Author’s response to reviews

Title: Non-Familial Cardiomyopathies in Lebanon: Exome Sequencing Results for Five Idiopathic Cases

Authors:

Marwan Refaat (mr48@aub.edu.lb)
Sylvana Hassanieh (shassanieh@mgh.harvard.edu)
jad ballout (jad.ballout1@gmail.com)
Patrick Zakka (pzakka@gmail.com)
Mostafa Hotait (mah58@mail.aub.edu)
Fadi Bitar (fbitar@aub.edu.lb)
Mariam Arabi (ma81@aub.edu.lb)
Samir Arnaout (sarnaout@aub.edu.lb)
Hadi Skouri (hs13@aub.edu.lb)
Antoine Abchee (aa14@aub.edu.lb)
Bernard Abi-Saleh (ba47@aub.edu.lb)
Maurice Khoury (mk04@aub.edu.lb)
Andreas Massouras (andreas.massouras@saphetor.com)
Georges Nemer (gn08@aub.edu.lb)

Version: 1 Date: 06 Sep 2018

Author’s response to reviews:

POINT-BY-POINT RESPONSE

We thank the reviewers for their positive feedback and constructive comments and critiques and provide a point-by-point response to the comments and an explanation of how they were addressed in the revised manuscript. We hope that the manuscript in its modified version finds its way for acceptance and publication in your esteemed journal.
Reviewer 1

Major concerns

1. Variant filtering process and results are not clear and sounds like some of the patients who had 'pathogenic' findings in one of the 85 genes didn't really get a full exome analysis.

All got exome sequencing, and we have now added a figure to explain the methodology and the results before filtering.

   a. Were variants called with a target-interval? If so, please specify

   No target-interval is included.

   b. A (supplementary) table should be added to show for 'each' sample the exact number of variants identified: 1) total, 2) within coding region (is it coding region or variants with coding changes in page 5 line 7? If only 8500 SNVs and 450 Indels were identified in the coding region, that's concerning as it's only ~half the number of variants typically identified in exome coding region), 3) with minor allele frequency <X% (for rare variants and 5% seem too high even for cardiomyopathy), 4) within 85 known genes

   This is now explained in the figure as well as methodology section. We thank the reviewer for pointing out to a mistake in the writing up describing the 8500 SNVs and the 450 indels in the coding region. In fact the 8500 SNVs are non-synonymous variants in the coding regions whereas the 450 indels are the ones in the coding regions. The whole section is re-written now with an additional supplementary figure 1.

   c. I would actually suggest that after filtering out common variants (>1%), authors count all 'homozygous', 'potential compound heterozygous' and 'heterozygous' variants before intersecting with the 85 known genes.

   This is what we did for the second round of filtering and we now explicitly said it in the text, and it is shown in the figure.

   d. I would also suggest for the heterozygous variants, keep the missense variants in genes with ExAC missense Z-score >3 in addition to the loss of function variants.

   This is what is displayed in all supplementary tables.

   e. Reporting all variants in 85 known genes in the tables seem unnecessary. Only the rare variants should be reported and should include more information like population allele frequency, clinvar/HGMD classification, ExAC missense/LoF Z-score/pLI, etc that could help the variant interpretation and classification. Variant classification base on the ACMG guideline can be added too.
We only reported the variants with a MAF<5% for the 85 genes. We adjusted now the tables to include only variants with a MAF<1%. Each variant is classified by an algorithm developed by Saphetor using the 5-class score suggested by the ACGS guidelines (this is added to the methods section). All tables are now reformatted.

2. Patients MR38 and MR22 seem to have additional features than cardiomyopathy and seem not such a great fit for this study. MR38 maybe since LMNA is associated with AD cardiomyopathy but MR22 has multiple additional features such as severe developmental delay.

This is the cohort we have seen under the IRB approved protocol. We have recruited additional patients, however due to budget limitations, we weren’t able to perform exome sequencing on them.

3. Based on what’s described the text, patient MR22 cannot be classified as 'molecularly diagnosed' with mutations in known genes

We agree, and the term “molecularly diagnosed” was removed.

4. Authors should describe more in detail why genetic diagnosis of these patient cohorts is needed and why they think exome sequencing is a preferred method to other tests (i.e. gene panel) and also mention what are some drawbacks of doing exome sequencing instead of panels.

Done in the Discussion part.

Minor concerns

1. "Next-generation exome sequencing" and "whole-exome sequencing" should be changed to "exome sequencing"

Done.

2. References needed for page 3 line 37

Done.

3. Was this study approved by the IRB? It should be mentioned in the patient selection paragraph
This is now added to the Material and Methods section as follows “The Institution Review Board (IRB) at the American University of Beirut approved the protocol of the study (Protocol Number: IM.MR.01). An informed consent was obtained from all participants in the study”.

4. The methods for DNA extraction and exome sequencing are too lengthy. Both used commercially available kits so there is no need to describe in detail how the methods work unless there were modifications made.

We removed unnecessary information as requested.

5. In page 4 line 43-44, which exome/genome databases were used and which in silico prediction scores were used? Add exact name and references.

It is now clarified. The Gnomad database is used and 13 predictive in silico tools found at varsome have been used. The Saphetor online tool includes all these, and this is now highlighted in the text and a link to their methodology is also added.

6. Page 5 first paragraph should either move to methods>data analysis paragraph or become a separate section under results

It is now moved to the material and methods section.

Reviewer: 2

General Comment: “The study is well design and technically performed. The conclusions are according to the results.”

General Answer: We thank the reviewer for the positive feedback. The following are our answers to the queries.

1.- ARVC or ARVD are term not currently used. Now, the disease is named ACM (Arrhytmogenic Cardiomyopathy).

Thank you, we corrected it the term throughout the manuscript.

2.- After exome analysis and posterior re-analysis of 86 genes, all exons of all genes were amplified? At least at 30x?

Yes we made sure that the coverage is >30X of all exons of all genes.
3.- All rare variants were confirmed by Sanger method?

Only relevant variants from the results’ section were confirmed by Sanger sequencing (this was added to the text).

4.- Exons not amplified at least at 30x, were amplified using Sanger?

As mentioned above, all were amplified more than 30X.

5.- Why not use global data of gnomeAD? It is more specific that ExAC or EVS.

This is now incorporated in the Saphetor new software version, and this is what we used.

6.- Concerning in silico prediction, only Mutation Taster was used? Nowadays, at least 3 to 5 in silico methods are required in order to support the prediction.

Indeed, 13 in silico predictive algorithms are now incorporated in the analysis by Saphetor/Varsome, and this is what has been used to predict significance (the link and description of Varsome used by Saphetor is now incorporated in the text).

7.- The classification of pathogenicity follows ACMG recommendations? This is the official classification in genetic field to date. A previous report of a variant do not suppose its pathogenic role.

We do agree.

8.- What about CNV (Copy Number Variants)? Do you analyse these kind of alterations? Any result?

Unfortunately, this is not included because we didn’t have as much matched controls in our study.

9.- Finally, the main limitation of the manuscript is family segregation. It is a crucial point in order to clarify the role of variants. Why not analysis of relatives?

Unfortunately, members of the families didn’t consent to do so.

Please, could you include all mentioned data in the manuscript?

All are tracked in the new version of the manuscript.