Next generation sequencing reveals *NRAP* as a candidate gene for hypertrophic cardiomyopathy in elderly patients

Sharma Ankit[1], Rakesh Koranchery[2], Ranjith Rajendran[2], K S Mohanan[2], Jayaprakash Shenthar[3], Perundurai S Dhandapany[1,4,5]

[1]Centre for Cardiovascular Biology and Disease, Institute for Stem Cell Science and Regenerative Medicine, Bangalore, India.

[2]Department of Cardiology, Government Medical College, Kozhikode, Kerala- 673008, India

[3]Sri Jayadeva Institute of Cardiovascular Sciences and Research, Jayanagar, Bengaluru, Karnataka-560069, India

[4]The Knight Cardiovascular Institute, [5]Departments of Medicine, Molecular and Medical Genetics, Oregon Health and Science University, Portland, Oregon

Correspondence to:

Perundurai S Dhandapany, PhD

[1]Centre for Cardiovascular Biology and Disease, Institute for Stem Cell Science and Regenerative Medicine, GKVK campus, [4,5]Department of Molecular and Medical Genetics, Knight Cardiovascular Institute (KCVI), Oregon Health & Sciences University, 3181 SW Sam Jackson Park Road L103, Portland, Oregon, U.S.A. 97239-3098

E-mail: dhan@instem.res.in or perundur@ohsu.edu

CONFLICT OF INTEREST:

We declare no conflict of interest

ABSTRACT:

Hypertrophic cardiomyopathy (HCM) is a heterogenous heart muscle disease predominantly caused by sarcomeric protein encoding genes. However, the cause for a significant number of elderly patients remains unclear. Here, we performed whole-exome sequencing in a South Indian family with an elderly HCM proband. We identified a heterozygous missense variant in the Nebulin-Related-Anchoring Protein encoding gene *NRAP* (NM_001261463, c.1259A>G,

p.Y420C) in the proband. NRAP is a multi-domain scaffolding protein involved in cardiac muscle thin filament assembly, myofibril and actin cytoskeleton organization. The respective NRAP (p.Y420C) is predicted to be pathological by *in -silico* analysis and might be potentially influencing its functions. Targeted re-sequencing in an independent cohort resulted in identification of the same amino acid change in an unrelated eighty-six years old patient. The protein-protein interactions (PPI) network analysis revealed NRAP is strongly associated with other known elderly and late-onset HCM genes/proteins. In line with these data, both the study patients are late-onset in nature. Our study for the first time reveals the association of *NRAP* in the elderly and late-onset HCM patients and further expands the genotypic-phenotypic spectrum of HCM.

INTRODUCTION:

HCM is a disease of heart muscles, which is characterized by abnormally muscle thickening finally leading to systolic dysfunction (1, 2). More often, mutations in several sarcomere genes are primary cause of familial and non-familial HCM. Clinical expression of HCM starts generally in the early age (~20 years of age) and progress quickly in middle age of the patients (~40 years of age). In contrast, in the elderly patients the symptoms remain silent but develop after an average age of 66 years (3). Till date, little is known about the elderly and late-onset HCM genes and only a few sarcomeric genes such as Myosin binding protein C3 (*MYBPC3)*, cardiac troponin T (*TNNT2*), cardiac troponin I (*TNNI1* and *TNNI3*), myosin light chain 3 (*MYL3*), alpha-cardiac myosin heavy chain 6 (*MYH6*) and beta-cardiac myosin heavy chain 7 (*MYH7*) are shown to be associated with elderly and late-onset HCM (4-9). Here, we analyzed a clinically well characterized Indian family with a history of late-onset HCM and performed whole-exome sequencing (WES) and analysis. We identified a novel missense variant in *NRAP* (NM_001261463, c.1259A>G, p.Y420C) in the proband compared to the healthy familial controls. Targeted re-sequencing in an independent cohort resulted in identification of a separate elderly patient with the same gene variant. Computational analysis using PPI and pathway analysis also suggest strong association of *NRAP* with known elderly and late-onset HCM genes.

MATERIALS AND METHODS:

Study participants:

Blood samples from study participants and their clinical reports were obtained from Government Medical College, Kozhikode and Sri Jayadeva Institute of Cardiovascular Sciences and Research, Bangalore, India. We analyzed five subjects from a single HCM family with informed consents. Echocardiograms (Echo) were performed on all the study participants. The study was conducted in accordance with the 'Declaration of Helsinki' and was approved by the respective hospitals and Institutional ethics committees. Genomic DNA was extracted from peripheral blood samples using standard procedures (phenol-chloroform method) for molecular analysis (5).

Genetic analysis:

To identify potential causal genetic variant in the patient, whole exome sequencing (WES) was performed for a proband (2II) who is negative for sarcomeric gene mutations and an unaffected available healthy family member (2IV) (**Figure 1A**). Sequencing was performed using the respective service providers including institutional sequencing facility after de-identifying the patient details. For each exome paired end 100 base pair reads (100X coverage) was employed and 6 Gb data was obtained. Exome analysis and variant calling were performed using in-house custom pipeline as outlined in the **Figure 1B**. Low quality reads were filtered and adapters were removed using Trimmomatic. Exomes were first mapped on to the human reference genome (GRCh38) using Burrows-Wheeler aligner (BWA-MEM). Variant calling was performed using HaplotypeCaller algorithm from Genome Analysis Tool Kit (GATK) (software.broadinstitute.org/gatk/). Variants were annotated using web interface of ANNOVAR software (wannovar.wglab.org/). To maximize our disease variant findings, we included ultra-rare (MAF ≤ 0.01%) and deleterious variants that are present in Genome Aggregation Database (gnomAD), 1000 Genomes Project, Korean Exome Information Database (KoEXID), Exome Aggregation Consortium (ExAC) and in-house database of healthy aging Indian exomes (www.idhans.org). Deleterious variants affecting potential protein functions, were based on Polymorphism Phenotyping v2 (PolyPhen2) (http://genetics.bwh.harvard.edu/pph2/) and Sorting Intolerant From Tolerant (SIFT) (https://sift.bii.a-star.edu.sg/www/SIFT_dbSNP.html) methods. Further, cardiovascular-related genes were filtered out using International Mouse Phenotyping Consortium (IMPC) (www.mousephenotype.org/) and Mouse Genome Informatics (MGI) (www.informatics.jax.org/) databases. Finally, a comprehensive literature survey was performed for understand the role of the particular genes in HCM. To replicate the association of identified

new gene variant, we performed targeted re-sequencing in the remaining unaffected family members and an independent cohort comprising 288 unrelated HCM patients.

Protein-protein interaction (PPI) analysis:

HCM-associated genes were extracted (10) and protein-protein interaction network was constructed with identified potential gene (*NRAP*) from patient's exome data. PPI network using STRING database (based on protein coexpression, experimentally determined interaction and various database annotation) and pathway analysis were performed using ClugoVs from Cytoscape 3.6.0 (www.cytoscape.org/). Only significant pathways with ≤0.005 were considered for pathway enrichment. Conservation alignment was performed using Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/).

RESULTS

Exome sequencing and analysis: The WES resulted in an average of 702.65 million clean reads in 2 subjects (2II and 2IV). We obtained 75.6 thousand single nucleotide polymorphisms (SNPs) and 9.1 thousand insertion/deletion/frameshift variants (InDels) in the proband, and 79.9 thousand SNPs and 9.2 thousand InDels in the healthy family member. By using the variant calling and filtering pipeline outlined in the methods section, we identified a missense heterozygous mutation in the *NRAP* as a potential candidate in the proband (**Figure 1B**). The *NRAP* mutation (c.1259A>G) results in the amino acid change from tyrosine to cysteine at 420th position (p.Y420C) in the respective protein. Notably, NRAP Y420 residue is highly conserved among different species (**Figure 1C**). *NRAP* variant is absent in various existing exomes and genomes databases including 1000 Genomes Project, KoEXID, ExAC, ESP and in-house Indian exomes database. However, it is present in the gnomAD with an ultra-rare frequency (MAF= 0.000003%, 1 in 251232 alleles). Notably, gnomAD comprises various exomes of heart failure cases. Sanger re-sequencing the particular *NRAP* exon in the remaining unaffected family members and healthy aging Indians above 80 years old (www.idhans.org) confirmed the absence of the Y420 residue variant further strengthening its pathological role in elderly patients.

Targeted sequencing and analysis: We analyzed additional 288 unrelated HCM patients (age: 51.66±0.47 years, Sex: 60.7% male, Left ventricular internal diameter end diastole (LVIDD): 40.9±4.75 mm, Left ventricular internal diameter end systole (LVIDS): 24.77±4.76 mm,

interventricular septum thickness (IVS): 18.81±3.90 mm and left ventricular ejection fraction (LVEF): 54.67±5.36 %), and observed the Y420 residue change in an eighty-six years old patient (**Figure 1D**). The variant was absent in 320 population stratified South Indian healthy individuals from the same ethnicity and geographical region. The baseline details of the control subjects are described previously by us (5). It was also absent in the 94 healthy ageing Indians above 80 years old without any medical conditions (www.idhans.org).

Genotype-phenotype correlations: Patient (2II) was diagnosed with HCM at the age of 61 with LVIDD and LVIDS being 44 mm and 24 mm respectively, LVEF was 68% and IVS was 22 mm. For an unrelated HCM patient in the replication cohort, the age of onset was 70 years with the respective echo parameters: LVIDD: 43 mm, LVIDS: 29 mm, LVEF: 60% and IVS: 21 mm.

PPI network analysis:

We performed PPI network analysis between *NRAP* and known HCM-associated genes to understand the potential molecular connection of NRAP with HCM. Network analysis revealed *NRAP* is strongly connected with HCM-associated genes including six known elderly and late-onset HCM genes as highlighted in the **Figure 2A**. In addition, the pathway enrichment analysis also showed a significant link between NRAP and HCM pathways including sarcomere organization, cardiac muscle cell development, cardiac myofibril assembly, striated muscle myosin thick filament, muscle alpha-actinin binding and cellular homeostasis (**Figure 2A**). Notably, the Y420 residue is located in the simple repeat domain of the NRAP (**Figure 2B**), which is known to interact with HCM-associated proteins α-actinin, actin, muscle LIM protein (MLP) and Kelch-like family member 41 (KLHL41).

DISCUSSION:

Here, we provide first evidence that a *NRAP* missense mutation is associated with elderly and late-onset HCM. NRAP is crucial for cardiac muscle thin filament assembly, actin cytoskeleton organization, and myofibril assembly in cardiomyocytes. NRAP is a multidomain scaffolding protein with a N-terminal LIM domain; a linker region with single repeating modules; and a C-terminal super repeats domain (11). During the myofibril assembly NRAP binds to the membrane helping cardiac muscle contraction and relaxation through complex containing β-integrins, talin and vinculin. NRAP super repeats domain and the vinculin tail help in positioning

the Z-line components (MLP) in single repeating domain. These super repeats are expected to assign directionality to the actin filaments where α-actinin, binds at Z-lines of simple repeat domain of NRAP to imparts its functions (12).

The mutated residue that we observed in the patients (Y420) is located in the single repeat domain, which contains the binding sites for α-actinin, actin, MLP and KLHL41 (**Figure2B**) (13-15). Previous studies have shown the association of LIM, single repeats and super repeats domains of NRAP are important for different types of cardiomyopathy. For example, frame-shift mutation at 448 amino acid position in the simple repeat domain and 1502 amino acid position in the super repeat domain cause loss of function in NRAP and results in childhood Dilated Cardiomyopathy (DCM) (16). Furthermore, compound heterozygous missense mutations leading to amino acid changes at the position 282 from alanine to threonine, 519 from asparagine to isoleucine and 1225 from alanine to valine have been found in a single patient of myofibril myopathy (MFM) (17). A recent study has shown the pathologically increased levels of NRAP contributing to muscle dysfunction in Nemaline myopathy (NM) by interacting KLHL41. Of note, NM has been shown to have strong association with HCM (18, 19). In addition, overexpression of NRAP is found to be associated with ventricular hypertrophy and dysfunction in a mouse model (20). Thus, loss of function NRAP mutations are associated with DCM whereas overexpression (or) gain of function leads to cardiac hypertrophy. So far, all the *NRAP* variants reported (**Figure 2C**) are loss-of- function mutations with respective to early-onset diseases. In contrast, our study for the first time suggests a missense amino acid change in the *NRAP* at 420th residue leading to elderly and late-onset HCM. We presume, further analysis of *NRAP* mutations in the larger cohorts, will establish the causal link for *NRAP* in the HCM.

ACKNOWLEDGEMENTS:

This study is funded by American Heart Association-Scientist Development Grant (15SDG23250005) and The Wellcome Trust/DBT India Alliance (IA) Intermediate Grant (IA/I/16/1/502367). We thank all the patients and families for the participation in this study. We also thank Ms. Paulami Dey for her kind help in Sanger sequencing,

DATA ACCESSIBILITY:

Supporting data is available upon reasonable request

References:

1. Fay WP, Taliercio CP, Ilstrup DM, Tajik AJ, Gersh BJ. Natural history of hypertrophic cardiomyopathy in the elderly. Journal of the American College of Cardiology. 1990;16(4):821-6.
2. Kubo T, Kitaoka H, Okawa M, Nishinaga M, Doi YL. Hypertrophic cardiomyopathy in the elderly. Geriatrics & gerontology international. 2010;10(1):9-16.
3. Lewis JF, Maron BJ. Elderly patients with hypertrophic cardiomyopathy: a subset with distinctive left ventricular morphology and progressive clinical course late in life. Journal of the American College of Cardiology. 1989;13(1):36-45.
4. Niimura H, Patton KK, McKenna WJ, Soults J, Maron BJ, Seidman J, et al. Sarcomere protein gene mutations in hypertrophic cardiomyopathy of the elderly. Circulation. 2002;105(4):446-51.
5. Dhandapany PS, Sadayappan S, Xue Y, Powell GT, Rani DS, Nallari P, et al. A common MYBPC3 (cardiac myosin binding protein C) variant associated with cardiomyopathies in South Asia. Nature genetics. 2009;41(2):187.
6. Anan R, Niimura H, Takenaka T, Hamasaki S, Tei C. Mutations in the genes for sarcomeric proteins in Japanese patients with onset sporadic hypertrophic cardiomyopathy after age 40 years. The American journal of cardiology. 2007;99(12):1750-4.
7. Choi JO, Yu CW, Chun Nah J, Rang Park J, Lee BS, Jeong Choi Y, et al. Long-Term Outcome of 4 Korean Families With Hypertrophic Cardiomyopathy Caused by 4 Different Mutations. Clinical cardiology. 2010;33(7):430-8.
8. Kubo T, Kitaoka H, Okawa M, Baba Y, Hirota T, Hayato K, et al. Genetic screening and double mutation in Japanese patients with hypertrophic cardiomyopathy. Circulation Journal. 2011;75(11):2654-9.
9. Tian T, Liu Y, Zhou X, Song L. Progress in the molecular genetics of hypertrophic cardiomyopathy: a mini-review. Gerontology. 2013;59(3):199-205.
10. Xu J, Li Z, Ren X, Dong M, Li J, Shi X, et al. Investigation of pathogenic genes in Chinese sporadic hypertrophic cardiomyopathy patients by whole exome sequencing. Scientific reports. 2015;5:16609.
11. Luo G, Zhang JQ, Nguyen TP, Herrera AH, Paterson B, Horowits R. Complete cDNA sequence and tissue localization of N-RAP, a novel nebulin-related protein of striated muscle. Cell motility and the cytoskeleton. 1997;38(1):75-90.
12. Carroll SL, Herrera AH, Horowits R. Targeting and functional role of N-RAP, a nebulin-related LIM protein, during myofibril assembly in cultured chick cardiomyocytes. Journal of cell science. 2001;114(23):4229-38.
13. Luo G, Herrera AH, Horowits R. Molecular interactions of N-RAP, a nebulin-related protein of striated muscle myotendon junctions and intercalated disks. Biochemistry. 1999;38(19):6135-43.
14. Lu S, Carroll SL, Herrera AH, Ozanne B, Horowits R. New N-RAP-binding partners α-actinin, filamin and Krp1 detected by yeast two-hybrid screening: implications for myofibril assembly. Journal of cell science. 2003;116(11):2169-78.
15. Ehler E, Horowits R, Zuppinger C, Price RL, Perriard E, Leu M, et al. Alterations at the intercalated disk associated with the absence of muscle LIM protein. The Journal of cell biology. 2001;153(4):763-72.
16. Vasilescu C, Ojala TH, Brilhante V, Ojanen S, Hinterding HM, Palin E, et al. Genetic Basis of Severe Childhood-OnsetáCardiomyopathies. Journal of the American College of Cardiology. 2018;72(19):2324-38.
17. D'Avila F, Meregalli M, Lupoli S, Barcella M, Orro A, De Santis F, et al. Exome sequencing identifies variants in two genes encoding the LIM-proteins NRAP and FHL1 in an Italian patient with BAG3 myofibrillar myopathy. Journal of muscle research and cell motility. 2016;37(3):101-15.
18. Mir A, Lemler M, Ramaciotti C, Blalock S, Ikemba C. Hypertrophic cardiomyopathy in a neonate associated with nemaline myopathy. Congenital heart disease. 2012;7(4):E37-E41.

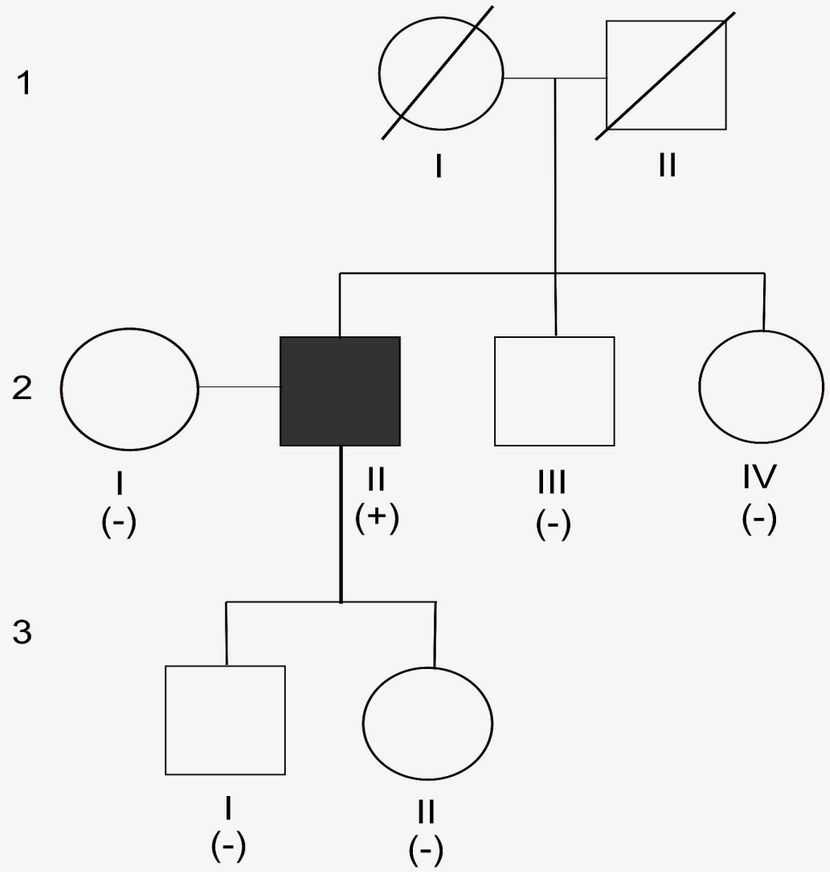
19. Jirka C, Pak JH, Grosgogeat CA, Marchetti MM, Gupta VA. Dysregulation of NRAP degradation by KLHL41 contributes to pathophysiology in Nemaline Myopathy. bioRxiv. 2018:487454.
20. Lu S, Crawford GL, Dore J, Anderson SA, DesPres D, Horowits R. Cardiac-specific NRAP overexpression causes right ventricular dysfunction in mice. Experimental cell research. 2011;317(8):1226-37.

Figure Legends:

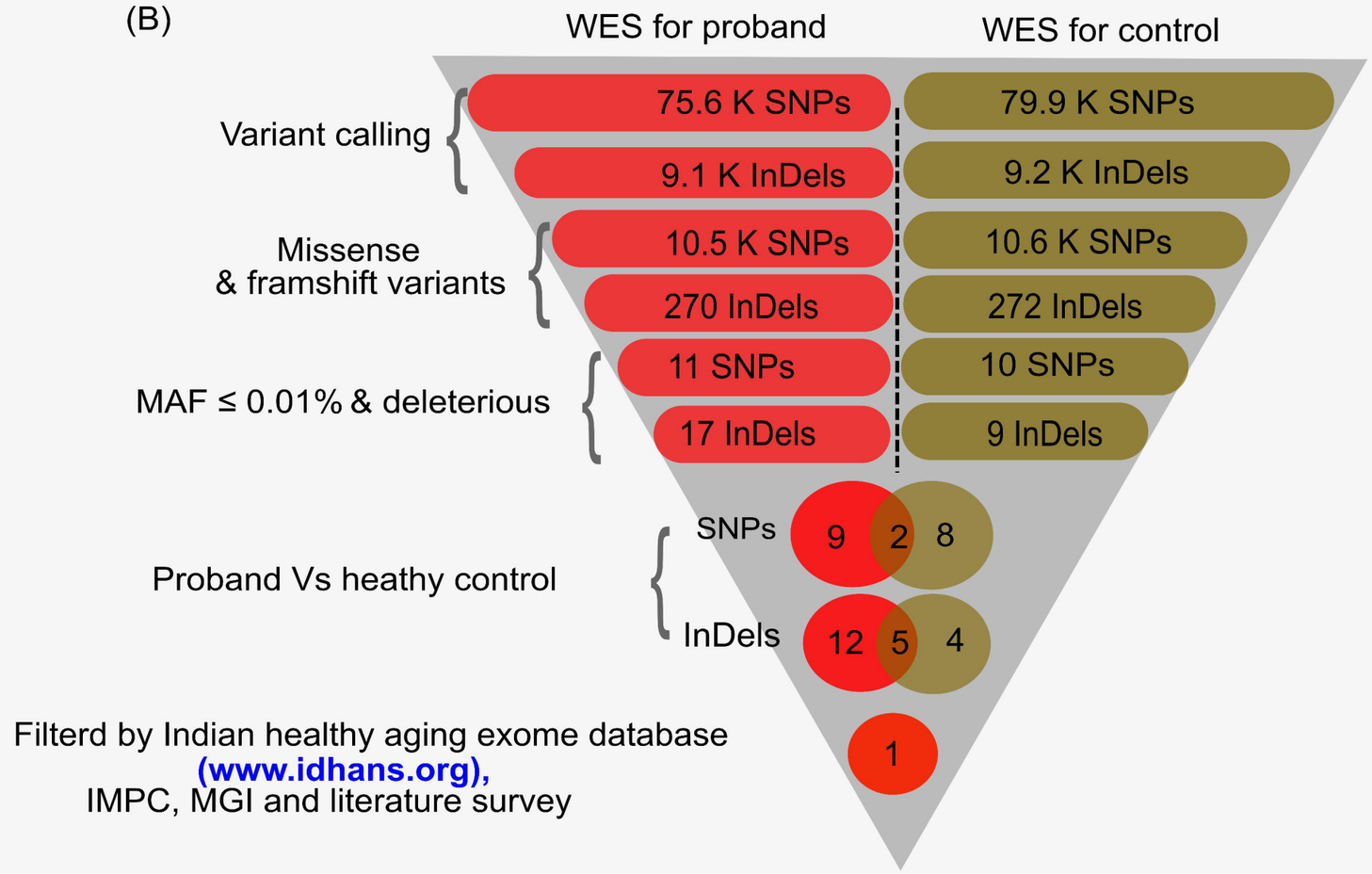
Figure 1: (A) Pedigree of the HCM family, square represents the male and circle represents the female subjects, symbols with a slash represent member**s** who had died. Solid color represents the proband, plus (+) indicates the presence of NRAP Y420C, and minus (-) indicates the absence of the respective amino acid change. (B) Variant filtering from WES analysis of HCM patient and an unaffected subject from the same family. (C) Alignment of NRAP in different species showing the conservation of Y420 region. (D) Sequence analysis showing NRAP Y420C in the Proband 2II and his family members and an unrelated HCM patient.

Figure 2: (A) PPI network analysis between NRAP and HCM-associated genes/proteins, and pathway enrichment analysis. NRAP shows strong association with elderly and late-onset HCM genes (highlighted in red color) (B) Schematic representation of NRAP domains and location of disease-associated amino acid changes from previous and current studies. (C) NRAP amino acid changes and associated diseases

(A)



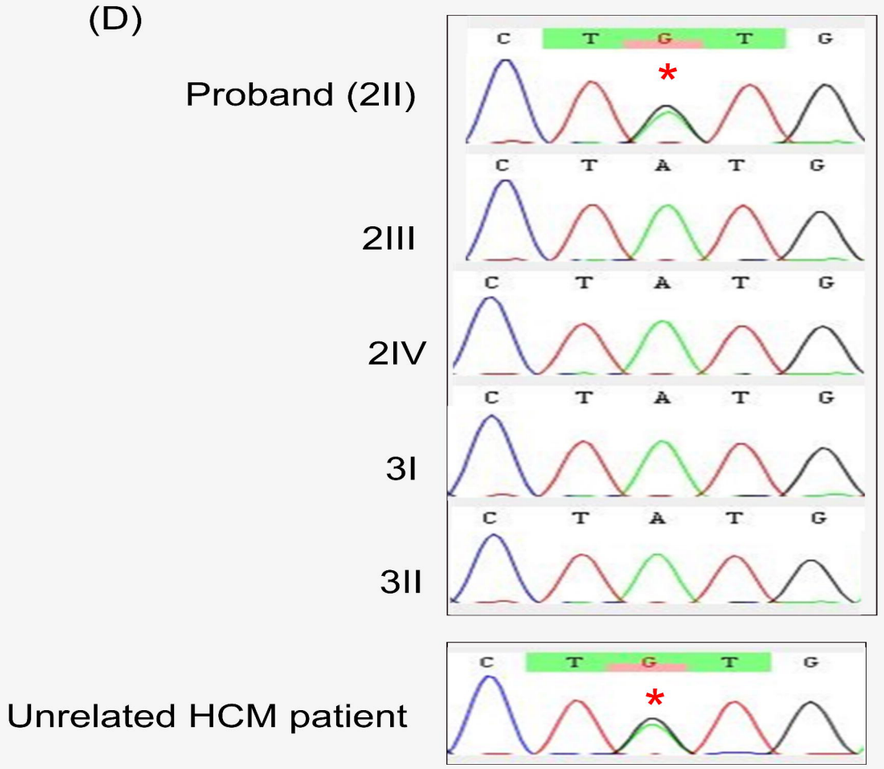
(B)



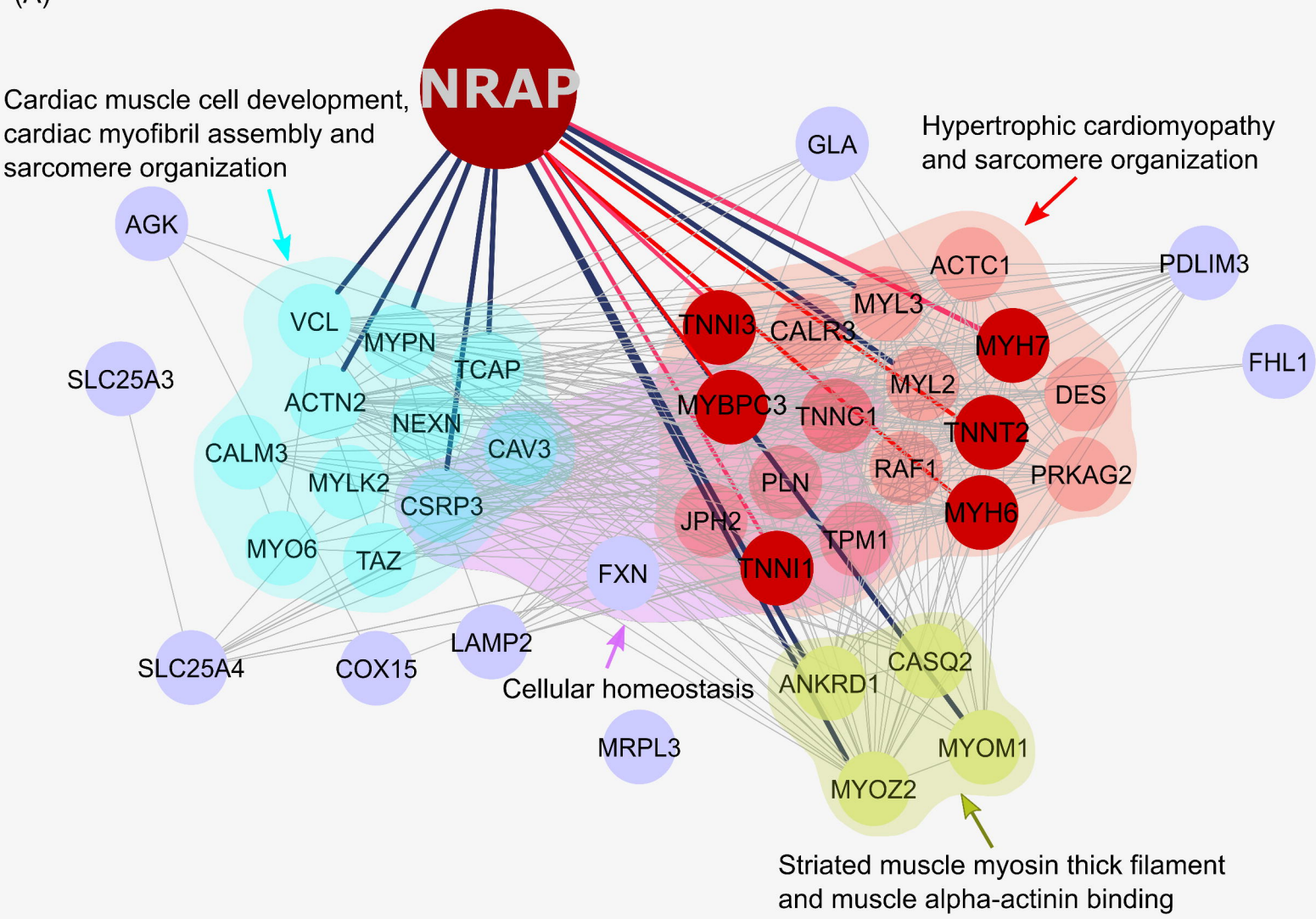
(C)

Homo sapiens (NP_932326.2)	MRGRYEGVG
Pan troglodytes (XP_009457540.3)	MRGRYEGVG
Macaca mulatta (XP_015003698.1)	MRGRYEGVG
Canis lupus (XP_851973.1)	MRGHYEGVG
Equus caballus (XP_001495958.2)	MRGRYEGVG
Bos taurus (DAA14731.1)	MRGRYEGVG
Rattus norvegicus (NP_001107215.1)	LRGHYEGVG

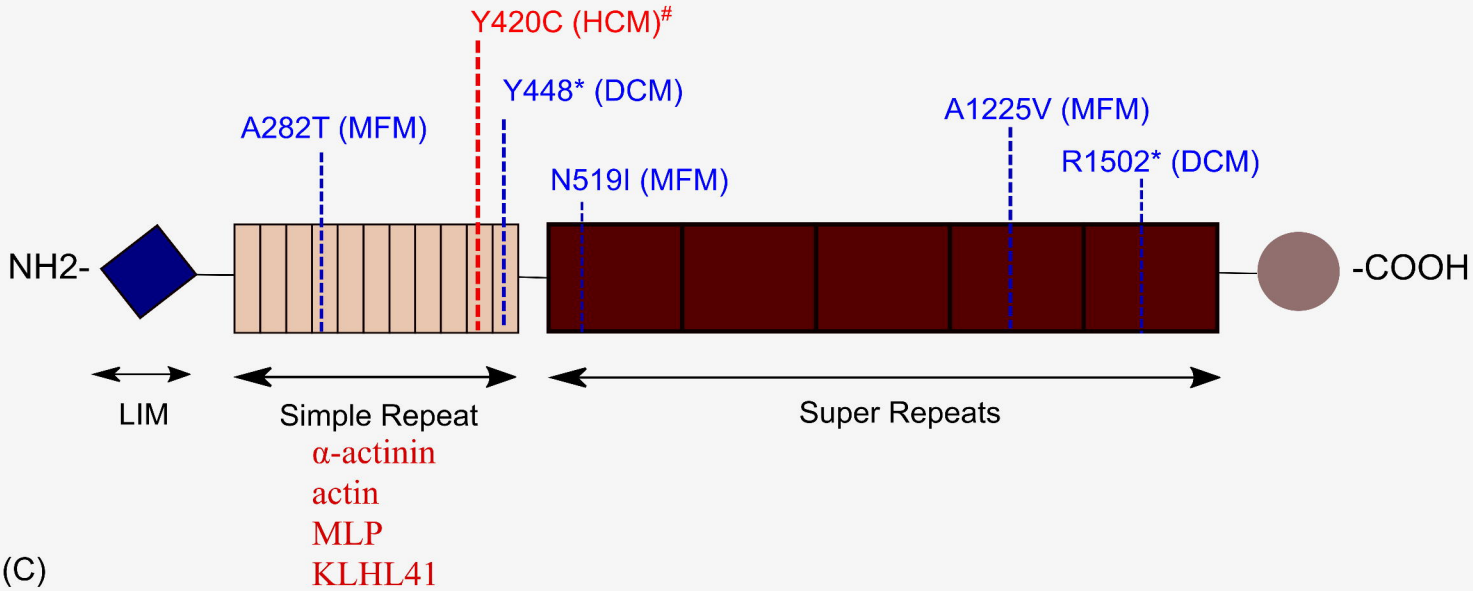
(D)



(A)



(B)



(C)

NRAP amino acid change(s)	Age	Phenotypes	Reference
p.Y448*d	3.5 years	Pediatric DCM	Vasilescu C et al. 2018
p.R1502*d	26 years	Adult DCM	Truszkowska GT et al. 2017
p.A288T, p.N519I, p.A1225V (Compound heterozygotes)	26 years	MFM	Francesca D'Avila et al. 2016
p.Y420C	70 and 86 years	Elderly HCM	Present study