Title: Use of a Targeted, Combinatorial Next-Generation Sequencing Approach for the Study of Bicuspid Aortic Valve

Authors:

Elizabeth M Bonachea (elizabeth.bonachea@nationwidechildrens.org)
Gloria Zender (gloria.zender@nationwidechildrens.org)
Peter White (peter.white@nationwidechildrens.org)
Don Corsmeier (don.corsmeier@nationwidechildrens.org)
David Newsom (david.newsom@nationwidechildrens.org)
Sara Fitzgerald-Butt (Sara.Fitzgerald-Butt@nationwidechildrens.org)
Vidu Garg (vidu.garg@nationwidechildrens.org)
Kim L McBride (kim.mcbride@nationwidechildrens.org)

Version: 4
Date: 31 July 2014

Author's response to reviews: see over
Reviewer: Gordon Huggins

1. Two individuals were described as carrying de novo mutations. Does this mean the mutations had not previously been reported? Or that the mutations were present but not found in the parental DNA? Please clarify. Table 1 indicates that the two de novo mutations actually rs ID numbers and were identified in 1000G and EVS.

In this manuscript, we use de novo to indicate that the given variant is present in the proband and absent in both parents.

2. The cohort was for the most part non-Hispanic white in racial and ethnic background. Were the few non-whites more likely to be found to have a mutation? The concern is that the few non-whites might carry gene variants that are related to their ancestry and not related to having BAV.

A new table (Supplemental Table 2) has been added to the manuscript to provide further clinical details of the probands in whom significant variants were identified. This table includes ethnicity- only one of the 31 variants reported in Table 1 was present in a non-Caucasian.

3. The basis for selecting the candidate genes for sequencing is not described. Perhaps the identification of Wnt pathway genes carrying mutations is because the gene selection process was biased towards the Wnt pathway?

Further details have been provided in the methods and discussion. The comparison group used in the DAVID analysis was the targeted capture gene list, not the whole genome. Thus, even if more WNT pathway genes were included in the capture, the representation of WNT variants remains higher than expected based on the capture design.

4. The mutation discovery approach is called “gene” sequencing; however, it is not clear whether the sequencing approach is inclusive of introns and promoter regions or whether it is restricted to just the exons of the candidate genes. If just the exons were sequenced then that point needs to be made much clearer to the reader.

The sequencing was performed with a whole gene interval approach including flanking regions surrounding the genes of interest. However, analysis was confined to exons for the purpose of this manuscript.

5. Would this approach be likely to discover a copy-number variation within the candidate genes?
The methodology presented here cannot be used to identify copy-number variation.

6. The methods section indicates that parental DNA was also available—was the parental DNA also screened? Or was a mutation found in a proband also confirmed in the parents?

The manuscript Methods section, subheading Sequencing, has been amended to reflect that parental DNA was only screened for those significant variants identified by next generation sequencing.

7. References are used inconsistently. For example the first sentence of the Results: “We studied a previously described cohort of 78...” that sentence must reference the manuscript that first reported the cohort in order for the reader to then go and learn more about the cohort. Please review closely and provide references throughout the manuscript where appropriate.

The manuscript Results section, subheading Identification of sequence variants, has been amended to refer appropriately to the original report of the study cohort.

8. Since there is no replication cohort these findings should be clearly described as “hypothesis generating”

The discussion section has been amended to reflect the nature of this data as hypothesis-generating and the need for further testing in replication cohorts.

9. A table describing the clinical characteristics of the patients found to carry a mutation would add value. Currently there is no description of the cohort within this manuscript.

A new table (Supplemental Table 2) has been added to the manuscript to describe the clinical characteristics of the probands in whom significant variants were identified. Characteristics described in this new table include cardiac phenotype, BAV cusp fusion phenotype, gender, ethnicity, and presence of known familial congenital heart disease.
**Reviewer:** Maria Grazia Andreassi

**How was the ascertainment of BAV confirmed in each patient? How many patients had a family history of BAV? Please provide more details on the clinical characteristics of study population.**

Information regarding ascertainment of BAV, family history of BAV, and other clinical characteristics of the study cohort can be found in the Methods, under the subheading Study Population.

“Each subject had undergone clinical echocardiography with images sufficient to identify associated cardiac malformations and aortic valve cusp fusion morphology (Table 2). Fifty of the 78 subjects (64%) had isolated BAV while the remainder had BAV-CoA. Forty-six subjects (59%) had R-L cusp fusion, 39% had R-NC fusion, and 2% had L-NC fusion. Eighteen of the 78 subjects had a family history of a left ventricular outflow tract defect”

**The authors state that of two variants were de novo, both present in the same individual with a family history of coarctation of the aorta. The authors should better clarify these variants. The table 1 is not clear because it reports two de novo and two unknown variants, but only p.T545M variant in MCTP2 gene seems to not have a minor allele frequency in 1000G. Please clarify.**

In this manuscript, we use *de novo* to indicate that the given variant is present in the proband and absent in both parents. Further clinical information relevant to these variants has also been added in Supplemental Table 2.

**The authors discussed a cost analysis of this pooling technique compared to both whole exome sequencing and whole genome sequencing, but what about a comparison of cost respect to targeted gene sequencing done on individual pools wherein each sample is labeled with a unique genetic “barcode”?**

The cost analysis has been expanded in the discussion as well as updated to the most up-to-date cost information available. A brief clarification was also made in the abstract, subheading conclusion.