Reviewer's report

Title: Two novel missense substitutions in the VSX1 gene: Clinical and Genetic analysis of families with Keratoconus from India

Version: 1
Date: 2 September 2014

Reviewer: Forbes Manson

Reviewer's report:

1. Major Compulsory Revisions

I. In Figs. 3A and 3B identify the family member who is shown in the sequence and compare it to the unaffected family member.

II. Need at least 105 controls (210 chromosomes) for sufficient power (80%). 340 chromosomes would be preferable (95%). See Collins & Schwartz (2002). AJHG 71; 1251-2.

III. Variants in VSX1 are an extremely rare cause of keratoconus. However in this study 2/8 families (25%) were found to have a putative mutation. This high ‘success’ rate needs discussed, as does the fact that the novel mutation was the same in both families. Considering the number of studies and KC patients that have been screened previously, finding the same novel mutation in 2/8 families strongly suggests these families are related. If this cannot be determined from family histories it should be addressed by genotyping flanking markers.

2. Minor Essential Revisions

I. P.3 ‘Genetic analysis of KC patients with different ethnic backgrounds has revealed multiple mutations in the VSX1 gene.’ This suggests a lot of mutations that in turn implies the locus is more important than it is. Probably only 4 confirmed mutations (5 at a push). Rephrase accordingly.

II. P.3. Locus is known more accurately than 20p11-q11 (20p11.21). Correct.

III. P.5. The text identifying the probands does not match the indicated probands on Figs. 1A and 1B. Correct.

IV. P.5. ‘suggestive of KC’. Is it KC or not? Clarify.

V. The level of English is poor in several places and needs corrected.

VI. At position c.767 in VSX1 the residue is a ‘C’ in the codon GCC (Ala). The correct codon for p.Leu268 is c.802-804.

VII. P.6. ‘c.767 T>A mutation, CTC (Leu 268) to CAC (His 268)’. This is clumsy. Better to list amino acid change after the cDNA change as p.Leu268His.

VIII. Table 2 shows the change as at c.767 as T>C instead of T>A. cDNA numbering is wrong as well, as noted above.

IX. P.7. In silico analysis are predictions (as acknowledged in first line), so cannot then say in next line that a variant ‘is pathogenic’. This has to be demonstrated
empirically. Based on in silico analysis a variant is no more than a putative mutation. This terminology must be corrected throughout the manuscript.

X. P.7. ‘However, leucine 268 amino acid residue was located in the C-terminal region of the CVC CVC (Chx10/Vsx-1 and ceh-10) domain of VSX1 and has been previously reported in familial KC patients [16, 24]. This reads as if the variant in Leu268 has been reported twice before in references 16 and 24. This is not the case and is probably not what is meant. Re-write for clarity.

XI. P.7. ‘Moreover, the pathogenic nature of this mutation also comes from leucine’. This does not make sense. Rephrase for clarity.

XII. P.8. ‘Furthermore, in our study, this potentially damaging mutation was detected in all affected individuals with a dominant inheritance of KC without PPCD’. This is inaccurate. You screened 20 KC patients and found it in two families consisting of 5 affected individuals.

XIII. P.8. ‘This is consistent with previously identified mutation studies in an Iranian family [24].’ It is not clear what you think is consistent.

XIV. P.8. ‘This is in agreement with previous studies suggesting that missense substitutions in the VSX1 may or may not be a disease-causing variant [16, 36]’. This gets to the crux of the matter with VSX1 variants. They are a very rare cause of KC and some authors have dismissed VSX1 as a causative gene. More needs to be given to this discussion point highlighting the difficulties in being certain that VSX1 is causative. Many studies find no putative pathogenic variants or find that previously reported mutations are actually polymorphisms.

XV. P.8. ‘In this study we were able to identify genetic changes in three out of eight families’. This is close to over-playing what has been found. A single, putative mutation, was found in 2/8 families. A second sequence variant of unknown significance (but probably a benign polymorphism) was found in a third family.

XVI. P.8. ‘In summary, we add two novel missense variations in the coding region of VSX1 to the existing spectrum of VSX1 mutations observed in Indian patients with the characteristic phenotype of KC’. As previously noted, from your data you have identified one putative mutation and one VUS. The mutation can only be considered as such because it has been found in two families. However if the families cannot be shown to be unrelated than this variant will also have to be considered as a VUS.

3. Discretionary Revisions

I. P.5. ‘Another missense substitution (S251T) was identified in both patients and their family members’ would be clearer as ‘both affected siblings and their affected father’

II. P.6. Suggest ‘Amino acid conservation analysis revealed that the wild type leucine 268 was highly conserved among the VSX1 orthologs in other species’ would be better as ‘Amino acid conservation analysis revealed that leucine at position 268 is conserved in 9 vertebrate orthologs’ (a residue is either conserved or it is not).
III. P.7. Only need to expand CVC on first use.

**Level of interest:** An article of limited interest

**Quality of written English:** Needs some language corrections before being published

**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