Author's response to reviews

Title: Mutations in the 3'-untranslated region of GATA4 as molecular hotspots for congenital heart disease (CHD)

Authors:

StellaMarie Reamon-Buettner (reamon-buettner@item.fraunhofer.de)
Si-Hyen Cho (borlak@item.fraunhofer.de)
Juergen Borlak (borlak@item.fraunhofer.de)

Version: 2 Date: 20 April 2007

Author's response to reviews: see over
Response to Reviewers' Comments:

Reviewer: Giuseppe Novelli

A. Major Compulsory Revisions

1. The paper is interesting and offers new important suggestions for a molecular pathogenenic mechanism for CHD. Nevertheless, I find the paper confused. In fact the authors have conducted on the same 68 formalin-fixed hearts, mutation analysis of other cardiac transcription factors (see references 16, 17, 18, 40) but it is not clear if mutations in these other transcription factors are present even if in the presence of GATA4 mutations. Moreover, it is not clear if, in the 68 malformed hearts, are present at the same time both nonsynonymous mutations in the coding region of GATA4 (reference 19) and sequence variations of the 3'UTR.

Response:

To help elucidate genetic alterations in affected tissues of malformed hearts, we carried out genetic analysis of a panel of cardiac-specific transcription factor genes from the same 68 formalin-fixed hearts of patients with complex cardiac malformations. Surprisingly, direct sequencing of genes revealed mutations in diseased heart tissues, which were mainly absent in normal heart tissues of the same CHD patients. Results of these studies enabled us to put forward a hypothesis of somatic mutations as a novel molecular cause of CHD. While certain transcription factor genes (NKX2-5, GATA4) exhibited a high rate of mutations, others were basically not affected (HEY2, MEF2C). Furthermore, we found malformed hearts containing combination of mutations in several transcription factors.

For instance, from ref 37 (present manuscript):

<table>
<thead>
<tr>
<th>Table 1. Combined Mutations in the Binding Domains of HEY2, NKX2-5, TBX5 and GATA4 in Two Patients with Atrioventricular Septal Defects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

We also identified 21 patients with combined mutations in the GATA4 zinc fingers and the homeodomain of NKX2-5 (unpublished results).
2. Table 2 must be changed with a patient code to identify if a malformed heart carries different somatic mutations in the coding and 3'UTR of GATA4.

**Response:**
Instead of changing Table 2, we added a new Table (in the present version of the manuscript Table 3) which contains GATA4 nonsynonymous (ns) mutations and 3'-UTR in diseased tissues of the malformed hearts. We identified 15 malformed hearts with both ns and 3'-UTR mutations.

3. In the discussion section, the authors must specify better if they have found malformed heart carrying mutations in different genes.

**Response:**
In Discussion, page 12, paragraph 2, we included:

"To help elucidate genetic alterations in affected tissues of malformed hearts, we have investigated in the past a panel of cardiac transcription factor genes from the same 68 malformed hearts [16-19,37]. Direct sequencing revealed mutations in diseased tissues, which were basically absent in matched normal heart samples. Common occurring mutations were identified, especially in the binding domains of transcription factors, which could affect DNA-protein or protein-protein interactions leading to CHD. While certain transcription factor genes (NXX2-5, GATA4) exhibited a high rate of mutations, others were not or rarely affected (HEY2, MEF2C). Results of these studies enabled us to put forward a hypothesis of somatic mutations as a novel molecular cause of CHD. Furthermore, we found malformed hearts containing combination of mutations in the binding domains of several transcription factors. As example, we identified in two patients with AVSD combined mutations in the binding domains of HEY2, NKX2-5, TBX5 and GATA4 [37]. We also identified 21 patients with combined mutations in the GATA4 zinc fingers and the homeodomain of NKX2-5 (unpublished results)."

4. At page 7 line 20, the authors affirm they amplify the first 609 of the 3'UTR; why they don't analyse the entire 3'UTR (1,525 bp).

**Response:**
In the blood samples affected by CHD, we analyzed the entire 3'-UTR located in exon 7, but in the formalin-fixed material we were successful only in amplifying the first four fragments of exon 7 covering the first 609 of the 3'UTR. There is also suggestion that the whole 3'-UTR is not necessary for localization, but that localization signals lie within the regions of <100-200 nt [ref 25].

**B. Minor Essential Revisions**

1. At page 8, line 12 the nucleotide position of the polyadenylation signals (c.+1505-1507) does not correspond to the length of the signal (6 bases).

**Response:**
Thank you for pointing out this typographical error. We revised the manuscript on page 8, starting line 12:
"None of the conserved motifs reported so far have been detected, but the highly conserved motif (AATAAA) associated with polyadenylation signals were located at positions c.+678-683 and c.+1502-1507."

2. At page 8 line 16 is reported that 17/68 malformed heart do not carry nucleotide changes in the 3'-UTR. By summing the mutated malformed heart from Table 2 (column Total), the number of mutated heart is 47; 47+17=64! Is this analysis failed in 4/68 formalin-fixed malformed heart?

Response:
It does not tally because the numbers correspond to positive cases for each of the mutations given in Table 2, in which a malformed heart can carry more than one mutations. Please see Table 3 of the present version of the manuscript.

In the present Table 3, there are 38/68 hearts negative for the 3'-UTR mutations reported here. As mentioned in Discussion, page 10, 2nd paragraph : "Additional sequence alterations in the 3'- UTR of GATA4 were detected, but these were mostly isolated cases and therefore not reported here." Overall for all sequence alterations detected in the 3'-UTR region, there were 17/68 without any sequence alterations except the dbSNPs in analyzed fragments.

3. At page 8 lines 17-19 are mentioned fragments 7-1, 7-3, 7-4. They have not described these fragments before. What is their location?

Response:
In Results, page 7, 2nd paragraph, we revised the manuscript to give the location of the four fragments which we were able to amplify in the formalin-fixed material.

"Notably, exon 7 of GATA4 consists of 1,708 bp, the majority (1,525 bp) being untranslated. From our collection of formalin fixed hearts, we could amplify the first four fragments of exon 7 (see Fig. 1A and Table 1 for the primer sequences designated GTx7-1 to GTx7-4). These primer sequences cover a part of the coding region c.1221 to c.1329 (nt 1739-1847, NM_002052) and 3'-UTR from c.+1 to c.+609 (nt 1848-2456, NM_002052). We found 9 sequence alterations, occurring in 3 to 7 patients, and 6 dbSNPs in the analyzed 3'-UTR region of GATA4 (Table 2; Table 3 see additional file 1)."

In Results, page 8, 2nd paragraph, we further revised the manuscript,

"Thus for this comparison, we analyzed 21 patients for GTx7-1 (21 x 2 x 289 bp =12,138 nucleotides); 25 patients for GTx7-3 (25 x 2 x 281 bp =14,050 nucleotides) and 24 patients for GTx7-4 (24 x 2 x 270 bp = 12,960 nucleotides)."

In formalin-fixed tissues, the fragments amplified in exon 7 based on NM_002052 are located in:
- GTx7-1 (nt 1739- 2027)
- GTx7-2 (nt 1842-2132)
- GTx7-3 (nt 1985-2265)
- GTx7-4 (nt 2187-2456)

4. Figure 3B must report the haplotype described in the text (page 9 lines 11-13).
Response:
We included the genotypes for (c.+119A>T) in Figure 3B, assuming that this is being asked instead of the haplotype.

Reviewer: Vidu Garg

Minor Essential Revisions

1. Consider commenting on the significance of finding the c.+119A>T variation in an unaffected individual. The two possibilities of incomplete penetrance versus no causal association should be clearly stated.

Response:
In Discussion, page 10, 2nd paragraph, we revised the manuscript:

"Further, we identified a germline mutation (c.+119A>T) in which two affected family members were positive. Similarly, it would alter RNA folding. Since the unaffected father carried c.+119A>T, we cannot discount the possibilities of a low penetrance mutation or no causal association at all."

2. The authors may consider expanding the discussion to comment more on the relevance of the haplotypes. According to the authors' response, the presence of multiple haplotypes may be the result of chromosomal duplications and/or rearrangements in a subpopulation of cardiomyocytes near the diseased tissue. This implies multiple levels of DNA damage (point mutations to chromosomal rearrangements) to the diseased cardiac tissue and calls into question if the genetic abnormalities are causal versus an artifact of DNA damage from tissue processing. The authors should comment upon this finding in the discussion.

Response:
In Discussion, pages 13-14, we added:

"This result is similar to our previous observations on different cardiac transcription factors conducted on the same set of malformed hearts [16-19,37]. After cloning amplified fragments with several closely-spaced 'heterozygous' mutations as marker loci within a single gene, we observed several haplotypes in individual hearts, instead of two haplotypes as expected of a diploid genome. (Note, we used the term haplotype to define the set of alleles within an investigated gene locus). The cause of the observed multiple haplotypes is unknown, but may be explained by a mixed population of cardiomyocytes carrying different mutations or de novo chromosomal rearrangements and gene duplications in the heart tissues of patients affected by CHD. Our results for NKX2-5 by using a yeast-based assay to determine function suggest that different haplotypes can lead to different cardiac disease phenotypes [41]. Furthermore, the presence of combined mutant alleles may alter/modify pattern of mRNA folding. We found, for instance, patients who were either positive for c.+218C>T or c.+259 A>G or both. Two patients were heterozygous for both mutations, while one patient was homozygous for both variant alleles. Singly c.+218C>T or c.+259 A>G would lead to misfolding (see Fig. 2), but if combined only the pattern observed for c.+259 A>G would result. Moreover, different
haplotypes due to combinations of closely-spaced polymorphisms in the 3'-UTR of genes can result in different mRNA stabilities in transient expression assays [35]. Nonetheless, as the malformed hearts were conserved in formalin, possibilities exist that the observed multiple haplotypes could be PCR errors resulting from fragmented DNA. We believe the contrary, however, as specific haplotypes were detected in several malformed hearts after cloning, but were absent in matched unaffected heart tissue (unpublished results)."