Reviewer's report

Title: Analysis of positional candidate genes in the AAA1 susceptibility locus for abdominal aortic aneurysms on chromosome 19

Version: 1 Date: 21 August 2010

Reviewer: Ynte M Ruigrok

Reviewer's report:

- Major Compulsory Revisions

In this study the authors selected nine candidate genes from the 19q13 abdominal aortic aneurysm (AAA) locus based on known function, as well as mRNA expression levels in the aorta. 394 AAA cases and 419 controls was genotyped for 41 SNPs located in or around the selected nine candidate genes. Several SNPs were nominally associated with AAA (p < 0.05). The SNPs with most significant p-values were located near the CCAAT enhancer binding protein (CEBPG), peptidase D (PEPD), and CD22. These three genes were selected for further analysis including DNA sequencing of the exonic regions in 23 cases. Immunohistochemical staining of aortic tissue sections from AAA and control individuals was carried out for the CD22 and PEPD proteins with specific antibodies.

The major problem of this study relies in the fact that the authors performed an association study to find further evidence for the involvement of the candidate genes in the pathogenesis of AAA. Based on the results of this association study the candidate genes are subjected to further analysis, including DNA sequencing of the exonic regions and immunohistochemical staining. However, there are some concerns on the association study:

1. the number of 394 patients and 401 controls used is low. The authors also show this in their power calculation in the Methods section on pages 8 and 9. The question arises whether the authors were able to show a difference in allele frequency between patients and controls using this sample size.

2. a large number of SNPs (i.e. 55 SNPs) were analysed for association but no form of multiple correction was applied. Applying a simple Bonferroni correction to the p-values of the genetic association study as shown in Table 2 on page 37 then only one SNP located in PEPD rs7248389) remains statistically significant.

3. since no multiple correction was performed one may expect the authors to use replication instead to validate their findings. However, this was also not performed or discussed.

4. 11 of the 55 genotypes SNPs were excluded for further analysis since these 11 SNPs had low call rates. So, 20% of the tested SNPs were excluded which is a high proportion. After exclusion of these SNPs can the authors ensure that
using the remaining SNPs enough genotyping information is generated to demonstrate association of the nine candidate genes with AAA?

5. five of the remaining 44 SNPs showed deviation from the Hardy Weinberg equilibrium (HWE). The authors state that HWE deviation can result from association and therefore three of the five SNPs showing association were excluded. However, another important reason for HWE deviation is genotyping errors. In my opinion it is therefore more important to remove the other two of the five SNPS not being in HWE.

The authors do not comment on these statistical problems in their manuscript. They should discuss these issues in the Discussion section. To overcome problem numbers 2 and 3 the authors should consider genotyping the associated SNPs in a replication cohort.

When considering only PEPD as a candidate gene, since after multiple correction only the association of this gene holds, then not much further evidence is found to support involvement of this gene in the pathogenesis of AAA. PEPD protein was expressed in fibroblasts and myofibroblasts in the media-adventitia border in both aneurysmal and non-aneurysmal tissue samples. Seven sequence variants were identified in PEPD, including three not present in the NCBI SNP (dbSNP) database. However, after investigation for predicted functional consequences, all of the sequence changes identified in the coding regions or 3'-UTR of CD22 and PEPD appeared to be tolerated.

- Minor Essential Revisions

The results of the sequencing study for PEPD are not convincing. For the sequencing study the authors used cases and controls. However, the amount of controls used is low, for example in the PEPD analysis only 2 controls were used. In one of the controls (number 43) two of the seven identified sequence variants are also found (variants four and six). The authors do not commend on this in the manuscript. Three of the seven variants are not found in dbSNP. From the manuscript is does not become clear whether these three variants include the two variants also found in the control subjects of this study.

During the introduction en methods sections it does not become clear how the analysed candidate genes were selected. This was finally explained in the results section. This explanation belongs to introduction and/or methods section.

In the first paragraph of the method section on page 7 the authors describe the patients included in their study. They have included four patients from a previous study described in reference 13. In this study described in reference 13 119 families with AAA were included. Why did the authors only include four patients of this study? Why did they not include one patient from each family, totalling 119 patients?

Overall the manuscript is rather long. At some points information is repeated in different sections. For example the way the analysed candidate genes were selected is described in the results section and again in the Discussion section.
The introduction section can be shortened. For example, paragraphs 3 to 5 describing possible candidate genes in the 19q locus can be shortened or moved to the Discussion section.

Discussion section: I miss a first paragraph describing the main findings of the study. Instead, the way the analysed candidate genes were selected is described again in this paragraph. This was already explained previously and there is no need to repeat this.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests