Author's response to reviews

Title: Association of HLA Class I with Severe Acute Respiratory Syndrome Coronavirus Infection

Authors:

Prof Marie Lin (marilin@ms2.mmh.org.tw)
Hsiang-Kuang Tseng (drtseng@anet.net.tw)
Mr Jean A Trejaut (itrejaut@ms1.mmh.org.tw)
Hui-Lin Lee (leehl@ms1.mmh.org.tw)
Jun-Hun Loo (jjunhun@ms1.mmh.org.tw)
Dr Chen-Chung Chu (chucc@ms1.mmh.org.tw)
Pei-Jan Chen (g8640@ms13.hinet.net)
Ying-Wen Su (yingwen_su@yahoo.com.tw)
Ken Hong Lim (khl@seed.net.tw)
Zen-Uong Tsai (tsaizu@ms1.mmh.org.tw)
Ruey-Yi Lin (lin-8424@mail.hoping.gov.tw)
Ruey-Shiung Lin (rueyshiung@yahoo.com)
Chun-Hsiung Huang (chhuang@ms2.mmh.org.tw)

Version: 4 Date: 29 Jul 2003

Dear editor,

I am very thankful to the reviewers. Their very constructive comments enlightened some important points. All revisions and advices were considered and discussed among us, and whether discretionary or compulsory have been addressed in the paper.
Please see our answers intercalated in italics in relation to each reviewer comment.
A copy of the revised article has been resubmitted.
I hope the article is now fit for publication and answers your requirements

Thank you.
Regards

Professor LIN

Answers to Reviewer's report

Title: Association of HLA Class I with Severe Acute Respiratory Syndrome Coronavirus Infection

Authors: Prof Marie Lin (marilin@ms2.mmh.org.tw) et al.
Version: Date: 3 15 Jul 2003
Reviewer: Satoshi Horai

1) A total of 190 healthy unrelated Taiwanese are used as control B (one of two control groups) in this study. They were just analyzed for HLA-A and -B loci at the serological level (Lin et al. 2001, Tissue Antigens, 57:192-199) as described in the manuscript, and hence it must be impossible to define HLA-B allele types for control B according to the previous serological data only. The authors need to explain completely how to determine individual HLA-B allele types for control B in the present study, because this case-control association study (probable SARS vs. control B) is
performed at the allele type level (as shown in Table 2 and the text).

- To answer this question the following comment was added in the "Statistical analysis" paragraph: When analyzing data containing results obtained at the serological level, corresponding data in the patient group were converted to serology typing using the World Health Organization nomenclature (WHO) table of correspondence (Marsh et al. 2002).

Consequently HLA allele names in column one of table 2 have been amended to indicate both level of analysis. Finally reference "Marsh et al. 2002" was added to the reference list.

2) In the research field of HLA and disease association, it is highly possible that an observed association comes from a statistical artifact due to multiple comparisons (e.g., Svejgaard and Ryder 1994, Tissue Antigens, 43:18-27). Consequently, it is important (and necessary) to conclude HLA association with SARS on the basis of the P-values corrected for the number of comparisons (Pc-values). In the present study, only significant association of HLA-B*4601 with the severity of SARS is observed with the Pc-values (vs. control A [non-infected health care workers]). With the uncorrected P-values, on the other hand, four HLA-B alleles (B*4601, B*1301, B*5401 and B*3901/B39) may be positively or negatively relevant to SARS infection in some case-control combinations (Table 2 and 4). Taken together, therefore, the association of HLA-B*4601 with SARS is less erroneous, but the authors should mention the need for further independent studies to test this hypothesis. This is because there is no statistical significance in any comparison with control B or non-SARS patients with fever, even with the uncorrected P-values (Table 2 and 4). Next, other three alleles (B*1301, B*5401 and B*3901/B39) are unlikely associated with SARS infection. Thus, the authors should explain more fully why they conclude that HLA-B*1301 appears to be protective against SARS infection.

- "The need for further independent studies to test the hypothesis presented" in the paper has now been added in the result section of the Abstract, and in the last paragraph of the discussion. Furthermore a few lines in the last paragraph of the discussion have been devoted to HLA-B*1301 explaining why we thought B*1301 was preferred to be a potential protective element in Asian populations.

3) The epidemiological discussion (pp. 4-5) in this manuscript is very interesting, but additional information (such as appropriate references) is necessary for the readers to understand it more easily. For example, there is no reference about the following points: genetic characteristics of Taiwanese (p. 4, bottom line), allele frequency of HLA-B*4601 in Taiwan indigenous people (p. 5, lines 13-14), and the degree of differences in immune repertories between Southeast Asian and Europe/African populations (p. 5, lines 19-20).

- References were added were deemed necessary or as recommended by the reviewer. As for the degree of differences in immune repertories between Southeast Asian and Europe/African populations, the Variation was inferred from the heterozygosity (h = 1 - sum of squares of allele frequencies) seen in these populations, and h was calculated (data not shown) from data of Proceedings of the 11th International Histocompatibility Workshop and Conferences. This last comment was added in the text (see last paragraph of discussion).

4) On p. 7, the authors should remove the SARS patient I.D. (probably initials for the patients) from Table 1, in order to keep personal genetic information confidential. Also, the patient information such as gender and age in table 3 (p. 8) is not necessary in the context of this manuscript, and can be removed.

- Although gender and age in table 3 is not a necessity we think the information may interest members of medical community and be of relevance for the following point. When constituting the "5 deceased or intubated probable SARS patients" group, one 91 years old patient was not included.
5) Other comments
This manuscript has frequent lack of description as to which control group is used to obtain the
probability and odds ratio of HLA-B allele association. This information should be mentioned
explicitly. In the Abstract (p. 1), for example, there is no information about the following parts in the
Results section: HLA-B*4601 (OR=2.08, P=0.05), HLA-B*5401 (OR=5.44, P=0.02) and
HLA-B*1301 (OR=0.16, P=0.05).

- Particular attention was drawn to describe appropriately the provenance of each P-value when it
appeared in the text/abstract.

On p. 4, line 3, the number of healthy unrelated Taiwanese examined (control B) must be 190 (not
101).

- The number of healthy unrelated Taiwanese examined (control B) was changed to 190.

On p. 5, line 15, the cited article (reference no. 8) must be inadequate because the article contains
no date on allele frequencies of HLA-B*1301 among Taiwan indigenous tribes.

- HLA-B*1301 is the most frequent allele in the B13 type among Southern Asian and most
particularly among Taiwan Aborigines. This data has not been published and can only be inferred
from Lin et al. (2000). This has now been made clear in the text (see discussion)

In Table 2 (p. 8), why is the P-value for B*5401 (vs. control B) not available?

- The P-value for B*5401 was not significant (n.s.). This error was corrected to n.s.

In Table 3 (p. 8), HLA-B alleles (B*4601, B*5601) for patient no. 1 are incompatible with those
(B*4601, B*5801) presented in Table 1 (p. 7). Which is correct?

- This error was changed to B*5801

There are many mistakes in citation (pp. 5-6). Reference nos. 2, 4, 5, 8, 9, 10, 15 and 16 should be
revised appropriately or updated.

- All citations were corrected and new references added accordingly.

****
Answers to Reviewer's report
Title: Association of HLA Class I with Severe Acute Respiratory Syndrome Coronavirus Infection
Authors: Prof Marie Lin (marilin@ms2.mmh.org.tw) et al.
Version: Date: 3 20 Jul 2003
Reviewer: Alicia Sanchez-Mazas

Discretionary revisions:
1. The "patients" section in materials and methods is very confusing and should be shorten or the
information put in a table. In fact, the first paragraph of the Results is a summary of that section and
could practically replace the latter.

- This section was left unchanged as the authors thought that details were important to members of
the medical community.
Compulsory revisions:

2. The authors used the powerful Fisher's exact test for their analyses, which is a reasonable choice with low sample sizes (patients). On the other hand, it's not clear whether the tests are done on the number of individuals or on the number of alleles detected. If it's on the number of alleles, the fact to assume no blanks, i.e. to consider two identical alleles in individuals showing only one allele (e.g. case no 33) at typing may bias the results, especially when serological typing is used, which is the case for HLA class I in the Taiwanese control sample B. As the P values are not highly significant (often equal to 0.05), such a bias could dramatically change the conclusions.

- The tests described in the paper were conducted according to the number of alleles detected without blanks, and therefore also applied assuming no blanks when serology results were used. This has now been precised in paragraph labelled Statistics. Although not described in the paper, we also conducted the test at the phenotypic level (with blanks). These results are available on request. As expected there were monor differences, but the same conclusions could be drawn. For information, when comparing the 33 probable SARS cases (table 2) and control A (101 high risk non-infected health care workers) the following results were obtained: B*4601, OR=2.2, P=0.06; B*1301, OR=0.15, P=0.03; B*5401, OR=5.83, P=0.02; B*3901, n.s.; again similar results were obtained to previous serotype analysis when using control B.
- When reconstructing table 4, the following was obtained HLA B*4601 in 5 deceased or intubated probable SARS patients and the 28 excluded fever patients: OR = 10.0, P=0.05 and Pc = n.s.; with control A: OR = 13.47, P=0.014 and Pc=n.s.

3. Line 7 of page 4 "Both control A and B showed significant odds ratios (OR=0.16, P=0.05)": in fact one P-value is 0.04 according to table 2.

- This was changed to (OR=0.16, P=0.05 and OR=0.16, P=0.04 respectively)

4. Lines 10-11 of the Discussion should be revised (not clear).

- The passage was changed to:

Although HLA-B*4601 frequency in the 33 probable SARS patients (table 2) was higher than the frequencies seen in control A or control B (P=0.05 and P=0.06 respectively), the frequency HLA-B*4601 in the five severe cases (table 4) was significantly increased when compared to the 28 excluded fever patients or control A (P=01, Pc=n.s. and P=0.0008, Pc=0.03 respectively) and appeared to be associated with the severity of SARS.

5. Line 6 of page 5: replace "Caucasian" by "Europeans" or "of European origin", otherwise specify better.

- The term European origin was used as recommended by the reviewer.

6. What justifies saying, on line 7 of page 5, that many peoples in Beijing are "genetically different to southern Chinese"?

This comment has now been withdrawn. This gives < Many peoples in Beijing, generally considered as northern Chinese [16], were also infected.>

7. Also, in line 2 of the second paragraph of page 5, "Taiwan indigenous peoples are genetically distinct from Taiwanese": Lin has shown that Taiwanese tribes were very distinct from each other, so we cannot consider them together to compare them to Taiwanese.
This passage is a general statement, and does imply that although indigenous peoples are distinct from each other they are also very different from the Taiwanese. This passage has nevertheless been changed to: This is most likely due to the fact that the HLA make-ups of the Taiwan indigenous peoples have little in common to the Taiwanese [9] and/or may not contain factors favouring SARS infection.

8. Next line, "HLA-B*4601 is a southern Asian gene": we are not allowed to assign specific genes to specific populations as sample sizes are never very high. What the authors mean is probably that southern Asians exhibit a higher frequency of this allele or that the allele has only been detected in those populations thus far. Please specify.

- We did as suggested and the passage became as follow: For example, HLA-B*4601/B46 is a gene seldom seen in Europeans populations thus far, and when compared to Northern Han, Southern Asian populations HLA-B*4601/B46 displays high frequencies (Northern Han 2.8%, Southern Han 15.4%, Singapore 15.1%, Vietnamese 13.2%) [18]. Furthermore, HLA-B*4601 has rarely been seen among indigenous peoples [9] except in children of intermarriage between Taiwanese and indigenous peoples [9].

9. Seven lines below, "Asian peoples (...) also have less variation in their immune repertoires than European or African populations": explain what justifies this sentence and give references.

- The passage was changed to: Certainly southern Asian peoples, not only live in highly densely populated regions, but also have less variation in their HLA related immune repertoires than European or African populations [Variation was inferred from the heterozygosity (h = 1 - sum of squares of allele frequencies) seen in these populations, and h was calculated (data not shown) from data of Proceedings of the 11th International Histocompatibility Workshop and Conferences][reference 18], which as a whole creates a favourable factor for rapid settling of any epidemic.

10. Table 1: - Explain in the note what I.D. is (second column)
- Last line in the note: "Most", not "ost".

The note I.D. and the column underneath were removed as deemed unnecessary. "ost" was changed to "Most".

11. Table 2:
- OR*: explain what "*" is.
- Note: Add "(OR)" next to Odds ratio the first time the term appears.
- Do the cells correspond to numbers of individuals or numbers of alleles? The fact to use both kinds of numbers in the table headings is confusing.

- The table has been corrected as suggested and indicates clearly that cells in the second row correspond to the total number of alleles. The table has also been modified to avoid confusion introduced by the P-value. Since no corrected P-values were significant, P-values were not indicated.

- Note: "all correction factors were greater than 27": why using "27"?

- This note has been withdrawn. The idea here of indicating the smallest correction factor observed among the loci was to allow the readers to confirm for themselves that none of the Pc-value were significant by using the minimal estimate: Pc-value >= smallest correction factor x P-value.
12. Table 4:
Pc(33): explain better the correction factor.

- The following note was added to Table 4: The correction factor (number of different alleles or number of comparisons) is shown within parenthesis;
Pc-value = P x Correction factor (Edward, 1974).