Reviewer's report

Title: Role of Helicobacter pylori cagA EPIYA motif and vacA genotypes for the development of gastrointestinal diseases in Southeast Asian countries

Version: 1 Date: 4 June 2012

Reviewer: Stacy Duncan

Reviewer's report:

This manuscript ["Role of Helicobacter pylori cagA EPIYA motif and vacA genotypes for the development of gastrointestinal diseases in Southeast Asian countries"] by Sahara et al. seek to evaluate the association between genotypes of H. pylori virulence factors, cagA and vacA, and disease susceptibility in persons infected with H. pylori in Southeast Asia. Previous studies have demonstrated that strains of H. pylori expressing virulence factors vacA (type s1, i1 and m1) and East Asian forms of cagA (encoding an EPIYA-D motif) are associated with a higher risk of H. pylori related gastric cancer. In this study, the authors perform an evaluation of 14 previous studies that have investigated the cagA and vacA genotypes in 1,355 H. pylori strains from Southeast Asian populations.

The authors validate that the prevalence of cagA and vacA genotypes are variable among H. pylori strains from Southeast Asian countries. Furthermore, they present evidence suggesting a significant association between genotypes of H. pylori virulence factors, cagA and vacA, and the development of gastric cancer and peptic ulcer disease in Southeast Asia. However, inclusion of a number of details would improve the manuscript.

Minor Essential Revisions:

1. Statistical revisions
   
o Page 6. Provide statistics (p-value) to support the following: “In the analysis of the vacA s1 subtype, most strains isolated from Chinese patients were the s1c subtype (75%), which was significantly higher than that found in Thai and Thai-Chinese patients.”

   o Page 7. Provide statistics (p-value) to support the following: “In Malaysia, the prevalence of vacA m1 genotypes was similar among Malay (64%), Chinese (55%), and Indian (70%) patients.”

   o Throughout the manuscript the authors should include the statistic used to calculate p-values.


   o " Greater than 90% of H. pylori strains isolated from East Asian populations carry the cagA..." This should read " Greater than 90% of H. pylori strains isolated from East Asian populations carry the cagA gene.”
o Provide some background regarding the effects of cagA on gastric epithelial cells.

o “As only reports from Nguyen TL, et al. [28] and Schmidt HMA,…” This should read “Reports from Nguyen TL, et al. [28] and Schmidt HMA,…”

o Describe 1a and 2a types of EPIYA motifs.

4. Page 8. In “Malaysia within the last 200 years resulting in the spread of hspEasia strains…”, hspEasia should read hspEAasia.


o Define what is meant by (cagA type 2a/cag right-end junction type III) and (cagA type 1a/cag right-end junction type II).

o Consider moving the following from the Discussion section to the Results section: “The examination of strains isolated from different ethnic groups provided an opportunity to examine the hypothesis that ethnic groups tend to retain their ancestral H. pylori genotype after migration [7, 41]. Intermarriage among ethnic Thai and Chinese in Thailand also allowed examination of the transmission pathways in relation to the ethnic status of the parents. When the mother was Chinese, 84% of the offspring possessed H. pylori with an East Asian genotype. By contrast, when mother was Thai and father Chinese, only 29% of strains among offspring were of East Asian types. This finding is consistent with the notion that the primary caretaker of the children is the most likely source of infection.”

7. Page 11. “Southeast. Importantly, most of the H. pylori strains isolated from countries…”. This should read “Southeast Asia. Importantly, most of the H. pylori strains isolated from countries…”

8. Throughout the manuscript, there needs to be consistency when referring to cagA EPIYA-C, cagA EPIYA-D, East Asian, Western, ABD and ABC. For example, if a sentence refers to ABC and East Asian in the same sentence it would be better to state Western (EPIYA-ABC) and East Asian (EPIYA-D).

9. Throughout the manuscript the use of peptic ulcer should read peptic ulcer disease and should be restated as in the following example from page 7: “Nguyen et al. reported that the frequency of vacA m1 strains in peptic ulcer” should read similar to the following: “Nguyen et al. reported that the frequency of vacA m1 alleles among H. pylori strains isolated from patients with peptic ulcer disease…”.

10. Table 1. Italicize cagA, insert a header above ABC, ABCC, ABD, 2a and 1a indicating EPIYA motifs
11. Table 2-4. Insert a header above Western and East Asian indicating cagA genotype and include EPIYA-C and EPIYA-D under Western and East Asian, respectively.

12. Tables 1-4. Replace Strain with vacA genotype.

**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.