms or modifiers that remain undetected in lowland cohorts.

Therefore, to address the significant gap in knowledge regarding ASD genetics in high-altitude populations and leveraging the unique context of the Tibetan Plateau, this study employs WES technology to screen for potential pathogenic mutations in 17 Tibetan children diagnosed with ASD. Our primary objective is to identify ASD-associated variants specific to or enriched within this population. This work aims to deepen the understanding of the genetic underpinnings of ASD, potentially uncovering novel pathways or targets influenced by high-altitude adaptation, and provide foundational data for future research into targeted therapeutic strategies for ASD in diverse populations, including those in high-altitude regions.

Materials and Methods

Study Population

This study enrolled 17 Tibetan pediatric ASD cases consecutively and randomly selected between June 2023 and October 2024 at the Second People's Hospital of Tibet Autonomous Region. A simple random sampling approach using computer-generated random numbers was applied to the pool of eligible consecutive patients meeting inclusion criteria to select the final cohort of 17 participants. This ensured unbiased selection from the available patient flow within the defined timeframe. All patients were echocardiographically confirmed by two independent clinicians. Inclusion criteria are: (1) isolated secundum ASD confirmed by transthoracic echocardiography (TTE) with a defect diameter ≥5 mm; (2) Tibetan ethnicity. Exclusion criteria comprised: (1) concurrent congenital cardiac anomalies (VSD, Tetralogy of Fallot); (2) genetic syndromes (Down syndrome, Marfan syndrome) or chromosomal abnormalities; (3) active malignancies, autoimmune disorders, or inflammatory conditions; (4) previous cardiac or catheter intervention; (5) Han or other ethnicity. Ethical approval was obtained from the Hospital of Chengdu Office of People's Government of Xizang Autonomous Region (Hospital.C.X). (Approval No.: Med-Eth-Re [2022] 77). All procedures strictly adhered to the Declaration of Helsinki (1964) and subsequent amendments. Legal guardians provided written informed consent after comprehensive study disclosure, with additional verbal assent obtained from participants aged ≥6 years. Detailed clinical characteristics of the study population are presented in Table 1.

Sample Collection

Professional healthcare personnel collected 5 mL whole blood specimens from 17 children diagnosed with ASD using sterile single-use vacuum blood collection devices. Venous blood samples were drawn into lavender-top vacuum tubes containing K2-EDTA anticoagulant. Following collection, the tubes were immediately inverted gently 8–10 times to ensure thorough mixing of the anticoagulant. Each specimen container received permanent labeling with distinct identifiers and phlebotomy timestamps. Processed samples underwent rapid cryopreservation in −80°C ultra-low temperature storage.

Whole-Exome Sequencing

Genomic DNA was extracted from blood samples using the Gentra Puregene Blood Kit (Qiagen) following the manufacturer's protocol. DNA integrity was assessed by agarose gel electrophoresis, while concentration and purity were measured with a NanoDrop 2000 spectrophotometer. Qualified samples met the following criteria: distinct electrophoretic bands without smearing, concentration ≥50 ng/μL, total quantity ≥1.5 μg, and OD260/280 = 1.8–2.0. Qualified DNA underwent whole-

Table 1 The Basic Information for ASD Patients

Sample ID	Age	Gender	Ultrasonic Electrocardiogram Report
COHD3_1	1 years and 4 months	Female	Anomalous pulmonary venous drainage + ASD + status post tricuspid valvuloplasty: Right atrial and right ventricular enlargement; CDFI: No shunt flow signals detected in the atria, ventricles, or great vessels. Mild tricuspid regurgitation is present.
COHD5_1	7 months	Male	CHD: CHD of the atrium; ASD (central type, 11mm); Left-to-right shunt at the atrial level; Enlarged right atrium and right ventricle.
COHD6_1	9 years old	Female	CHD; ASD (Central type, 8mm); Left-to-right shunt at the atrial level; Enlargement of the right atrium and right Ventricle.
COHD20_1	7 years old	Male	CHD; After repair of ASD: No color flow jet of shunt at the atrial level.
COHD22_1	7 years old	Male	CHD; After repair of ASD: No color doppler flow jet of shunt at the atrial level; Pulmonary artery hypertension (43 mmHg); Tricuspid regurgitation (minimal).
COHD29_1	1 years and 6 months	Male	CHD; ASD (central type, 8mm) with left-to-right shunting at the atrial level.
COHD35_1	12 years old	Male	CHD; After repair of ASD: No color doppler flow jet of shunt at the atrial level; Pulmonary artery hypertension (41 mmHg); Tricuspid regurgitation (minimal).
COHD39_1	11 years old	Male	CHD: Displacement of the tricuspid valve (with the distance between the septal leaflet of the tricuspid valve and the atrioventricular valve being 26mm); ASD (central type, 6mm) with left-to-right shunting at the atrial level.
COHD40_1	7 years old	Male	CHD; After repair of ASD: No color doppler flow jet of shunt at the atrial level.
COHD42_1	2 years and 5 months	Female	CHD; After repair of ASD: No color doppler flow jet of shunt at the atrial level; Tricuspid regurgitation (minimal).
COHD43_1	2 years and 7 months	Male	CHD; ASD (3mm); CDFI: Left-to-right shunting at the atrial level; Dilated coronary sinus (6mm); Persistent left superior vena cava.
COHD45_1	12 years old	Female	CHD; ASD (central type, 20mm); CDFI: Bidirectional shunting at the atrial level; Right heart enlargement; Pulmonary artery hypertension (66mmHg); Tricuspid regurgitation (moderate).
COHD46_1	14 years old	Male	CHD; ASD (4mm); CDFI: Left-to-right shunting at the atrial level; Tricuspid Regurgitation (Minimal).
COHD57_1	8 years old	Male	CHD; After repair of ASD: No color doppler flow jet of shunt at the atrial level.
COHD59_1	4 years old	Female	CHD; After repair of ASD: No color doppler flow jet of shunt at the atrial level.
COHD63_1	9 years old	Male	CHD; After repair of ASD: No color doppler flow jet of shunt at the atrial level; Left atrial enlargement; Pulmonary artery hypertension (36 mmHg); Tricuspid regurgitation (minimal).
COHD66_1	11 years old	Female	CHD; After repair of ASD: No color doppler flow jet of shunt at the atrial level; Tricuspid regurgitation (minimal).

Abbreviations: CHD, Congenital heart disease; ASD, atrial septal defect; CDFI, color doppler flow image.

exome sequencing (WES) through the following steps: 1) Genomic DNA was sheared into 100–500 bp fragments using a Covaris ultrasonicator. 2) DNA fragments were treated with T4 DNA polymerase for blunt-end formation and polynucleotide kinase for 5'-phosphorylation, followed by purification with Agencourt AMPure XP beads. 3) A single adenine nucleotide was added to the 3'-ends using a terminal transferase reaction. 4) Illumina sequencing adapters were ligated to DNA fragments using T4 DNA ligase in a buffer system. 5) Libraries were size-selected (300–400 bp) and purified with Agencourt SPRIselect beads. 6) Libraries were amplified with high-fidelity polymerase and quantified concentrations via Qubit fluorometry. 7) Prepared libraries were hybridized with biotinylated probes from the SureSelectXT Human All Exon V6 Kit using the SureSelectXT Reagent Kit. 8) Probe-bound DNA fragments were captured using Dynabeads® MyOne™ Streptavidin T1 magnetic beads. 9) Enriched libraries were amplified with high-fidelity polymerase and purified with magnetic beads. 10) Final library concentration (>5 ng/μL) and fragment size distribution (300–400 bp) were verified using Qubit and Agilent 2100 Bioanalyzer. 11) Libraries were sequenced on the Illumina HiSeq platform with 2 × 150 bp paired-end reads to generate FastQ files.

Mapping to Reference Sequences, Variants Detection, and Annotation

Quality assessment of raw sequencing data was performed using FastQC (v0.11.9), followed by filtering of low-quality reads (Phred score <20) and adapter sequences using Trimmomatic (v0.39). The processed reads were aligned to the human

reference genome (GRCh37/hg19) with BWA-MEM (v0.7.17). Duplicate reads were marked using Picard (v2.27.5). Base Quality Score Recalibration (BQSR) was conducted through GATK (v4.2.6.1) for local realignment optimization. Single-sample gVCF files containing single nucleotide variations (SNVs) and insertion-deletions (InDels) were generated via GATK HaplotypeCaller (https://software.broadinstitute.org/gatk/best-practices/). All detected variants were annotated using ANNOVAR (http://annovar.openbioinformatics.org/en/latest/) through comparison with updated population databases (1000Genomes, ExAC, gnomAF), functional annotation resources (SIFT, PolyPhen-2, MutationTaster, CADD, DANN), and disease-associated repositories (OMIM, HPO, HGMD, MGI). This annotation process evaluated variant allele frequencies, functional impacts, evolutionary conservation scores, and clinical pathogenicity predictions.

Filtering and Priority Classification of SNV/InDel

For all SNV/InDel variants, the following filtering criteria were applied to retain qualified sites for subsequent analysis: variants with frequencies below 0.01 in 1000Genomes, ExAC03 East Asian populations, and gnomAD Asian populations; variants with frequencies under 0.05 in GeneskyExonDB_Freq; and non-synonymous mutations located in exonic regions or splice sites. SNV/InDel variants were prioritized into four confidence tiers (First 1, First 2, Second, Third) based on aggregated evidence from these computational tools, with descending order reflecting predictive reliability. Pathogenicity predictions integrated SIFT, PolyPhen-2, MutationTaster, CADD, and DANN scores. Variants were defined as pathogenic if they met all the following criteria: variants were prioritized into First 1, CADD Raw score ≥ 4, CADD Phred score ≥ 25, DANN score ≥ 0.93, and concurrently annotated as "damaging" by SIFT, PolyPhen-2, and MutationTaster.

Candidate Susceptible Variants and Gene Selection

Three computational tools (SIFT, PolyPhen-2, and MutationTaster) were employed to predict the functional impact of amino acid substitutions. Variants unanimously classified as pathogenic by all three algorithms were prioritized as candidate susceptibility variants. Candidate genes are screened based on the functional analysis of genes corresponding to the candidate variants, combined with annotations from the OMIM, HPO, HGMD, and MGI databases, and integrated with information from Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways.

Results

Clinical Characteristics of the ASD Cases

This study recruited 17 ASD patients (11 males and 7 females) with isolated cardiac anomalies. As shown in Table 1, echocardiography confirmed ASD diameters ranging from 3 to 26 mm. Apart from ASD, other structural cardiac abnormalities or congenital malformations were detected in some patients. For example, in COHD3_1, anomalous pulmonary venous drainage and status post tricuspid valvuloplasty were observed. Besides, pulmonary artery hypertension and tricuspid regurgitation were present in COHD45_1. Detailed clinical characteristics of each patient are provided in Table 1.

WES Data

In this study, we employed Illumina HiSeq sequencing technology on 17 samples of ASD patients for comprehensive exon sequencing. The total raw reads obtained were 1,629,482,722, which underwent quality control filtering to ensure high-quality reads (Q20 > 99%, Q30 > 97%). This indicates the reliability of the sequencing data. Reads lengths ranged from 15 to 151bp, and detailed information is available in Supplementary Table S1. The average sequencing depth and coverage rate of the target region (exons) were 143.24× and 98.27% (≥10×), respectively. Within the exon region, 84.82% of the bases had a coverage depth of ≥30×. According to the SNV detection standard, sites with a coverage depth exceeding 10× have high reliability, indicating that more than 98% of the sites in the exon region of this study met the reliability threshold. It should be noted that there were some variations in the sequencing data quality among samples. For example, in COHD29_1, the Q30 value was relatively lower than other samples, which might affect the accuracy of variant detection to some extent. However, the overall high-quality sequencing data ensured the reliability of the subsequent analysis.

The samples were analyzed by alignment using the Picard software, and the results are summarized in Supplementary Table S2. The genome alignment generated an average of 95,851,925 reads, among which 95,374,877 were effectively aligned

(average proportion of 99.50%). The average number of bases covered in the target region (exons) was 60,767,983, with an average coverage depth of 143.24 times. In the alignment process, some samples showed slightly lower alignment rates than others. For instance, COHD39_1 had an alignment rate of 99.57%, while COHD29_1 had an alignment rate of 99.38%. These differences might be attributed to the inherent characteristics of the samples or the technical variations during the sequencing process. Despite these minor fluctuations, the overall high alignment rate ensured the accuracy of the subsequent variant calling.

A total of 350,398 mutation sites were identified using the GATK HaplotypeCaller method, including 291,532 SNV and 58,866 InDel sites. Based on the locations of these SNV and InDel sites in the genome, a classification and statistical analysis was carried out, and the distribution proportions of SNV and InDel sites in different regions are shown in Figure 1. Furthermore, the SNV and InDel sites were functionally classified, and the distribution proportions of SNV and InDel sites of different functional types are presented in Figure 2. Meanwhile, the SNV and InDel sites were classified and counted according to genotype (homozygous/heterozygous), and the numbers of SNV and InDel sites with different genotypes are provided in Supplementary Table S3. During the classification and counting process, we found that the number of heterozygous SNVs was generally higher than that of homozygous SNVs in most samples. For example, in COHD3_1, there were 56,638 heterozygous SNVs and 38,369 homozygous SNVs. This pattern might be related to the genetic background of the patients and the nature of the disease. However, further investigation is needed to fully understand the underlying mechanisms.

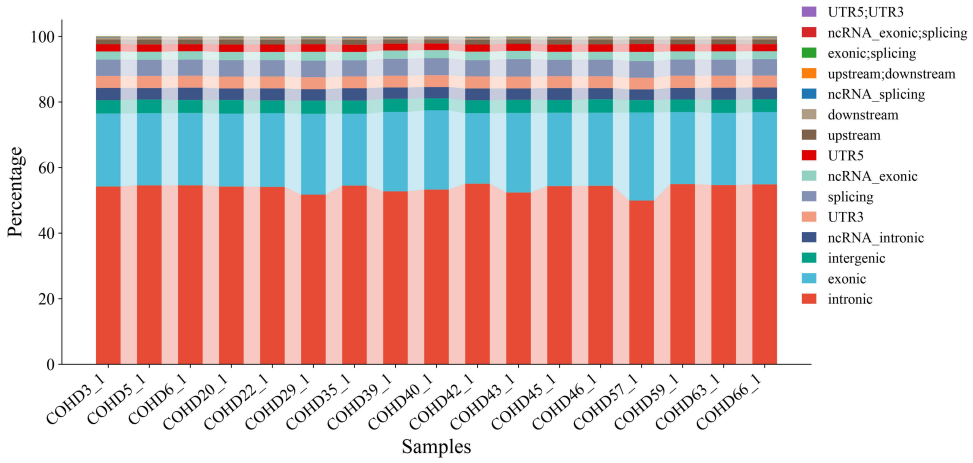

Figure 1 Analysis of variant distribution in genomic regions.

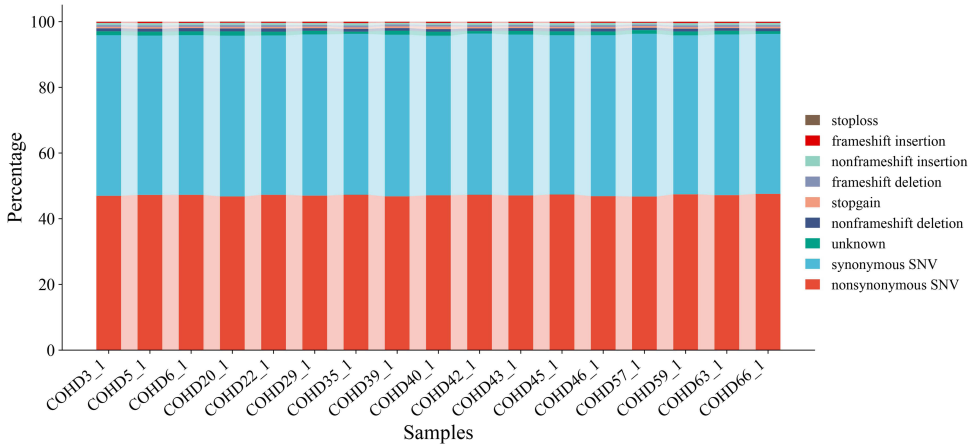

Figure 2 Categorization of functional types across all variants.

https://doi.org/10.2147/PGPM.S525556

Identification of ASD-Associated Pathogenic Gene Variants

In this study, 4,087 variants classified as "Category I" were identified in the proband's exome through comprehensive analysis of detected SNVs/InDels using multiple filtering approaches. These highest-priority variants were selected based on stringent criteria: minor allele frequencies below 0.001 in the 1000 Genomes dataset, less than 0.01 in ExAC03, gnomAD, and ESP6500 databases, coupled with an SNP calling quality score of H. Following this filtration process, 156 loci remained for subsequent evaluation. Pathogenicity was strictly defined by meeting all of the following computational criteria: (1) classification as "First 1" priority, (2) CADD Raw score ≥ 4, (3) CADD Phred score ≥ 25, (4) DANN score ≥ 0.93, and (5) unanimous "Damaging (D)" predictions from SIFT, PolyPhen-2, and MutationTaster. Applying these criteria, exactly 9 variants across 9 genes (*ALKAL1, AVL9, C5, CRYAB, DOCK8, NTN3, PIWIL1, PLEKHG4, TSC1*) fulfilled every condition and were classified as pathogenic (Table 2). Critically, all 9 variants exhibited: (i) "First 1" priority status, (ii) unanimous "D" calls from SIFT, PolyPhen-2, and MutationTaster, (iii) CADD Raw scores between 4.007 (*TSC1*) and 4.623 (*C5*) (all ≥4), (iv) CADD Phred scores between 27.0 (*TSC1*) and 32.0 (*C5*) (all ≥25), and (v) DANN scores between 0.989 (*TSC1*) and 1.000 (*DOCK8, PLEKHG4*) (all ≥0.93).

Detailed characteristics of these pathogenic variants are shown in Table 3. These mutations are located in the exon region and belong to nonsynonymous SNVs, including *ALKAL1* (rs145116532: c.287G >A: p. R96Q), *AVL9* (rs374798430: c.1267G >A: p.D423N), *C5* (rs138933092: c.4432C >T: p.R1478W), *CRYAB* (rs141638421: c.470G>A: p.R157H), *DOCK8* (rs147287319: c.989G>A: c.1193G>A: p.R330Q, p.R398Q), *NTN3* (rs141616597: c.1243C>T: p. R415C), *PIWIL1* (rs117506395: c.2207C>T: p.T736M), *PLEKHG4* (rs142533677: c.2246G>A: p.R749Q), and *TSC1* (rs118203532: c.1460C>G: p.S487C). Besides, functional annotation analysis of these 9 pathogenic gene variants revealed distinct cardiovascular associations: *C5* was linkage to Takotsubo (stress) cardiomyopathy, *CRYAB* exhibited correlations with cardiomyopathy, monogenic dilated cardiomyopathy, and myofibrillar myopathy, *PIWIL1* showed relevance to congenital heart disease, while *TSC1* displayed associations with cardiac conduction disorder & epilepsy and cardiac rhabdomyoma (Table 4). Beyond cardiovascular links, functional annotation uncovered broader biological

Table 2 Prediction of Damaging Effects on Protein Functions of Nine-Candidate Pathological Variations

Gene	First priority	Freq_Alt (1000g)	ExAC03	esp6500	gnomAD_exome	SIFT Pred	POLYPhen V2 Pred	MutationTaster Pred	Cadd Raw	Cadd Phred	Dann
ALKAL1	First1	< 0.001	< 0.001	< 0.001	< 0.001	D	D	D	4.097	27.7	0.999
AVL9	First1	< 0.001	< 0.001	< 0.001	< 0.001	D	D	D	4.384	31	0.999
C5	First1	< 0.001	< 0.001	< 0.001	< 0.001	D	D	D	4.623	32	0.999
CRYAB	First1	< 0.001	0.001	< 0.001	< 0.001	D	D	D	4.096	27.7	0.999
DOCK8	First1	< 0.001	0	< 0.001	< 0.001	D	D	D	4.158	28.3	1
NTN3	First1	< 0.001	0.0001	< 0.001	0.0013	D	D	D	4.287	29.5	0.999
PIWIL1	First1	< 0.001	0	< 0.001	< 0.001	D	D	D	4.095	27.7	0.999
PLEKHG4	First1	< 0.001	0	< 0.001	< 0.001	D	D	D	4.057	27.4	1
TSC1	First1	< 0.001	0.002	< 0.001	< 0.001	D	D	D	4.007	27	0.989

Abbreviations: Freq_Alt (1000g), alternative allele frequency data in 1000 Genomes Project; gnomAD exome, Genome Aggregation Database exome; SIFT Pred, Sorting Intolerant From Tolerant Prediction; POLYPhen V2 Pred, Polymorphism Phenotyping v2 Prediction; Cadd Phred, Combined Annotation Dependent Depletion Prediction; Cadd Raw, Combined Annotation Dependent Depletion Raw Score; Dann, Deleterious Annotation of Genetic Variants.

Table 3 Genetic Information for Predicted Genes

Gene	SNP ID	Chr	Position	Gene Region	cDNA Change	Protein Change	Functions
ALKAL1	rs145116532	chr8	52539869	Exon3	c.287G >A	p. R96Q	Nonsynonymous SNV
AVL9	rs374798430	chr7	32570071	Exon11	c.1267G >A	p.D423N	Nonsynonymous SNV
C5	rs138933092	chr9	120962743	Exon36	c.4432C >T	p.R1478W	Nonsynonymous SNV
CRYAB	rs141638421	chr11	111908822	Exon3, 4	c.470G>A	p.R157H	nonsynonymous SNV
DOCK8	rs147287319	chr9	334292	Exon 10, 11	c.989G>A, c.1193G>A	p.R330Q, p.R398Q	Nonsynonymous SNV
NTN3	rs141616597	chr16	2473110	Exon3	c.1243C>T	p.R415C	Nonsynonymous SNV
PIWIL1	rs117506395	chr12	130367144	Exon19	c.2207C>T	p.T736M	Nonsynonymous SNV
PLEKHG4	rs142533677	chr16	67285340	Exon 12, 14,15	c.2246G>A	p.R749Q	Nonsynonymous SNV
TSC1	rs118203532	chr9	132906118	Exon14, 15	c.1460C>G	p.S487C	Nonsynonymous SNV

Abbreviations: SNP, Single nucleotide polymorphism; Chr, Chromosome; cDNA, Complementary DNA; Q, Glutamine; N, Asparagine; W, Tryptophan; H, Histidine; C, Cysteine; M, Methionine; SNV, Single nucleotide variant.

Table 4 Function Analysis for the Candidate Genes

Gene	OMIM	HPO	HGMD_gene	MGI	GO_BP	GO_MF	GO_CC	KEGG_Pathway
ALKAL1	–	–	–	–	positive Regulation of neuron projection development	Protein binding	Extracellular region	–
AVL9	–	–	–	–	Cell migration	–	Cytoplasm	–
C5	Autosomal recessive	Autosomal recessive inheritance	Takotsubo (stress) cardiomyopathy	Respiratory system phenotype	Positive regulation of vascular endothelial growth factor production	Endopeptidase inhibitor activity	Extracellular region	Complement and coagulation cascades
CRYAB	Cardiomyopathy, dilated, 1II	EMG abnormality, EMG: myopathic abnormalities, Cardiomyopathy, Hypertrophic cardiomyopathy	Cardiomyopathy, Monogenic Dilated Cardiomyopathy, Myofibrillar myopathy	Cardiovascular system phenotype	Regulation of cell death, negative regulation of cell growth	Structural molecule activity	Cardiac myofibril	Protein processing in endoplasmic reticulum
DOCK8	Autosomal recessive	Cerebral vasculitis	XY disorder of sex development, Antibody deficiency	Nervous system phenotype	Immunological synapse formation, small GTPase mediated signal transduction	Protein binding	Cytoplasm	–
NTN3	–	–	Autism spectrum disorder	–	Axon guidance	Signaling receptor binding	Cellular_component	Axon guidance
PIWIL1	–	–	Congenital heart disease	Endocrine/ exocrine gland phenotype	Regulation of translation	Single-stranded RNA binding	Nucleus	–
PLEKHG4	–	Motor deterioration, Sensory axonal neuropathy	Autism spectrum disorder, Cerebellar ataxia, autosomal dominant	–	Regulation of catalytic activity	Guanyl-nucleotide exchange factor activity, protein binding	–	–
TSC1	Lymphangioleiomyomatosis, Autosomal dominant, Tuberous sclerosis-1	Cardiac rhabdomyoma, arrhythmia	Cardiac conduction disorder and epilepsy, Cardiac rhabdomyomas	Cardiovascular system phenotype	Cardiac muscle cell differentiation	Protein binding	Nucleus, cytoplasm, 1,lipid droplet	AMPK signaling pathway, PI3K-Akt signaling pathway

roles. For example, *CRYAB* participated in regulation of cell death and negative regulation of cell growth (GO_BP), while *ALKAL1* modulated neuron projection development (GO_BP), DOCK8 was involved in immunological synapse formation (GO_BP), and *NTN3* functioned in axon guidance (GO_BP, KEGG). This pleiotropy highlights the diverse potential mechanisms by which these pathogenic variants may contribute to ASD. Further investigation is essential to delineate their specific roles in ASD pathogenesis and interactions with genetic/environmental factors.

Discussion

Congenital heart disease (CHD), encompassing conditions such as ASD, arises from multifactorial origins involving both environmental and genetic determinants. Epidemiological evidence indicates that environmental exposures, including gestational viral infections, may elevate CHD susceptibility. However, genetic contributors are increasingly recognized as pivotal, particularly in ASD cases,[17] where genomic variants and rare mutations in cardiac-associated genes exhibit prominent associations with disease pathogenesis.[18] Thus, this study performed WES to identify novel genetic alterations in ASD patients within the Tibetan population. Our study showed that nine high-confidence candidate variants in the *ALKAL1, AVL9, C5, CRYAB, DOCK8, NTN3, PIWIL1, PLEKHG4*, and *TSC1* genes that were firstly reported in ASD patients of Tibetan descent. Function annotation analysis suggested potential associations of *C5, CRYAB, PIWIL1*, and *TSC1* with CHD, supporting further investigation into their role in ASD pathogenesis within this population.

ALKAL1 is a member of the ALK family, primarily regulates downstream signaling pathways via receptor tyrosine kinases, including cell proliferation, differentiation, and tissue development.[19,20] Mutations in *ALKAL1* ligands are closely associated with nervous system dysfunction (eg, neuropathic pain),[21] and immune system dysfunction.[19] While *ALKAL1* has not been extensively studied in CHD, its known roles in development and immune modulation align with the multifactorial nature of CHD. Our identification of the rs145116532 mutation in ASD patients suggests a potential, previously unexplored link between *ALKAL1* dysregulation and cardiac maldevelopment, warranting further investigation into its specific mechanisms in heart formation.

AVL9 is a member of the adenylate cyclase-associated protein family, primarily functioning through the regulation of the cytoskeleton, involvement in signal transduction, and cell migration. Abnormal expression of A*VL9* is closely linked to the development of various diseases. For instance, high expression of AVL9 in colorectal cancer has been considered a marker of tumor malignancy.[22] In non-small cell lung cancer, *AVL9* interacts with miR-203a-3p to regulate tumor cell proliferation and migration, thus influencing the clinical features of the cancer.[23] Additionally, mutations in the *AVL9* gene may be associated with certain hereditary eye diseases, such as myopia.[24] The pathogenic *AVL9* rs374798430 mutation identified in this study represents a novel association with ASD. Given *AVL9*'s involvement in fundamental cellular processes critical for morphogenesis, this finding opens a new avenue for exploring how cytoskeletal regulation contributes specifically to septal development defects.

C5 is a key component of the complement system and plays a significant role in immune responses and tissue inflammation. Its excessive activation is implicated in cardiovascular pathologies like acute myocardial infarction (endothelial damage)[25] and venous thromboembolism.[26] Genetic variants causing *C5* dysfunctional are linked to increased cardiovascular disease risk, potentially via enhanced inflammation. For instance, *C5a* overproduction promotes atherosclerosis.[27] Our discovery of the pathogenic C5 p.R1478W variant in ASD patients expands the role of the complement system beyond acquired cardiovascular diseases to encompass congenital cardiac malformations. This finding strongly supports the emerging hypothesis that dysregulated immune/inflammatory pathways during development are critical etiological factors in CHD.[28,29] We speculate that this mutation might disrupt normal cardiac development by inducing localized inflammation or interfering with developmental signaling, ultimately leading to structural defects like ASD.

CRYAB (αB-crystallin) is a small molecular heat shock protein widely distributed across various tissues in the human body. *CRYAB* is vital for protein homeostasis and cellular protection, particularly in the heart where it guards against ischemia-reperfusion (I/R) injury by inhibiting apoptosis and ferroptosis.[30] *CRYAB* is also closely associated with cardiovascular diseases. It also modulates cardiac fibroblast proliferation and fibrosis via the LBH pathway, contributing to cardiac repair.[31,32] Mutations in *CRYAB* established causes of cardiomyopathy. For instance, the A527G mutation of *CRYAB* has been linked to dilated cardiomyopathy (DCM) and congenital cataracts.[33] Additionally, these mutations may result in abnormal protein function, contributing to heart failure.[34,35] Our identification of a pathogenic *CRYAB* mutation (rs141638421) in ASD patients links this cardioprotective and cardiomyopathy-associated gene to congenital septal

defects. This finding broadens the phenotypic spectrum of *CRYAB* variants and highlights its essential role not only in sustaining adult heart function but also in ensuring proper cardiac development. Potential mechanisms include impaired protection of immature cardiomyocytes against stress or dysregulation of pathways vital for septal formation.

PIWIL1, interacting with piRNAs, regulates gene expression and cellular functions. It demonstrates neuroprotective effects.[36] In cardiovascular contexts, *PIWIL1* mutations may disrupt miRNA biogenesis, increasing atherosclerosis and stroke risk.[37] The pathogenic *PIWIL1* p.T736M variant found in our study suggests a potential novel role for this RNA regulatory protein in CHD pathogenesis, specifically ASD. This finding warrants exploration into whether *PIWIL1*-dependent regulatory pathways are involved in cardiac development, potentially through modulating gene expression networks critical for septal formation.

TSC1 complexes with TSC2 to suppress the mTOR signaling pathway, a key regulator of cell growth, proliferation, and metabolism. Mutations in *TSC1* are a primary cause of tuberous sclerosis complex, a condition frequently involving cardiac rhabdomyomas.[38] Such mutations can compromise cardiac myocyte performance[39] and perturb critical processes including oxidative stress response and autophagy, potentially leading to heart failure.[40] The deleterious *TSC1* p.S487C variants identified in our study connects this central mTOR signaling pathway, fundamental to TSC pathology, to isolated ASD within the Tibetan population. Although cardiac rhabdomyomas are a hallmark feature of TSC, our finding suggests that specific *TSC1* mutations may confer susceptibility to isolated septal defects independently of the full syndromic phenotype. This highlights the pleiotropic effects of mTOR pathway dysregulation on cardiac development and suggests overlapping pathogenic mechanisms between syndromic and non-syndromic CHD.

Our study provides the first comprehensive report of ASD-associated genetic variants in the Tibetan population. The identification of pathogenic mutations in genes like *C5, CRYAB, PIWIL1*, and *TSC1*, which have established or emerging roles in cardiovascular biology, significantly strengthens the genetic basis of ASD. While environmental factors specific to the Tibetan plateau (chronic hypoxia) may contribute to CHD risk, our findings highlight the critical and potentially predominant role of genetic susceptibility, even in this unique high-altitude adapted population. Due to their distinct genetic background shaped by adaptation,[41–43] Tibetans may exhibit a unique spectrum and prevalence of CHD-associated mutations, making this population particularly valuable for identifying genotype–phenotype relationships and gene–environment interactions specific to cardiac development under hypoxic stress.

However, it is necessary to acknowledge the limitations of this study. While WES effectively detects rare variants, functional validation of identified mutations remains essential to confirm their pathogenicity and clarify the precise molecular mechanisms. Furthermore, replication in larger, independent cohorts, including other Tibetan ASD patients and diverse populations, is crucial to assess the findings' generalizability and population-specificity. Such expanded research will be vital for developing population-specific genetic screening strategies and elucidating CHD' broader pathophysiological mechanisms.

Conclusion

This study identified nine high-confidence candidate variants in genes (*ALKAL1, AVL9, C5, CRYAB, DOCK8, NTN3, PIWIL1, PLEKHG4*, and *TSC1*) potentially associated with ASD in the Tibetan population through WES. Functional annotation analysis indicated plausible biological links between *C5, CRYAB, PIWIL1*, and *TSC1* genes with CHD. Collectively, these findings provide preliminary insights into the genetic underpinnings of ASD within this specific population and suggest candidate markers warranting further investigation for diagnostic and therapeutic development. We explicitly acknowledge that key limitations require attention: While WES effectively detects rare variants, functional validation of these variants is essential to confirm their pathogenicity and elucidate their precise roles in ASD and CHD pathogenesis. Furthermore, replication of these findings in larger, independent cohorts, encompassing diverse Tibetan ASD patients as well as other populations, is necessary to robustly assess their generalizability and potential population-specificity. Addressing these limitations through structured future research will be vital not only for validating our findings but also for translating them into clinical practice. Specifically, extended investigation may enable the development of population-tailored genetic screening approaches and elucidate shared pathological mechanisms between ASD and CHD.

Data Sharing Statement

Data are available from the corresponding author upon reasonable request.

Ethics Statement

Our study was approved by the Ethics Committee of the Hospital of Chengdu Office of People's Government of Xizang Autonomous Region (Hospital.C.X.).

Acknowledgments

We would like to express gratitude to the Second People's Hospital of Tibet Autonomous Region for the provision of blood samples.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by the Science and Technology Projects of Xizang Autonomous Region, China (No. XZ202402ZY0003), the Science and Technology Major Project of Tibetan Autonomous Region of China (No. XZ202201ZD0001G01), and the Science and Technology Projects of Xizang Autonomous Region, China (No. XZ202403ZY0006).

Disclosure

There are no conflicts of interest related to this study.

References

1. Brida M, Chessa M, Celermajer D, et al. Atrial septal defect in adulthood: a new paradigm for congenital heart disease. *Eur Heart J*. 2022;43(28):2660–2671. doi:10.1093/eurheartj/ehab646
2. Tobler D, Greutmann M. Simple cardiac shunts in adults: atrial septal defects, ventricular septal defects, patent ductus arteriosus. *Heart*. 2020;106(4):307–314. doi:10.1136/heartjnl-2019-314700
3. Kaisbain N, Lim WJ, Kim HS. Atrial septal defect with Crochetage sign presenting with pulmonary artery thrombosis. *BMJ Case Rep*. 2021;14(7):e244180. doi:10.1136/bcr-2021-244180
4. Larsen LA, Hitz MP. human genetics of atrial septal defect. *Adv Exp Med Biol*. 2024;1441:467–480.
5. Chaithra S, Agarwala S, Ramachandra NB. High-risk genes involved in common septal defects of congenital heart disease. *Gene*. 2022;840(146745).
6. Rozqie R, Satwiko MG, Anggrahini DW, et al. NKX2-5 variants screening in patients with atrial septal defect in Indonesia. *BMC Med. Genomics*. 2022;15(1):91.
7. Bu H, Sun G, Zhu Y, et al. The M310T mutation in the GATA4 gene is a novel pathogenic target of the familial atrial septal defect. *BMC Cardiovascul Disorders*. 2021;21(1):12. doi:10.1186/s12872-020-01822-5
8. Maran S, Ee R, Faten SA, et al. Mutations in the tail domain of MYH3 contributes to atrial septal defect. *PLoS One*. 2020;15(4):e0230982. doi:10.1371/journal.pone.0230982
9. Broberg M, Ampuja M, Jones S, et al. Genome-wide association studies highlight novel risk loci for septal defects and left-sided congenital heart defects. *BMC Genomics*. 2024;25(1):256. doi:10.1186/s12864-024-10172-x
10. Morton SU, Quiat D, Seidman JG, Seidman CE. Genomic frontiers in congenital heart disease. *Nat Rev Cardiol*. 2022;19(1):26–42. doi:10.1038/s41569-021-00587-4
11. Cai R, Tan Y, Wang M, et al. Detection of novel pathogenic variants in two families with recurrent fetal congenital heart defects. *Pharmacogenomics Personalized Med*. 2023;16:173–181. doi:10.2147/PGPM.S394120
12. Sevim Bayrak C, Zhang P, Tristani-Firouzi M, Gelb BD, Itan Y. De novo variants in exomes of congenital heart disease patients identify risk genes and pathways. *Genome Med*. 2020;12(1):9. doi:10.1186/s13073-019-0709-8
13. Dong W, Kaymakcalan H, Sc J, et al. Mutation spectrum of congenital heart disease in a consanguineous Turkish population. *Mol Genetics Genomic Med*. 2022;10(6):e1944. doi:10.1002/mgg3.1944
14. Liu Y, Cao Y, Li Y, et al. Novel genetic variants of sporadic Atrial Septal Defect (ASD) in a Chinese Population Identified by Whole-Exome Sequencing (WES). *Med Sci Monit*. 2018;24:1340–1358.

15. González-Andrade F. High Altitude as a Cause of Congenital Heart Defects: a Medical Hypothesis Rediscovered in Ecuador. *High Alt Med Biol*. 2020;21(2):126–134. doi:10.1089/ham.2019.0110
16. Chun H, Yue Y, Wang Y, et al. High prevalence of congenital heart disease at high altitudes in Tibet. *Eur J Preventive Cardiol*. 2019;26(7):756–759. doi:10.1177/2047487318812502
17. Gabriel GC, Ganapathiraju M, Lo CW. The role of cilia and the complex genetics of congenital heart disease. *Ann Rev Genomics Hum Genet*. 2024;25(1):309–327. doi:10.1146/annurev-genom-121222-105345
18. Raggi F, Cangelosi D, Becherini P, et al. Transcriptome analysis defines myocardium gene signatures in children with ToF and ASD and reveals disease-specific molecular reprogramming in response to surgery with cardiopulmonary bypass. *J Transl Med*. 2020;18(1):21. doi:10.1186/s12967-020-02210-5
19. Katic L, Priscan A. Multifaceted roles of ALK family receptors and augmentor ligands in health and disease: a comprehensive review. *Biomolecules*. 2023;13(10):1490. doi:10.3390/biom13101490
20. Fadeev A, Mendoza-Garcia P, Irion U, et al. ALKALs are in vivo ligands for ALK family receptor tyrosine kinases in the neural crest and derived cells. *Proceedings of the National Academy of Sciences of the United States of America* 2018, **115**:E630–e638.
21. Defaye M, Iftinca MC, Vm G, et al. The neuronal tyrosine kinase receptor ligand ALKAL2 mediates persistent pain. *J Clin Invest*. 2022;132(12). doi:10.1172/JCI154317.
22. Li D, Zeng Y, Shen P, et al. AVL9 is upregulated in and could be a predictive biomarker for colorectal cancer. *Cancer Manage Res*. 2021;13:3123–3132. doi:10.2147/CMAR.S301844
23. Liang J, Sun T, Wang G, Zhang H. Clinical significance and functions of miR-203a-3p/AVL9 axis in human non-small-cell lung cancer. *Personalized Medicine*. 2020;17(4):271–282. doi:10.2217/pme-2019-0108
24. Cl S, Am M, Ry C, et al. Myopia in African Americans Is Significantly Linked to Chromosome 7p15.2-14.2. *Invest Ophthalmol Visual Sci*. 2021;62(9):16. doi:10.1167/iovs.62.9.16
25. Vahldieck C, Löning S, Hamacher C, et al. Dysregulated complement activation during acute myocardial infarction leads to endothelial glycocalyx degradation and endothelial dysfunction via the C5a:C5a-Receptor1 axis. *Front Immunol*. 2024;15:1426526. doi:10.3389/fimmu.2024.1426526
26. Skjeflo EW, Brækkan SK, Ludviksen JK, et al. Elevated plasma concentration of complement factor C5 is associated with risk of future venous thromboembolism. *Blood*. 2021;138(21):2129–2137. doi:10.1182/blood.2021010822
27. Martínez-López D, Roldan-Montero R, García-Marqués F, et al. Complement C5 protein as a marker of subclinical atherosclerosis. *J Am College Cardiol*. 2020;75(16):1926–1941. doi:10.1016/j.jacc.2020.02.058
28. Wienecke LM, Cohen S, Bauersachs J, Mebazaa A, Chousterman BG. Immunity and inflammation: the neglected key players in congenital heart disease? *Heart Fail Rev*. 2022;27(5):1957–1971. doi:10.1007/s10741-021-10187-6
29. Zhao X, Gao C, Chen H, Chen X, Liu T, Gu D. C-reactive protein: an important inflammatory marker of coronary atherosclerotic disease. *Angiology*. 2024;33197241273360. doi:10.1177/00033197241273360
30. Wu A, Zhong C, Song X, et al. The activation of LBH-CRYAB signaling promotes cardiac protection against I/R injury by inhibiting apoptosis and ferroptosis. *iScience*. 2024;27(5):109510. doi:10.1016/j.isci.2024.109510
31. Wu A, Zhang L, Chen J, et al. Limb-bud and Heart (LBH) mediates proliferation, fibroblast-to-myofibroblast transition and EMT-like processes in cardiac fibroblasts. *Mol Cell Biochem*. 2021;476(7):2685–2701. doi:10.1007/s11010-021-04111-7
32. Xu Y, Wu A, Chen J, Song X, Chen M, Liu Q. Limb-Bud and Heart (LBH) upregulation in cardiomyocytes under hypoxia promotes the activation of cardiac fibroblasts via exosome secretion. *Mediators Inflammation*. 2022;2022:8939449. doi:10.1155/2022/8939449
33. Ha C, Kim D, Bak M, et al. CRYAB stop-loss variant causes rare syndromic dilated cardiomyopathy with congenital cataract: expanding the phenotypic and mutational spectrum of alpha-B crystallinopathy. *J Human Genetics*. 2024;69(3–4):159–162. doi:10.1038/s10038-023-01218-1
34. Sadeh M, Rahat D, Meiner V, et al. Multi-system neurological disorder associated with a CRYAB variant. *Neurogenetics*. 2021;22(2):117–125. doi:10.1007/s10048-021-00640-x
35. Kelters IR, Verbueken D, Beekink T, et al. Generation of human induced pluripotent stem cell (hiPSC) lines derived from three patients carrying the pathogenic CRYAB (A527G) mutation and one non-carrier family member. *Stem Cell Res*. 2024;80(103497):103497. doi:10.1016/j.scr.2024.103497
36. Matei N, Camara J, McBride D, et al. Intranasal wnt3a attenuates neuronal apoptosis through Frz1/PIWIL1a/FOXM1 pathway in MCAO rats. *J Neurosci*. 2018;38(30):6787–6801. doi:10.1523/JNEUROSCI.2352-17.2018
37. Wang M, Gu J, Shen C, et al. Association of MicroRNA biogenesis genes polymorphisms with risk of large artery atherosclerosis stroke. *Cell Mol Neurobiol*. 2022;42(6):1801–1807. doi:10.1007/s10571-021-01057-8
38. Wang S, Sun H, Wang J, et al. Detection of TSC1/TSC2 mosaic variants in patients with cardiac rhabdomyoma and tuberous sclerosis complex by hybrid-capture next-generation sequencing. *Mol Genetics Genomic Med*. 2021;9(10):e1802. doi:10.1002/mgg3.1802
39. Uchiyama H, Masunaga Y, Ishikawa T, et al. TSC1 intragenic deletion transmitted from a mosaic father to two siblings with cardiac rhabdomyomas: identification of two aberrant transcripts. *Eur J Med Genet*. 2020;63(11):104060. doi:10.1016/j.ejmg.2020.104060
40. Cho CS, Kim Y, Park SR, et al. Simultaneous loss of TSC1 and DEPDC5 in skeletal and cardiac muscles produces early-onset myopathy and cardiac dysfunction associated with oxidative damage and SQSTM1/p62 accumulation. *Autophagy*. 2022;18(10):2303–2322. doi:10.1080/15548627.2021.2016255
41. Zhang G, Cui C, Wangdue S, et al. Maternal genetic history of ancient Tibetans over the past 4000 years. *J Genetics Genomics = Yi Chuan Xue Bao*. 2023;50(10):765–775. doi:10.1016/j.jgg.2023.03.007
42. Li X, Xu S, Li X, et al. Novel insight into the genetic signatures of altitude adaptation related body composition in Tibetans. *Front Public Health*. 2024;12:1355659. doi:10.3389/fpubh.2024.1355659
43. Ms P, Yh L, Qk S, et al. Genetic and cultural adaptations underlie the establishment of dairy pastoralism in the Tibetan Plateau. *BMC Biol*. 2023;21(1):208. doi:10.1186/s12915-023-01707-x