Molecular Genetics and Metabolism Reports 41 (2024) 101152

Contents lists available at ScienceDirect

Molecular Genetics and Metabolism Reports

journal homepage: www.elsevier.com/locate/ymgmr

Using artificial intelligence and promoter-level transcriptome analysis to identify a biomarker as a possible prognostic predictor of cardiac complications in male patients with Fabry disease

Hiroshi Kobayashi [a,c,*], Norio Nakata [b,d], Sayoko Izuka [a], Kenichi Hongo [e], Masako Nishikawa [f]

[a] Division of Gene Therapy, Research Center for Medical Sciences, The Jikei University of Medicine, 3-25-8, Nishi-shimbashi, Minato-ku, Tokyo 105-8461, Japan
[b] Division of Artificial Intelligence Medicine, Research Center for Medical Sciences, The Jikei University of Medicine, 3-25-8, Nishi-shimbashi, Minato-ku, Tokyo 105-8461, Japan
[c] Department of Pediatrics, The Jikei University of Medicine, 3-25-8, Nishi-shimbashi, Minato-ku, Tokyo 105-8461, Japan
[d] Department of Radiology, The Jikei University of Medicine, 3-25-8, Nishi-shimbashi, Minato-ku, Tokyo 105-8461, Japan
[e] Division of Cardiology, Department of Internal Medicine, The Jikei University of Medicine, 3-25-8, Nishi-shimbashi, Minato-ku, Tokyo 105-8461, Japan
[f] Clinical Research Support Center, The Jikei University of Medicine, 3-25-8, Nishi-shimbashi, Minato-ku, Tokyo 105-8461, Japan

ARTICLE INFO

Keywords:
Artificial intelligence
Cap analysis of gene expression
Fabry disease
Cardiac complication
Magnetic resonance imaging
CHN1

ABSTRACT

Fabry disease is the most frequently occurring form of lysosomal disease in Japan, and is characterized by a wide variety of conditions. Primarily, the three major types of concerns associated with Fabry disease observed during adulthood that must be prevented are central nervous system, renal, and cardiac complications. Cardiac complications, such as cardiomyopathy, cardiac muscle fibrosis, and severe arrhythmia, are the most common mortality causes in patients with Fabry disease. To predict cardiac complications of Fabry disease, we extracted RNA from the venous blood of patients for cap analysis of gene expression (CAGE), performed likelihood ratio tests for each RNA expression dataset obtained from individuals with and without cardiac complications, and analyzed the correlation between cardiac functional factors observed using magnetic resonance imaging data extracted using artificial intelligence algorithms and RNA expression. Our findings showed that CHN1 expression was significantly higher in male Fabry disease patients with cardiac complications and that it could be associated with many cardiac functional factors. *CHN1* encodes a GTPase-activating protein, chimerin 1, which is specific to the GTP-binding protein Rac (involved in oxidative stress generation and the promotion of myocardial fibrosis). Thus, CHN1 is a potential predictive biomarker of cardiac complications in Fabry disease; however, further studies are required to confirm this observation.

1. Introduction

Fabry disease (OMIM #301500) is classified as a lysosomal storage disease derived from the dysfunction of the enzyme alpha-galactosidase (GLA) localized in lysosomes, and is one of the most frequently occurring lysosomal diseases. Globotriaosylceramide, a substrate of glycosphingolipids, accumulates in the lysosomes of endothelial and smooth muscle cells, renal and cardiac cells, autonomic ganglia, sweat glands, cornea, and the central nervous system due to GLA dysfunction, resulting in various symptoms such as cerebral infarction, arrhythmia, cardiac hypertrophy, renal failure, hypohidrosis, angiokeratoma, corneal crowding, cataracts, and limb pain. This disease is inherited as an X-linked recessive trait, with male patients generally exhibiting severe symptoms; however, female patients are also frequently affected with symptoms as severe as male patients, although symptom onset may be delayed [1].

Cardiovascular complications, such as arrhythmia and heart failure, are the leading causes of death of Fabry disease. Cardiac magnetic resonance imaging (MRI) is an important diagnostic toolfor the evaluation of cardiac conditions. Characteristic findings include fibrosis of the left ventricular myocardium, as determined by late-gadolinium enhancement. The cardiac complications of Fabry disease are difficult to detect which delays treatment. Therefore, an improved method to diagnose Fabry disease is required.

* Corresponding author at: Division of gene therapy, Research Center for Medical Sciences, Department of Pediatrics, The Jikei University of Medicine, 3-25-8, Nishi-shimbashi, Minato-ku, Tokyo 105-8461, Japan.
E-mail address: hrkb2012@gmail.com (H. Kobayashi).

https://doi.org/10.1016/j.ymgmr.2024.101152
Received 10 May 2024; Received in revised form 6 October 2024; Accepted 7 October 2024

Table 1
Profiles of study participants.

		Fabry disease		Healthy control (C)
		Cardiac complication		
		Positive (A)	Negative (B)	
Number		15	8	5
Mean age		42.5 y (9–56 y)	24 y (9–36 y)	35 y (22–55 y)
Mean onset age		16.6 y (3–43 y)	9.8 y (5–16 y)	–
Age at start of ERT		32.6 y (9–47 y)	16.5 y (8–22 y)	–
Type of cardiac complication	LVH	12 (80 %)	–	–
	LAD	6 (40 %)	–	–
	Valvular disease	5 (33 %)	–	–
	LGE	3 (20 %)	–	–
	Arrhythmia	11 (73 %)	–	–
CNS complication		5 (33 %)	1 (12.5 %)	–
Renal complication		3 (20 %)	1 (12.5 %)	–
Mean DS3 score		10.5 points	6.0 points	NA

y; years-old
LVH; left ventricular hypertrophy
LAD; left atrial dilation
LGE; Late Gadolinium Enhancement
Gadolinium delay: delayed excretion after gadolinium contrast on MRI with implications for myocardial fibrosis.
CNS complications, such as complications of the central nervous system, refer to distinct complications or imaging changes, such as microcerebral hemorrhage or cerebral infarction, excluding parenchymal brain signal changes.
Renal complications were defined as overt proteinuria and decreased renal function.
DS3; Fabry Disease Severity Scoring System: CNS; central nervous system.

Cardiac MRI is the gold standard for evaluating myocardial function, volume, and scarring. Additionally, cardiac MRI is the optimal technique for comprehensive tissue characterization, including the assessment of myocardial edema, myocardial siderosis, myocardial perfusion, and diffuse myocardial fibrosis. Cardiac MRI is therefore an indispensable tool for evaluating congenital heart diseases and heart failure [2].

Herein, MRI was used for analysis for the following reasons: MRI does not require radiation exposure, children over a certain age can be subjected to testing, and it is more objective than ultrasound examination for evaluation of cardiac function.

Radiomic studies using MRI and gene expression profiling include studies on brain tumors [3,4] and breast cancer [5]. In addition, no radiomic studies on MRI and gene expression profiling for genetic disorder exist, except for studies on brain MRIs in patients with juvenile-onset cognitive impairment, white matter dystrophy [6], and single-gene disorders strongly associated with autistic spectrum disorders [7].

In this study, we used artificial intelligence (AI) and next-generation sequencing to investigate the prognostic factors of cardiac complications in Fabry disease. AI is an excellent tool for image recognition, and "segmentation" technology was used to analyze images of cardiac complications observed in Fabry disease, and evaluate its results with those of genetic analysis to detect prognostic factors for cardiac complications.

2. Materials and methods

2.1. Profiles of patients and individuals from the control group

The study population included 23 male patients with Fabry disease who visited the Jikei University Hospital between 2015 and 2021. Fabry disease was diagnosed using enzyme assays or genetic analysis. The patients were classified into: group A: Fabry disease with cardiac complications ($n = 15$); group B: patients without cardiac complications ($n = 8$); and group C: healthy patients ($n = 5$). Groups A and B included a certain number of patients with central nervous system and renal complications. Since these complications increase in frequency with age and are considered independent of cardiac complications, patients with these complications were also included in groups A and B. Cardiac complications were defined as left ventricular hypertrophy determined using echocardiography, late gadolinium enhancement observed using cardiac MRI, and a history of distinct arrhythmias (e.g., atrial fibrillation

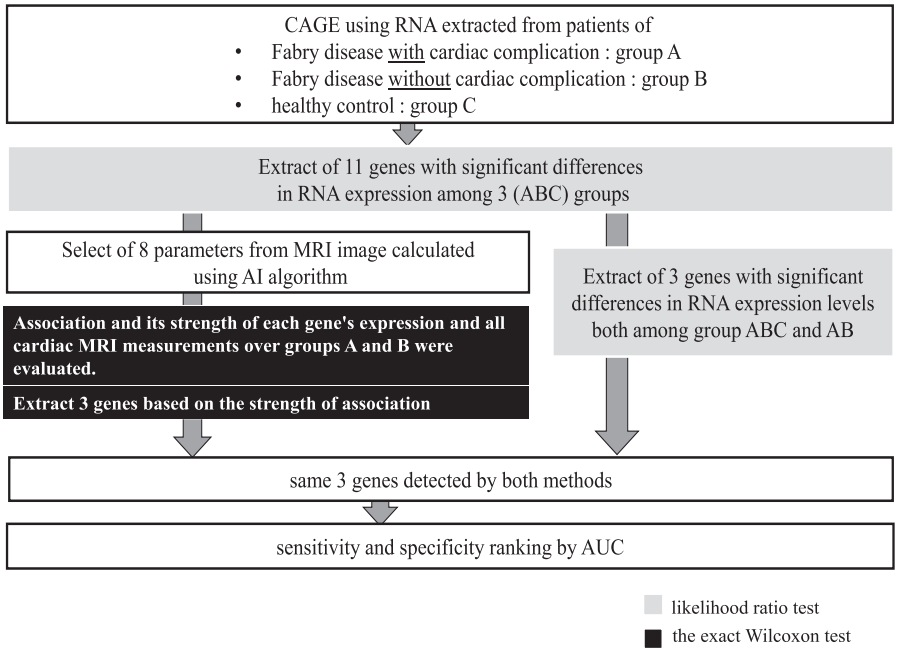

Fig. 1. Flow chart of this study.
The same three genes were detected as candidates for prognostic predictors of cardiac complications in patients with Fabry disease.
CAGE; Cap analysis of gene expression, AI; artificial intelligence, AUC; area under the receiver operating characteristic curve.

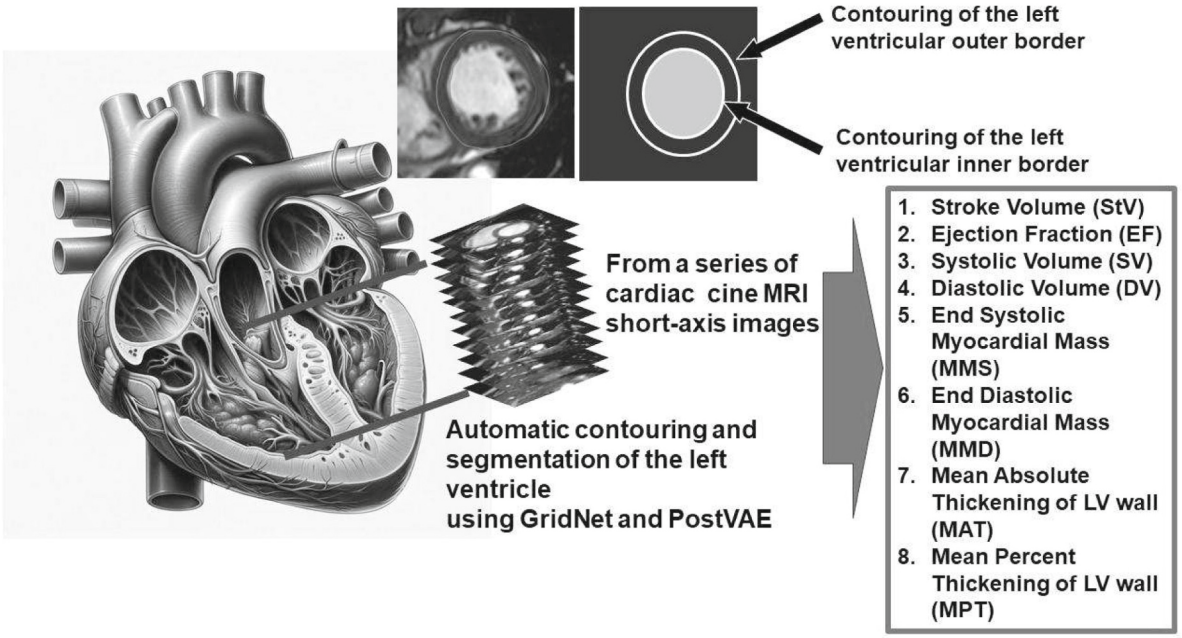

Fig. 2. AI algorithm to select and calculate MRI parameters.
The neural network mechanism using GridNet and Post-VAE predicts short-axis cine-MRI image data (classification). Then, it identifies the position of the heart (localization) and pixel in the heart (contouring and segmentation). Consequently, calculating the eight MRI parameters.
Post-VAE: post variational autoencoder.

and ventricular tachycardia). Classification of cardiac complications revealed a higher rate of left ventricular hypertrophy and arrhythmia (Table 1). The Fabry disease clinical severity score, DS3 shown at the bottomof the Table1, is characterized by its ease of scoring compared to that of the scoring system used by the multiple sclerosis society of India. The DS3 scoring system comprises five domains and 12 items based on four clinical assessment domains (peripheral nerves, kidneys, heart, and central nervous system) and one interview domain for the patient. The evaluation of DS3 requires estimated glomerular filtration rate, proteinuria, echocardiogram findings, electrocardiogram, and head imaging test results. Based on correlation with the Clinical Global Impression of Severity (CGI), a score of <8 is considered mild, <12 is moderate, and ≥ 12 is severe.

2.2. Summary of the study depicted using a flow chart

We extracted RNA from the sera of 28 subjects using the *PAX* gene blood RNA system (PreAnalytiX™ from Quigen and BD) and stored it at −80 °C, and transported samples to DNAFORM (Yokohama, Kanagawa, Japan) for CAGE analysis based on the following criteria: using likelihood ratio tests we listed the highly expressed genes, and identified genes that showed significant differences in expression among the ABC and AB groups. (Fig. 1).

2.3. CAGE library preparation and sequencing

CAGE library preparation, sequencing, mapping, and gene expression analyses were performed using the DNAFORM software. The quality of the total RNA extracted from the blood was assessed using the Bioanalyzer system (Agilent) to ensure that the RNA integrity number was over 7.0. After the depletion of globin mRNA using the Globin-Zero Gold kit (Illumina), cDNA was synthesized from total RNA using random primers. The ribose diols in the 5′ cap structures of RNAs were oxidized, and then biotinylated. Biotinylated RNA/cDNA was selected using streptavidin beads (cap-trapping). After RNA digestion using RNase ONE/H and adaptor ligation to both ends of the cDNA, double-stranded cDNA libraries (CAGE libraries) were constructed. The CAGE libraries were sequenced using single-end reads of 75 nt using the NextSeq 500 system (Illumina). Reads obtained (CAGE tags) were mapped to the human hg38 genome using BWA [8] (version 0.7.17), and unmapped reads were mapped using the HISAT2 [9] (version 2.0.5).

CAGE tag clustering was performed using the RECLU pipeline [10]. Tag count data were clustered using a modified Paraclu program. Clusters with count per million (CPM) <0.1 were disregarded. Regions with 90 % overlap between replicates were extracted using BEDTools [11] (version 2.12.0). The clusters with irreproducible discovery rate was ≥ 0.1, those longer than 200 bp were disregarded.

Differentially expressed genes among the three groups were detected using the DESeq2 package [12] (version 1.20.0) from size factor-normalized count data using the likelihood ratio test function. Individuals from each group were treated as replicates. Genes with Benjamini–Hochberg adjusted *p*-values (padj) <0.05 were treated as differentially expressed (i.e., candidate genes). For each differentially expressed gene, a pairwise comparison of expression levels between groups was also conducted. Genes with non-adjusted $p < 0.05$ were treated as differentially expressed in the pairwise comparison.

2.4. MRI parameters using AI (Fig. 2)

Of the enrolled patients, 16 underwent cardiac MRI, including 11 in group A and five in group B. The cardiac MRI data of patients with Fabry disease (AB group) were analyzed using an AI algorithm (deep learning) with the QIR software.

Cine-MRI images in short- and long-axis views are available for each case. Images were acquired using the Steady-state Free Precession (SSFP) MR imaging protocol with the following settings: typical thickness 7–8 mm, gap $\leq$2 mm, TR 40–161 ms, TE 1.36–1.48 ms, flip angle 50–80°, FOV 400 or 416 × 512 mm image matrix using the 3.0-T Skyra system (Siemens Healthcare) and 1.5-T Avanto system (Siemens Healthcare).

In this study, cine-MRI DICOM images were used for the segmentation and calculation of various quantified parameters using the CASIS

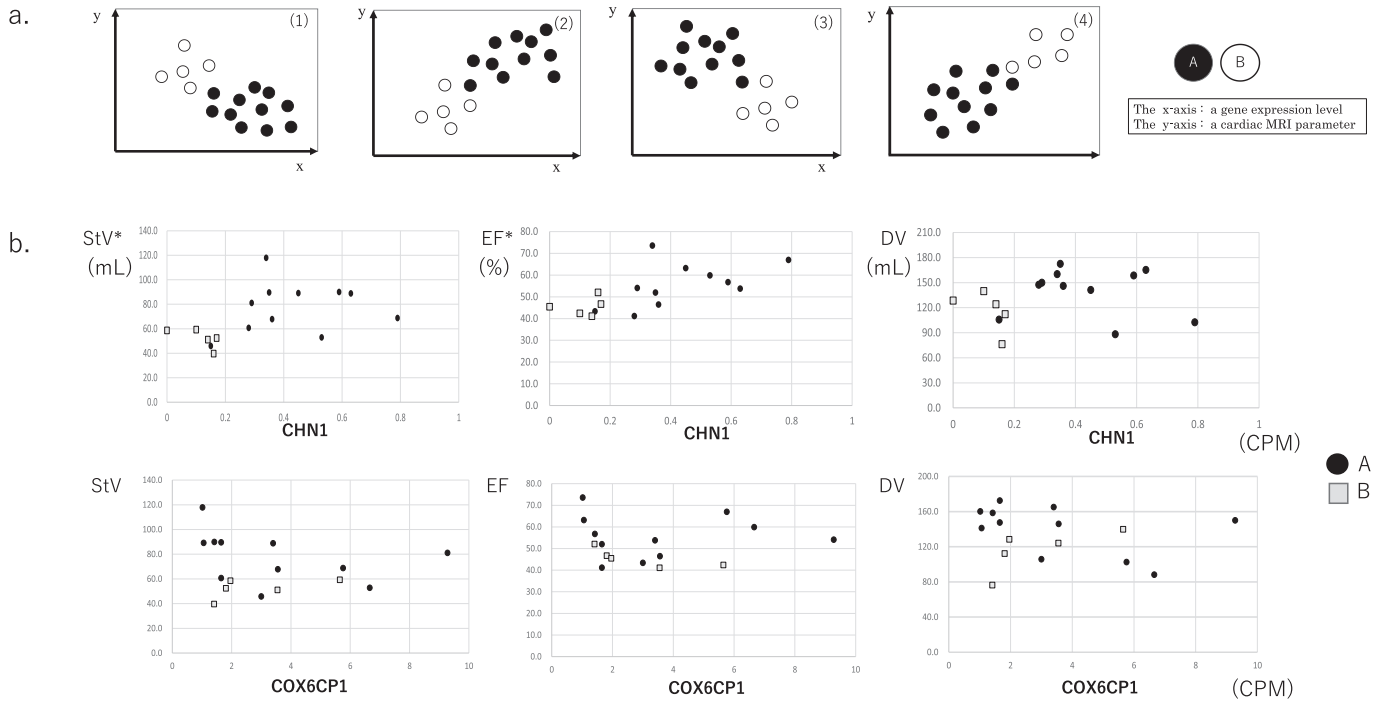

Fig. 3. Distribution of the levels of gene expression and MRI parameters.
a. Typical scatter plots with a gene expression to be detected on the x-axis and MRI parameters on the y-axis to distinguish between the A and B groups.
(1) In the group with cardiac complications (group A, black), X was greater and Y was lower than those in the group without cardiac complications (Group B, white).
(2) In the group with cardiac complications (group A, black), X was larger and Y was larger than those in the group without cardiac complications (Group B, white).
(3) In the group with cardiac complications (group A, black), X was smaller and Y was larger than those in the group without cardiac complications (Group B, white).
(4) In the group with cardiac complications (group A, black), X was smaller and Y was smaller than those in the group without cardiac complications (Group B, white).
b. Scatter plots with gene expression on the x-axis and MRI parameters on the y-axis.
In CHN, the MRI parameters StV and EF indicated by * have similar distributional patterns to a. (2), and the sum of the respective standardized values (X + Y) shows significant differences between groups A and B.
However, in COX6CP1, no distinction between groups A and B was possible.

QIR version 4 (Dijon, France) software for cardiac MRI studies.

We used cine-MRI are GridNet [13] and post-variational autoencoder (Post-VAE) [14] for the segmentation of the left ventricle (LV) in this study. GridNet is a novel deep convolutional neural network-based novel), fully automatic MRI cardiac segmentation method, and Post-VAE is a post-processing technique that automatically transforms the erroneous segmentation map into a valid one after segmentation.

LV cine MRI images were acquired over the cardiac cycle, which alternates between the relaxing phase (diastole) and contracting phase (systole). Segmentation of the LV and myocardium using cine MRI facilitates the quantification of the ventricular volume, mass, and ejection fraction (EF) [15]. In this study, stroke volume (StV), EF, diastolic volume (DV), systolic volume (SV), end-systolic myocardial mass (SMM), and end-diastolic myocardial mass (DMM) were calculated for all cardiac cine MRI examinations. Additionally, the mean absolute and percent thickening of the LV wall (MAT and MPT, respectively) were automatically calculated from the average value of LV thickness obtained using 16-segment polar mapping [16–18].

2.5. Statistics

The SAS 9.4 (SAS Institute Inc., Cary, NC, USA) and R software were used for statistical analyses.

To summarize the patients characteristics, we used mean values with standard deviations or medians with interquartile ranges for continuous variables and percentages for categorical variables. The association between a given candidate gene and all cardiac parameters on MRI (StV, EF, DV, SV, MMS, MMD, MAT, and MPT) over groups A and B was examined. MRI measurements were considered a surrogate of Fabry disease with/without cardiac complication.

If a candidate gene could be used as a prognostic predictive marker of cardiac complication in Fabry disease, then a correlation between some parameters of the MRI images and candidate gene should exist, and their scatter plot should show a similar trend to that of the four patterns displayed in Fig. 3 a (1)–(4).

To compensate for the lack of power due to the small sample size, we made use of information on a possibly predictive gene and an MRI parameter similar to that used by Yoshida et al. in their study [19].

The strength of the association between a given candidate gene and Fabry disease with/without cardiac complication, i.e., the ability of a gene as a predictive marker, is evaluated using the number of MRI parameters that are significantly different between groups A and B. This strength was measured as follows:

1) For each pair of a candidate gene and cardiac parameter from MRI images, calculate their rank correlation coefficient (RCCF) and record its sign [positive (+) or negative (−)].
2) Standardize each value of the candidate gene and of the cardiac parameter using respective means and standard deviations through groups, and denote their standardized value X and Y, respectively.
3) Select either X + Y or X − Y according to the sign of RCCF (+ or −), and compare the selected variable between groups A and B using the exact Wilcoxon test.
4) A two-sided p-value <0.05 was used as a reference criterion. As Fabry disease is rare and the sample size was small, the multiplicity of tests was not adjusted.

Table 2
Likelihood ratio test results for gene groups detected using CAGE.

Gene symbol	A vs B			C vs A			C vs B			Stat	p-value	p adj
	Base mean	Log2 fold change	p-value	Base mean	Log2 fold change	p-value	Base mean	Log2 fold change	p-value			
REXO2	545.841	−0.091	0.325	600.986	−0.421	<0.001	586.981	−0.547	<0.001	24.041	<0.001	**0.026***
ACVR1	293.388	−0.236	**0.001***	322.312	−0.249	0.005	302.446	−0.438	<0.001	27.352	<0.001	**0.007***
CHN1	5.843	−1.926	**<0.001***	6.626	1.381	0.003	2.856	−0.112	0.878	23.528	<0.001	**0.027***
SLC22A23	103.200	−0.318	0.231	143.946	−1.157	<0.001	139.742	−1.500	<0.001	22.004	<0.001	**0.029***
FUNDC2	264.764	0.129	0.594	370.366	−1.494	<0.001	411.093	−1.481	<0.001	22.618	<0.001	**0.029***
MT1F	247.871	0.575	**<0.001***	208.702	0.048	0.715	260.463	0.502	0.035	20.845	<0.001	**0.046***
HMBS	306.051	0.018	0.925	410.716	−1.278	<0.001	441.341	−1.322	<0.001	22.400	<0.001	**0.029***
COX6CP1	56.134	0.754	0.090	90.876	−2.348	<0.001	119.720	−1.825	0.001	22.227	<0.001	**0.029***
AC073476.4	492.763	0.129	0.157	435.822	0.565	<0.001	450.766	0.682	<0.001	28.894	<0.001	**0.005***
MIR3662	1.249	1.342	0.085	2.938	−3.563	<0.001	4.285	−2.454	0.005	21.967	<0.001	**0.029***
MIR1244–1	122.957	0.072	0.432	108.833	0.680	<0.001	111.461	0.785	<0.001	36.960	<0.001	**0.000***

Eleven genes with significant differences among the A, B, and C groups were extracted from over 3000 genes. * $p < 0.05$.

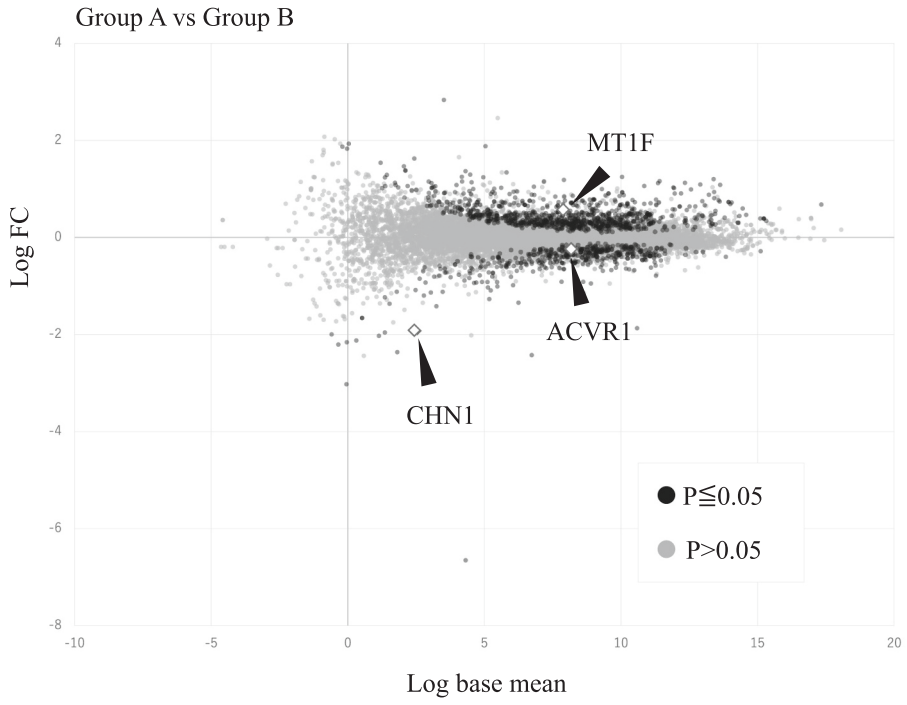

Fig. 4. MA-plot showing gene expression between groups A and B.
The x-axis represents the average gene expression (log CPM) and y-axis represents the ratio of variation in expression between groups (log FC), and each gene is plotted as a single point.
Gray dots: Non-differentially expressed genes; Black dots: Differentially expressed genes; Outlined squares: Three selected candidate biomarkers.

The area under the receiver operating characteristic (ROC) curve (AUC) was used as a performance evaluation index. The 95 % confidence interval (CI) of the ROC AUC was estimated using the method described in [20]. We expected the sensitivity and specificity to be ≥0.8 and ≥ 0.75, respectively, to determine the cut-off value of a gene expression level. If the expected values were not obtained, ordinary criteria neglecting the misclassification cost or prevalence (e.g., Youden index, distance to [0,1], sensitivity, and specificity equality) were considered.

3. Results

3.1. Genes with significant differences in RNA expression levels between the AB and ABC groups

We tested the likelihood ratios between the A, B, and C groups and identified the top 11 genes that differed significantly. Subsequently, out of these 11 genes, we identified three, *ACVR1*, *CHN1*, and *MT1F*, that significantly differed between the A and B groups (Table 2), and visualized them using the MA-plot for gene expression between groups A and B (Fig. 4).

3.2. Genes associated with MRI parameters of cardiac complication in Fabry disease

Several typical scatter plots are shown in Fig. 3. In Fig. 3a (2), the candidate gene *CHN* (x-axis) and MRI parameters StV and EF (y-axis) have similar distributional patterns to Fig. 3a (2). As described in Section 2.5, comparing X + Y using the exact Wilcoxon test distinguished group A from B. However, the candidate gene *COX6CP1*did not show any similar distributional pattern to that in Fig. 3a (1)–(4).

After comparing either X + Y or X - Y using the exact Wilcoxon test, *p*-values indicating significant differences between the AB groups for each MRI parameter are shown in bold in Table 3. Of the eight MRI parameters identified, the top three genes that exhibited significant differences

Table 3
Exact Wilcoxon test results for genes with significant differences between the AB groups regarding MRI parameters.

	Mean standardized values of MRI parameters		**Exact Wilcoxon test to compare group AB (*$p < 0.05$)**										
	Group A	Group B	**REXO2**	**ACVR1**	**CHN1**	**SLC 22A23**	**FUNDC2**	**MT1F**	**HMBS**	**COX6CP1**	**AC 073476.4**	**MIR 3662**	**MIR 1244–1**
Number	11	5											
Mean age (y)	47.5	30.8											
StV	***53.5***	***42.3***	0.1804	**0.0055***	**0.0018***	0.2674	0.6612	**0.0055***	0.0687	0.0687	0.913	**0.0087***	0.0897
EF	***53.5***	***42.3***	0.1451	**0.0032***	**0.0018***	0.2212	0.8269	**0.0009***	0.2212	0.038	0.1149	**0.0055***	0.4409
DV	***51.5***	***46.7***	0.8269	0.0897	0.2212	0.6612	0.2674	0.0897	0.2674	0.2674	0.913	**0.0380***	0.4409
SV	***48.4***	***53.6***	0.7427	**0.0087***	**0.0380***	0.913	1.0000	**0.0192***	0.8269	0.7427	0.3773	0.5096	0.8269
MMS	***54.3***	***40.6***	0.0897	**0.0018***	**0.0009***	0.0687	**0.0133***	**0.0009***	0.0517	**0.0275***	0.7427	**0.0032***	**0.0275***
MMD	***54***	***41.1***	0.0687	**0.0087***	**0.0018***	0.0897	**0.0192***	**0.0032***	0.0687	**0.038***	0.5833	**0.0055***	0.0517
MAT	***54.3***	***40.5***	0.0897	**0.0018***	**0.0005***	0.1451	**0.0087***	**0.0009***	0.2212	**0.0087***	0.1149	**0.0018***	0.3196
MPT	***52.8***	***42.7***	0.2212	**0.0133***	**0.0005***	0.5096	1.0000	**0.0087***	0.5833	0.8269	0.2674	0.332	0.7427

The 11 genes displayed in Table 2 were ranked using the number of MRI parameters significantly different between patients with Fabry disease with cardiac complications (group A) and without cardiac complications (group B).
P-values were calculated using the exact Wilcoxon test to compare the AB groups for each MRI parameter. *$p < 0.05$.
StV; stroke volume, EF; Ejection Fraction, SV; Systolic volume, DV; Diastolic Volume, MMS; End systolic myocardial mass, MMD; End diastolic myocardial mass, MAT; mean absolute thickening of the LV wall, MPT; Mean percent thickening of the LV wall.

Fig. 5. Sensitivity and specificity of *CHN1*, *ACVR2*, *MTF1* were estimated using the area under the ROC curve.
As a reference, *FUNDC2* expression, which did not differ significantly between the AB groups in the likelihood ratio test, was plotted.

in various items were selected.

A ROC curve was generated for each gene extracted in Step 3. If expression levels were below the limit of quantification, the expression was considered 0.

The gene expression levels of *ACVR1*, *CHN1*, and *MT1F* were associated with the distribution of seven of eight MRI parameters, suggesting that they could be used to distinguish between groups with and without cardiac complications. In the likelihood ratio test, *ACVR1*, *CHN1*, and *MT1F* were the same genes that showed significant differences between the A and B as well as A, B, and C groups (Table 2).

3.3. Sensitivity and specificity of the selected genes are potentially associated with cardiac complications

The ROC curves of the three candidate genes of groups A and B are shown in Fig. 5. Among the three candidate genes, *CHN1* had the largest ROC AUC (0.921, 95 % CI 0.806–1.000). *MT1F* and *ACVR1* showed similar ROC AUC values (0.842, 95 % CI 0.676–1.000 and 0.833, 95 % CI 0.658–1.000, respectively). In *CHN1*, a sensitivity and specificity of 0.8 were observed at the cutoff value (0.74). For *ACVR1*, these were 0.8 and 0.75, respectively, at the cutoff value (0.56).

3.4. Correlation diagram between age and expression levels of CHN1

No strong correlation existed between CHN1 expression and age. (Fig. 6) Thus, aging was not considered a significant inducer of CHN1 expression in patients with Fabry disease.

4. Discussion

Fabry disease exhibits various syndromes; however, the three major complications after adolescence are central neurological syndromes including cerebral infarction and cerebral hemorrhage; cardiac complications such as arrhythmia and cardiomyopathy; and renal complications. Cardiac complications account for more than half of all deaths

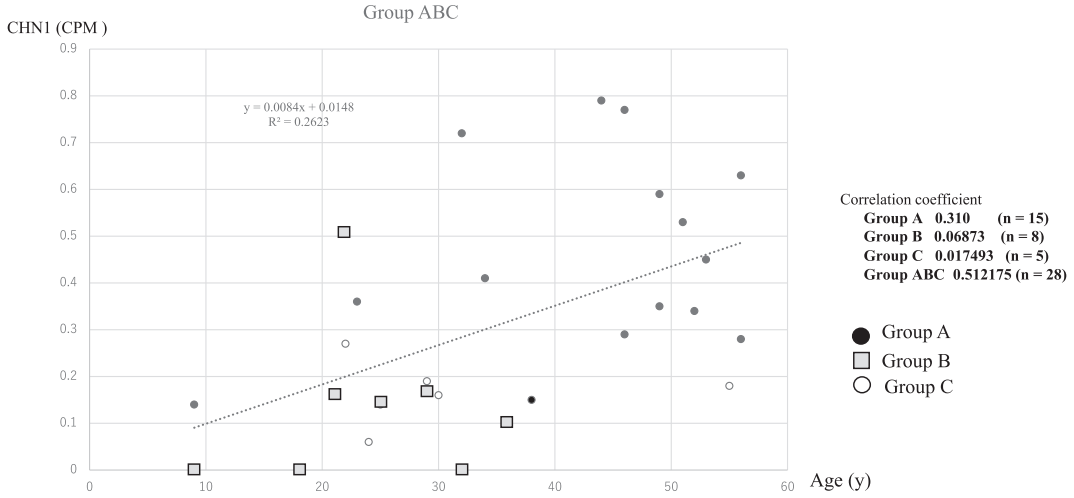

Fig. 6. Correlation between CHN1 expression (CPM) and age.
We plotted *CHN1* RNA expression (CPM) on the y-axis and patient age (years) on the x-axis, and discovered a weak correlation between *CHN1* RNA expression level and age.

by Fabry disease in Japan and globally. Cardiac complications are difficult to predict, and once they become apparent are often irreversible, even with treatments such as enzyme replacement therapy. Therefore, determining the prognosis of cardiac complications is crucial.

Of 11 genes with significant differences in expression levels among the three groups, three genes that significantly differed between the groups with and without concomitant cardiac complications in Fabry disease were identified. The MRI images of 12 patients in the Fabry disease group were re-analyzed with the Quantified Imaging Resources software using AI and RCCF between eight parameters.

The expression levels of each gene were determined, and significant differences were found in many items between the A and B groups. Ranking of genes revealed that the top three genes were identical. When the sensitivity and specificity of these three genes were tested, CHN1 was at the top of the list. In addition, AI technology has made remarkable progress, especially in the field of diagnostic imaging. The AI segmentation used in this study was extremely useful in the digitization of myocardium movement in the MRI and calculation of the eight parameters required for this study.

The *CHN1* gene is well-established as the primary cause of Duane's retraction syndrome [21,22], and encodes a GTPase-activating protein chimerin 1 specific for the GTP-binding protein Rac, which participates in oxidative stress generation and promotion of myocardial fibrosis, and for a signaling pathway that has an instrumental role in the early development of the central nervous system. Sustained activation of signaling pathways involving GTP-binding proteins may lead to cardiac hypertrophy and fibrosis [23,24].

Our findings suggest that *CHN1* is involved in the development of the cardiac complications of Fabry disease which are characterized by fibrosis and arrhythmias in the left ventricular myocardial base. The cutoff point for identifying gene clusters was 0.28 CPM, a sensitivity of 0.86, and specificity of 0.875. This point was not higher than that in healthy subjects, which appeared to be appropriate because any further increase in sensitivity would result in the detection of healthy subjects. If the expression of the CHN1 gene would have been higher in older age groups, the increase in expression could be attributed to age-related changes. However, the distribution of the *CHN1* gene in all 28 subjects surveyed and in group A showed no correlation between age and expression levels. Thus, the occurrence of cardiac complications in patients with Fabry disease can be predicted by examining *CHN1* expression.

Gene rankings based on significant differences among the A, B, and C three groups and gene rankings based on the strength of the association in many items in the MRI parameters were identical for the top three genes. The gene expression of *ACVR1* (2nd place) and *MT1F* (3rd place) may be related; however, the sensitivity for CHN1 alone was greater and more specific than that for the statistical results for the three genes combined. Thus, If the prognosis of future cardiac complications, especially arrhythmia and myocardial fibrosis, can be predicted by examining the RNA expression of CHN1 alone, it will be possible to predict the optimal timing for the initiation of cardiovascular medical intervention and enzyme replacement therapy. It is also considered to be a factor that has a large influence.

In Fabry disease, in particular, there are many cases of sudden death due to fatal arrhythmia, and based on these predictions, appropriate interventions, such as prescribing antiarrhythmic drugs and implanting cardioverter defibrillator devices, can be timed better, thus improving patient prognosis and functionality. It is also likely to have a direct impact on the quality of life.

However, future studies with an larger cohort size of between 100 and 1000 participants are required to determine whether this prognostic predictor can be used in clinical settings. To date, many studies have already been published on the automatic segmentation of atria and ventricles from cardiac MRI data using AI [25–30] and quantitative analysis of cardiac function using this method [31,32]; it is anticipated that more detailed studies will be conducted in this field in the future.

To the best of our knowledge, this is the first study to successfully demonstrate quantitative analysis of cardiac function by automatic segmentation using AI and the radiomics of gene expression profiles of genetic disorders using cardiac MRI.

5. Conclusion

The *CHN1* gene encoding chimerin 1, which is involved in myocardial fibrosis, was identified as a possible predictor of cardiac complications in Fabry disease based on AI-based MRI and correlation analyses of the transcriptome of RNA extracted from the peripheral blood of patients with Fabry disease and healthy individuals.

Funding

This study was supported by MEXT KAKENHI (grant number: JP18992403).

Details of ethics approval

This study was approved by the Research Ethics Committee of the Jikei University School of Medicine, Tokyo, Japan.

Data sharing statement

The data are not publicly available because of privacy and ethical restrictions.

CRediT authorship contribution statement

Hiroshi Kobayashi: Writing – review & editing, Writing – original draft, Project administration, Investigation, Funding acquisition, Data curation, Conceptualization. **Norio Nakata:** Writing – review & editing, Visualization, Supervision, Software, Project administration, Methodology, Investigation, Data curation, Conceptualization. **Sayoko Izuka:** Resources, Methodology. **Kenichi Hongo:** Writing – review & editing, Supervision, Project administration. **Masako Nishikawa:** Writing – review & editing, Validation, Supervision, Project administration, Methodology, Investigation, Data curation.

Declaration of competing interest

HK received grants for industry-academia collaborations from Sumitomo Pharma Co., Ltd., JCR Pharmaceuticals Co., Ltd., and Amicus Therapeutics, Inc. Other authors do not have any conflicts of interest.

Acknowledgements

We thank Dr. Toshihide Kato (K. K. DNAFORM, Research Services) for facilitating data analysis. We would also like to thank the patients with Fabry disease and healthy volunteers who provided valuable samples.

Data availability

The data that has been used is confidential.

References

[1] R. Izhar, M. Borriello, A.L. Russa, et al., Fabry disease in women: genetic basis, available biomarkers, and clinical manifestations, Genes 15 (2024) 37, https://doi.org/10.3390/genes15010037.
[2] M. Salerno, B. Sharif, H. Arheden, A. Kumar, L. Axel, D. Li, S. Neubauer, Recent advances in cardiovascular magnetic resonance: Techniques and applications, Circ. Cardiovasc. Imaging 10 (2017), https://doi.org/10.1161/CIRCIMAGING.116.003951.
[3] G. Li, L. Li, Y. Li, Z. Qian, F. Wu, Y. He, H. Jiang, R. Li, D. Wang, Y. Zhai, Z. Wang, T. Jiang, J. Zhang, W. Zhang, An MRI radiomics approach to predict survival and tumor-infiltrating macrophages in gliomas, Brain 145 (2022) 1151–1161, https://doi.org/10.1093/brain/awab340.
[4] P. Lohmann, M. Kocher, J. Steger, N. Galldiks, Radiomics derived from amino-acid PET and conventional MRI in patients with high-grade gliomas, Q. J. Nucl. Med. Mol. Imaging 62 (2018) 272–280, https://doi.org/10.23736/S1824-4785.18.03095-9.
[5] A.C. Yeh, H. Li, Y. Zhu, J. Zhang, G. Khramtsova, K. Drukker, A. Edwards, S. McGregor, T. Yoshimatsu, Y. Zheng, Q. Niu, H. Abe, J. Mueller, S. Conzen, Y. Ji, M.L. Giger, O.I. Olopade, Radiogenomics of breast cancer using dynamic contrast enhanced MRI and gene expression profiling, Cancer Imaging 19 (2019) 48, https://doi.org/10.1186/s40644-019-0233-5.
[6] Z. Chen, Y.J. Tan, M.M. Lian, M. Tandiono, J.N. Foo, W.K. Lim, N. Kandiah, E. K. Tan, A.S.L. Ng, High diagnostic utility incorporating a targeted neurodegeneration gene panel with MRI brain diagnostic algorithms in patients with young-onset cognitive impairment with leukodystrophy, Front. Neurol. 12 (2021) 631407, https://doi.org/10.3389/fneur.2021.631407.
[7] V. Frewer, C.P. Gilchrist, S.E. Collins, K. Williams, M.L. Seal, R.J. Leventer, D. J. Amor, A systematic review of brain MRI findings in monogenic disorders strongly associated with autism spectrum disorder, J. Child Psychol. Psychiatry 62 (2021) 1339–1352, https://doi.org/10.1111/jcpp.13510.
[8] H. Li, R. Durbin, Fast and accurate short read alignment with burrows-wheeler transform, Bioinformatics 25 (2009) 1754–1760, https://doi.org/10.1093/bioinformatics/btp324.
[9] D. Kim, B. Langmead, S.L. Salzberg, HISAT: a fast spliced aligner with low memory requirements, Nat. Methods 12 (2015) 357–360, https://doi.org/10.1038/nmeth.3317.
[10] H. Ohmiya, M. Vitezic, M.C. Frith, M. Itoh, P. Carninci, A.R. Forrest, Y. Hayashizaki, T. Lassmann, FANTOM Consortium, RECLU: a pipeline to discover reproducible transcriptional start sites and their alternative regulation using capped analysis of gene expression (CAGE), BMC Genomics 15 (2014) 269, https://doi.org/10.1186/1471-2164-15-269.
[11] A.R. Quinlan, I.M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features, Bioinformatics 26 (2010) 841–842, https://doi.org/10.1093/bioinformatics/btq033.
[12] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014) 550, https://doi.org/10.1186/s13059-014-0550-8.
[13] C. Zotti, Z. Luo, O. Humbert, et al., GridNet with automatic shape prior registration for automatic MRI cardiac segmentation, in: Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th international workshop, Revised Selected Papers 8, STACOM 2017. Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10–14, 2017, Springer International Publishing, 2018, pp. 73–81.
[14] N. Painchaud, Y. Skandarani, T. Judge, O. Bernard, A. Lalande, P.M. Jodoin, Cardiac segmentation with strong anatomical guarantees, IEEE Trans. Med. Imaging 39 (2020) 3703–3713, https://doi.org/10.1109/TMI.2020.3003240.
[15] A.A. Malayeri, W.C. Johnson, R. Macedo, J. Bathon, J.A. Lima, D.A. Bluemke, Cardiac cine MRI: quantification of the relationship between fast gradient echo and steady-state free precession for determination of myocardial mass and volumes, J. Magn. Reson. Imaging 28 (2008) 60–66, https://doi.org/10.1002/jmri.21405.
[16] M.D. Cerqueira, N.J. Weissman, V. Dilsizian, A.K. Jacobs, S. Kaul, W.K. Laskey, D. J. Pennell, J.A. Rumberger, T. Ryan, M.S. Verani, American Heart Association writing group on myocardial segmentation and registration for cardiac imaging standardized myocardial segmentation and nomenclature for tomographic imaging of the heart, Circulation 105 (2002) 539–542, https://doi.org/10.1161/hc0402.102975.
[17] M.T.P. Le, N. Zarinabad, T. D'Angelo, I. Mia, R. Heinke, T.J. Vogl, A. Zeiher, E. Nagel, V.O. Puntmann, Sub-segmental quantification of single (stress)-pass perfusion CMR improves the diagnostic accuracy for detection of obstructive coronary artery disease, J. Cardiovasc. Magn. Reson. 22 (2020) 14, https://doi.org/10.1186/s12968-020-0600-1.
[18] A. Nojiri, I. Anan, S. Morimoto, M. Kawai, T. Sakuma, M. Kobayashi, H. Kobayashi, H. Ida, T. Ohashi, Y. Eto, T. Shibata, M. Yoshimura, K. Hongo, Clinical findings of gadolinium-enhanced cardiac magnetic resonance in Fabry patients, J. Cardiol. 75 (2020) 27–33, https://doi.org/10.1016/j.jjcc.2019.09.002.
[19] E. Yoshida, Y. Terao, N. Hayashi, et al., Promoter-level transcriptome in primary lesions of endometrial cancer identified biomarkers associated with lymph node metastasis, Sci. Rep. 7 (1) (2017) 14160. https://www.nature.com/articles/s41598-017-14418-5.
[20] E.R. DeLong, D.M. DeLong, D.L. Clarke-Pearson, Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach, Biometrics 44 (1988) 837–845, https://doi.org/10.2307/2531595.
[21] N. Miyake, J. Chilton, M. Psatha, L. Cheng, C. Andrews, W.M. Chan, K. Law, M. Crosier, S. Lindsay, M. Cheung, J. Allen, N.J. Gutowski, S. Ellard, E. Young, A. Iannaccone, B. Appukuttan, J.T. Stout, S. Christiansen, M.L. Ciccarelli, A. Baldi, M. Campioni, J.C. Zenteno, D. Davenport, L.E. Mariani, M. Sahin, S. Guthrie, E. C. Engle, Human CHN1 mutations hyperactive alpha-2-chimerin and cause Duane's reaction syndrome, Science 321 (2008) 839–843, https://doi.org/10.1126/science.1156121.
[22] T.C. Zhou, W.H. Duan, X.L. Fu, et al., Identification of novel CHN1 p variant in a large Han Chinese family with congenital Duane retraction syndrome, Sci. Rep. 16225 (2020) 10, https://doi.org/10.1038/s41598-020-73190-1.
[23] K. Lorenz, J.P. Schmitt, E.M. Schmitteckert, M.J. Lohse, A new type of ERK1/2 autophosphorylation causes cardiac hypertrophy, Nat. Med. 15 (2009) 75–83, https://doi.org/10.1038/nm.1893.
[24] W. Parichatikanond, R. Duangrat, S. Mangmool, vGaq protein-biased ligand of angiotensin II type 1 receptor mediates myofibroblast differentiation through TGF-beta1/ERK axis in human cardiac fibroblasts, Eur. J. Pharmacol. 951 (2023) 175780, https://doi.org/10.1016/j.ejphar.2023.175780.
[25] O. Bernard, A. Lalande, C. Zotti, F. Cervenansky, X. Yang, P.A. Heng, I. Cetin, K. Lekadir, O. Camara, M.A. Gonzalez Ballester, G. Sanroma, S. Napel, S. Petersen, G. Tziritas, E. Grinias, M. Khened, V.A. Kollerathu, G. Krishnamurthi, M.M. Rohe, X. Pennec, M. Sermesant, F. Isensee, P. Jager, K.H. Maier-Hein, P.M. Full, I. Wolf, S. Engelhardt, C.F. Baumgartner, L.M. Koch, J.M. Wolterink, I. Isgum, Y. Jang, Y. Hong, J. Patravali, S. Jain, O. Humbert, P.M. Jodoin, Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE Trans. Med. Imaging 37 (2018) 2514–2525, https://doi.org/10.1109/TMI.2018.2837502.
[26] Z. Chen, J. Bai, Y. Lu, Dilated convolution network with edge fusion block and directional feature maps for cardiac MRI segmentation, Front. Physiol. 14 (2023) 1027076, https://doi.org/10.3389/fphys.2023.1027076.
[27] F. Ahmad, W. Hou, J. Xiong, Z. Xia, Fully automated cardiac MRI segmentation using dilated residual network, Med. Phys. 50 (2023) 2162–2175, https://doi.org/10.1002/mp.16108.
[28] C. Fan, Q. Su, Z. Xiao, H. Su, A. Hou, B. Luan, ViT-FRD: a vision transformer model for cardiac MRI image segmentation based on feature recombination distillation, IEEE Access. 11 (2023) 129763–129772, https://doi.org/10.1109/ACCESS.2023.3302522.

[29] C. Martin-Isla, V.M. Campello, C. Izquierdo, K. Kushibar, C. Sendra-Balcells, P. Gkontra, A. Sojoudi, M.J. Fulton, T.W. Arega, K. Punithakumar, L. Li, X. Sun, Y. Al Khalil, D. Liu, S. Jabbar, S. Queiros, F. Galati, M. Mazher, Z. Gao, M. Beetz, L. Tautz, C. Galazis, M. Varela, M. Hullebrand, V. Grau, X. Zhuang, D. Puig, M. A. Zuluaga, H. Mohy-Ud-Din, D. Metaxas, M. Breeuwer, R.J. van der Geest, M. Noga, S. Bricq, M.E. Rentschler, A. Guala, S.E. Petersen, S. Escalera, J.F. R. Palomares, K. Lekadir, Deep learning segmentation of the right ventricle in cardiac MRI: the M&Ms challenge, IEEE J. Biomed. Health Inform. 27 (2023) 3302–3313, https://doi.org/10.1109/JBHI.2023.3267857.

[30] Y. Li, Z. Liu, Q. Lai, S. Li, Y. Guo, Y. Wang, Z. Dai, J. Huang, ESA-UNet for assisted diagnosis of cardiac magnetic resonance image based on the semantic segmentation of the heart, Front. Cardiovasc. Med. 9 (2022) 1012450, https://doi.org/10.3389/fcvm.2022.1012450.

[31] F. Odille, A. Bustin, S. Liu, B. Chen, P.A. Vuissoz, J. Felblinger, L. Bonnemains, Isotropic 3D cardiac cine MRI allows efficient sparse segmentation strategies based on 3D surface reconstruction, Magn. Reson. Med. 79 (2018) 2665–2675, https://doi.org/10.1002/mrm.26923.

[32] S. Singh, S. Kaushik, R. Vats, A. Jain, N. Thakur, Right ventricle MRI image segmentation of heart, in: IEEE 5th International Conference for Convergence in Technology, Bombay, 2019, pp. 1–4, https://doi.org/10.1109/I2CT45611.2019.9033921.