PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Annexin A11 (ANXA11) gene polymorphisms are associated with sarcoidosis in a Han Chinese population - a case-control study
AUTHORS	Zhang, LG; Xiao, ZJ; Shi, GC; J, Huang; Zhang, CX

VERSION 1 - REVIEW

REVIEWER	António Morais1, Bruno Lima 1 Department of Pneumology, Hospital S˜ao Jo˜ao, Porto, Portugal; Department of Medicine, Faculdade de Medicina do Porto, Porto, Portugal 2 North Histocompatibility Centre, Porto, Portugal
REVIEW RETURNED	07-Jan-2014

GENERAL COMMENTS	In this manuscript, LG Zhang et al studied in a Han Chinese cohort a protective putative role of ANXA 11 polymorphisms regarding sarcoidosis susceptibility. This association, namely in what rs1049550 T allele is concerned had been previously described in American and in three European populations. The validation of a previously genetic association described in a certain population is crucial because of the significant variety of sarcoidosis prevalence and type of presentation in the different populations as also for the different genetic associations found in different ethnicities. The authors confirmed in a population with a different genetic background in comparison with the populations were the ANXA polymorphisms had been studied, the association with sarcoiodosis susceptibility (3 SNP), which emphasizes the alleged role of annexine A11 in this context. Although a concise well written manuscript, there are some questions that we would like to have a clarification by the authors: - In the introduction and in table 1 the authors describe the published sarcoidosis association studies for ANXA 11. However, there is one study published recently, carried out in a Portuguese population, which has a different genetic background in comparison with the other studied European populations that is absent. Annexin A11 gene polymorphism (R230C variant) and sarcoidosis in a Portuguese population Tissue Antigens, 2013, 82, 186–191 - The group of patients has a mean age of 53.6±4.6 years, which is particularly high taking into account the usual descriptions in different populations. The authors should address and explain this point.

	- Describe patients according with their demographic characteristics- How many are men? - Although the sample size is presented as a limitation (page 4) no calculations for it are given - After a characterization in material and methods of different clinical presentations and evolutions, strangely in the results we don't see any stratification of the patients according with these different phenotypes. This is very important in order to understand the real categorization of the population and if any of the significant SNP associations observed could be associated with a particular phenotype or a different outcome. The previous reports showed some different associations regarding for instance different type of clinical presentations or thoracic stage (Scadding criteria). - The authors should give eligibility criteria for the controls – How were they recruited? If consecutively, when they were recruited? Have they been recruited according to age or sex? - Although, mentioned in the statistical analysis, no results from logistic regression are given. - Page 9, line 31 – "Gabriel et al. 14" – reference 14 correspond to Mrazek et al, 2011 - Page 10, line 41 – authors should consider rewriting the sentence "The significantly more G-G-C haplotypes (block 4) (P = 0.027) were (…)"

REVIEWER	Albert M. Levin, Ph.D. Henry Ford Health System United States of America
REVIEW RETURNED	15-Jan-2014

GENERAL COMMENTS	The current paper describes a validation study of common ANXA11 variants and risk of sarcoidosis in Han Chinese. This is an important paper as it demonstrates association with sarcoidosis risk in a Chinese population, which is consistent with results found in European populations. Prior to publication, there are some key changes that need to be addressed that require additional text and analyses. Major comments 1) For the description of the tagSNP selection, there seem to be some key features left out. First, the metric for selection (for example, r2) and the threshold value for determining what SNPs are tagged are not presented. Second, the authors state that tagSNPs were selected from both the 1000 genomes and HapMap, but they do not specify the number of tagSNPs selected from each source. Third, there is no mention of the total number of SNPs that are tagged in either the HapMap or the 1000 genomes data by the 15 tagSNPs that were selected. A complete description should list the total number and proportion of variants (at a minor allele threshold of 0.05) that are tagged by the selected tagSNPs. Fourth, the base pair boundaries of the tagSNP selection should also be included. 2) The authors have based the linkage disequilibrium (LD) block

	definitions on the combined samples of cases and controls. LD patterns can differ between a sample of the base population (i.e. the controls) and the case subjects. The authors have shown different haplotype frequencies between cases and controls within the blocks that reflect these differences in LD between cases and controls. To ensure that the LD blocks most closely reflect the population level LD patterns, definition of the blocks should be based on the control samples alone or alternatively on the original CHB HapMap samples. Actually, it's not clear to me exactly which samples were used to define the blocks. This should all be described in the methods and should be based on either the CHB or the controls from their sample. 3) Based on the statistical methods section, the authors do not appear to have evaluated whether adjustment for covariates such as age and sex influence the SNP odds ratios. The manuscript would be improved by evaluating whether adjustment for these important potential confounding variables alters the SNP odds ratio estimates and significance. 4) In the statistical methods section, there is no description of the haplotype analyses performed. The readers will need to know what method was used, if there was adjustment for covariates, what genetic model used to estimate the haplotype odds ratios, how rare haplotypes were handled, ect. Also, the presentation of the haplotype results in Table 3 could be improved. First, it is not where the haplotype blocks and the corresponding haplotypes start and end. Clarity could be improved by putting the block number on the same row as the first haplotype and include an empty row between the blocks. Second, the first row of each block should be reserved for the reference haplotype, which are used to compute the odds ratios. Currently, the reference haplotype does not appear to be in the table. Third, the chi-square values can be deleted from the table as they are redundant given the presentation of the p-values. Fourth, another column should include the dbSNP ids for the SNPs in the block in the same order as they appear in the haplotype along with the alleles (major/minor) designated. Currently, the reader needs to refer to both the text and the LD figure to get this information. Finally, the Bonferroni thresholds are incorrectly stated. There should be a single Bonferroni threshold based on the 11 tests performed across all blocks. This should also be stated in the statistical methods. This also effects as statement made in the abstract, which should be altered. 5) The first full paragraph on page 10 of the results section states that there are three significant SNP associations, but it not clear if these results are independent of one another. The subsequent haplotype analysis also does not clarify this question. To answer this question, the authors should present a multi-SNP model of sarcoidosis risk. Starting with rs1049550 in the model, likelihood ratio tests should be performed for each of the other two SNPs. The most significant (if likelihood ratio $p<0.05$) should then enter the model, and the remaining third SNP should be tested in the same manner. This is the only way the reader can determine whether these SNPs are merely tagging one another or have independent added value. Based on the LD values, it looks like they may be independent, but regardless of the outcome, this is a critically important analysis. Also, like the others, it would be helpful to build the models both unadjusted and adjusted for age and sex. Minor Comments: 1) The Chi-square values are not necessary in the abstract. 2) The authors need to specify in the statistical methods the genetic

	model that was used to calculate the odds ratios. I suppose it was additive, but this needs to be made explicit. 3) The number of digits presented for proportions and 95% confidence intervals makes the tables a bit too busy for me. I would suggest truncating all proportions at one decimal place and the 95% confidence intervals at two. This should be a convention applied in the text as well. 4) In table 3, the upper and lower values of the 95% confidence intervals are separated by a "~". I would suggest using a "-" as was done in Table 2. 5) On line 37 of page 7, the authors state that "All participants were from a non-genetically related ..". I believe they want to state that all participants were genetically unrelated, correct? However, I don't think that this can be confirmed? Maybe it would be best to say that no familial relationship were known between the study participants. 6) There are some spelling mistakes and grammatical issues. The paper should be re-proofed.

VERSION 1 – AUTHOR RESPONSE

Reviewer #1:

Point 1: In the introduction and in table 1 the authors describe the published sarcoidosis association studies for ANXA 11. However, there is one study published recently, carried out in a Portuguese population, which has a different genetic background in comparison with the other studied European populations that is absent.

Response1: We have added this reference in table 1 and introduction section.

Point 2: The group of patients has a mean age of 53.6±4.6 years, which is particularly high taking into account the usual descriptions in different populations. The authors should address and explain this point.

Response 2: We have cross-checked the mean age of patients. Sarcoidosis most often occurs older than 45 years of age, with women being diagnosed more frequently than men. Our report is consistent with previous studies in Han chinese.

Reference 1: Yun Li, Stefan Pabst, Christian Kubisch, et al. First independent replication study confirms the strong genetic association of ANXA11 with sarcoidosis. Thorax. 2010 ;65(10): 939 - 940.

Reference 2: **王宇**, **曹悦鞍**, **彭朝**胜, **等**. 结节病患者155**例**临床分析. 临床肺科杂志, 2012, 17 (9): 1577-1578.

Point 3: Describe patients according with their demographic characteristics- How many are men?

Respondse 3: Four - hundred and twelve patients with sarcoidosis (155 men and 257 women) were recruited from our hospital. The control group consisted of 418 unrelated healthy subjects (160 men and 258 women) who underwent health examinations in the Medical Examination Center of the First Affiliated Hospital of the Xin'xiang Medical College. We have added it in subjects section (Page 7, lines 2-4; Page 7 lines 8-12).

Point 4: Although the sample size is presented as a limitation (page 4) no calculations for it are given.

Response 4: We performed power calculations for case-control genetic association analyses using PGA v2.0. Our sample size can detect SNP and haplotype associations with 90% and 85% power, respectively, at a false positive rate of 5%, disease prevalence of 1‰, disease allele/haplotype frequency of 0.05/0.03, and a presumed odds ratio (OR) of 1.5. We have added this description in dicussion section. According to Instructions for authors of the journal, "Strengths and limitations of this study" is needed. We have revised the description of limitations (Page 4, lines 21-22; Page 14, lines 13-17). Thank you very much!

Point 5: After a characterization in material and methods of different clinical presentations and evolutions, strangely in the results we don't see any stratification of the patients according with these different phenotypes. This is very important in order to understand the real categorization of the population and if any of the significant SNP associations observed could be associated with a

particular phenotype or a different outcome. The previous reports showed some different associations regarding for instance different type of clinical presentations or thoracic stage (Scadding criteria).
Response 5: We thank the reviewer for pointing out this important point. To assess particular disease phenotypes, further correlation analyses were carried out to explore potential association between ANXA11 genotype (rs2789679, rs1049550 and rs2819941) and chest radiographic (CXR) stages of sarcoidosis patients. The results revealed significant differences between stage I and stages II - IV for rs1049550. The significantly more CC genotype (P = 0.012) were found in the stages I of sarcoidosis patients. We have added the table 3 and the description in abstract, results and disccussion section (Page 2, lines 19-22; Page 11, lines 10-17; Page 12, lines 17-22). Thank you very much!
Point 6: The authors should give eligibility criteria for the controls - How were they recruited? If consecutively, when they were recruited? Have they been recruited according to age or sex?
Response 6: The criteria for the controls have been given. The control group consisted of 418 unrelated healthy subjects (mean ± SD age: 54.2 ± 5.3 years; 160 men and 258 women) who underwent health examinations in the Medical Examination Center of the First Affiliated Hospital of the Xin'xiang Medical College (Xinxiang, China) between Oct 2009 and Sept 2013. We have revised the subjects section (Page 7, lines 8-18). Thank you!
Point 7: Although, mentioned in the statistical analysis, no results from logistic regression are given.
Response 7: Unconditional logistic regression was used to calculate the odds ratio (OR) and 95% confidence interval (CI) in independent association between each locus and the presence of sarcoidosis. Generalized linear regression was used to evaluate the interaction effects between gene and gender or age. Gender and age of subjects were treated as covariants in binary logistic regression. We have revised table 2 and added this description in statistical analysis section (Page 9, lines 13-19). All results have been gived in table 2. Thank you very much!
Point 8: - Page 9, line 31 – "Gabriel et al. 14" – reference 14 correspond to Mrazek et al, 2011
Response 8: We apologize for this mistake. We have revised it.
Point 9: Page 10, line 41 – authors should consider rewriting the sentence "The significantly more G-G-C haplotypes (block 4) (P = 0.027) were (…)"
Response 9: We have deleted this sentence because of Global P > 0.05 (block 4) in revised manuscript.
Reviewer #2:
Point 1: For the description of the tagSNP selection, there seem to be some key features left out. First, the metric for selection (for example, r2) and the threshold value for determining what SNPs are tagged are not presented. Second, the authors state that tagSNPs were selected from both the 1000 genomes and HapMap, but they do not specify the number of tagSNPs selected from each source. Third, there is no mention of the total number of SNPs that are tagged in either the HapMap or the 1000 genomes data by the 15 tagSNPs that were selected. A complete description should list the total number and proportion of variants (at a minor allele threshold of 0.05) that are tagged by the selected tagSNPs. Fourth, the base pair boundaries of the tagSNP selection should also be included.
Response 1: We thank you for your recognition of our work and encouragement. Marker selection was done according to previous studies, and preliminary analysis was performed using the HapMap data and the following criteria. First, we examined tagSNPs in Haploview (v4.2), using the CHB population and a minor allele frequency cut - off (MAF) ≥ 5% (HapMap Data Release 27). We found that there were a total of 29 potential tagSNPs in all. As a first screen of the most common SNPs in the Eastern Han Chinese sarcoidosis sample, a MAF ≥ 20% with pair-wise tagging and r2 ≥ 0.8 was used as the cut - off when choosing tagSNPs. Second, the LD pattern of this gene was determined in the Chinese population using the preliminary data from HapMap. Five LD blocks across the ANXA11 were defined using Haploview's 'confidence intervals' method. 15 SNPs were further analyzed in an association study.
The rs2789679 and rs2245168 are located in 5' - UTR, rs2236558 in intron 6, rs7067644 and rs12763624 in intron 8, rs1049550 in exon10, rs2573351 in intron 13, rs10887581, rs11201989, rs2573353, rs2789695, rs2573356, rs2819945, rs11202059 and rs2819941 in intron 14. We have revised the SNP selection section (Page 8, lines 1-15).

Point 2: The authors have based the linkage disequilibrium (LD) block definitions on the combined samples of cases and controls. LD patterns can differ between a sample of the base population (i.e. the controls) and the case subjects. The authors have shown different haplotype frequencies between cases and controls within the blocks that reflect these differences in LD between cases and controls. To ensure that the LD blocks most closely reflect the population level LD patterns, definition of the blocks should be based on the control samples alone or alternatively on the original CHB HapMap samples. Actually, it's not clear to me exactly which samples were used to define the blocks. This should all be described in the methods and should be based on either the CHB or the controls from their sample.
Response 2: We apologize for missing description of block definitions and thank you for raising this question in time. In fact, Definition of the blocks were based on the control samples alone in our manuscript. We have revised figure 2 and described in the methods (Page 10, lines 1-3).
Point 3: Based on the statistical methods section, the authors do not appear to have evaluated whether adjustment for covariates such as age and sex influence the SNP odds ratios. The manuscript would be improved by evaluating whether adjustment for these important potential confounding variables alters the SNP odds ratio estimates and significance.
Response 3: Unconditional logistic regression was used to calculate the odds ratio (OR) and 95% confidence interval (CI) in independent association between each locus and the presence of sarcoidosis. Generalized linear regression was used to evaluate the interaction effects between gene and gender or age. Gender and age of subjects were treated as covariants in binary logistic regression. We have revised table 2 and added this description in statistical analysis section (Page 9, lines 13-19). All results have been gived in table 2. Thank you very much!
Point 4: In the statistical methods section, there is no description of the haplotype analyses performed. The readers will need to know what method was used, if there was adjustment for covariates, what genetic model used to estimate the haplotype odds ratios, how rare haplotypes were handled, ect. Also, the presentation of the haplotype results in Table 3 could be improved. First, it is not where the haplotype blocks and the corresponding haplotypes start and end. Clarity could be improved by putting the block number on the same row as the first haplotype and include an empty row between the blocks. Second, the first row of each block should be reserved for the reference haplotype, which are used to compute the odds ratios. Currently, the reference haplotype does not appear to be in the table. Third, the chi-square values can be deleted from the table as they are redundant given the presentation of the p-values. Fourth, another column should include the dbSNP ids for the SNPs in the block in the same order as they appear in the haplotype along with the alleles (major/minor) designated. Currently, the reader needs to refer to both the text and the LD figure to get this information. Finally, the Bonferroni thresholds are incorrectly stated. There should be a single Bonferroni threshold based on the 11 tests performed across all blocks. This should also be stated in the statistical methods. This also effects as statement made in the abstract, which should be altered.
Response 4: The haplotype frequencies were estimated using GENECOUNTING in revised manuscript, which computes maximum-likelihood estimates of haplotype frequencies from unknown phase data by utilizing an expectation-maximization algorithm. The significance of any haplotypic association was evaluated using a likelihood ratio test, followed by permutation testing that compared estimated haplotype frequencies in cases and controls. The Global P value of block1, block2, block3 and block 4 are respectively 0.075, 0.003, 0.045 and 0.221. We have revised table 4 and table 5. In the statistical methods section, we have added the description of the haplotype analyses performed (Page 10, lines 3-8). Thank you!
Point 5: The first full paragraph on page 10 of the results section states that there are three significant SNP associations, but it not clear if these results are independent of one another. The subsequent haplotype analysis also does not clarify this question. To answer this question, the authors should present a multi-SNP model of sarcoidosis risk. Starting with rs1049550 in the model, likelihood ratio tests should be performed for each of the other two SNPs. The most significant (if likelihood ratio $p<0.05$) should then enter the model, and the remaining third SNP should be tested in the same manner. This is the only way the reader can determine whether these SNPs are merely tagging one

another or have independent added value. Based on the LD values, it looks like they may be independent, but regardless of the outcome, this is a critically important analysis. Also, like the others, it would be helpful to build the models both unadjusted and adjusted for age and sex.
Response 5: Comparison of genotype and allele frequency distribution revealed significant differences between the patients with sarcoidosis and healthy controls for 3 SNPs: rs2789679, rs1049550 and rs2819941. We presented a multi-SNP model of sarcoidosis risk. Starting with rs1049550 in the model, likelihood ratio tests been performed for each of the other two SNPs. However, these results are independent of one another (P>0.05). We have added the description in results section (Page 11, lines 3-4). Thank you very much!
Point 6: The Chi-square values are not necessary in the abstract.
Response 6: We have deleted the Chi-square values in the abstract (Page 2, lines 15-18).
Point 7: The authors need to specify in the statistical methods the genetic model that was used to calculate the odds ratios. I suppose it was additive, but this needs to be made explicit.
Response 7: We have added the descrption in statistical methods (Page 9, lines 13-19). Thank you!
Point 8: The number of digits presented for proportions and 95% confidence intervals makes the tables a bit too busy for me. I would suggest truncating all proportions at one decimal place and the 95% confidence intervals at two. This should be a convention applied in the text as well.
Response 8: We have revised it in the text and table 2, table 4 and table 5.
Point 9: In table 3, the upper and lower values of the 95% confidence intervals are separated by a "~". I would suggest using a "-" as was done in Table 2.
Response 9: Table 3 in previous study has been changed to table 4 and table 5 in revised manuscript.
Point 10: On line 37 of page 7, the authors state that "All participants were from a non-genetically related ..". I believe they want to state that all participants were genetically unrelated, correct? However, I don't think that this can be confirmed? Maybe it would be best to say that no familial relationship were known between the study participants.
Response 10: We have revised it (Page 7, lines 17-18).
Point 11: There are some spelling mistakes and grammatical issues. The paper should be re-proofed.
Response 11: The manuscript has beed edited by native English speakers.

VERSION 2 – REVIEW

REVIEWER	Albert M. Levin, Ph.D. Henry Ford Health System United States of America
REVIEW RETURNED	17-Mar-2014

GENERAL COMMENTS	1. In the abstract, the last sentence of the Results (starting on page 3 line 4 of the marked version) now follows the chest x-ray case-only results. As I believe the haplotype results are from the case-control (i.e. risk) analyses, they should be moved up to follow the single SNP case control analyses. Otherwise, the reader wouldn't know which outcome the haplotype results correspond to. This re-ordering appears to be consistent with the way the results section is laid out. 2. The SNP selection description (on page 8 from lines 7 to 22 of the marked version) is too confusing. First, the sentence starting on line 11 "Marker selection was done according to previous studies." makes it sound like the 15 markers were selected as part of these previous studies, which they weren't. Second, based on the allele frequencies from Table 2, only SNPs with minor allele frequencies ≥ 0.2 in the HapMap CHB were genotyped. So, it's confusing why the authors are talking about a minor allele frequency threshold of 0.05 at all. I would suggest re-structuring this paragraph to make the following points:

	a. We used the CHB data from the HapMap (release 27) to select tagSNPs for ANXA11 b. We restricted our search for tagSNPs from base pair XXX to YYY. c. Further, we limited the SNPs to tag to those with a minor allele frequency ≥ 0.2 in the CHB d. Based on these restrictions, there were a total of ZZ SNPs. e. Using HAPLOVIEW with an r2 threshold ≥ 0.8, 15 tagSNP were selected and used for subsequent analyses. 3. In the revised Statistical Analysis section (page 10 of the marked manuscript), the sentence beginning on line 7 "Generalized linear regression was used to evaluate the interaction effects between gene and gender and age" is confusing. Logistic regression is what was used, correct? If so, there isn't any need to call logistic regression something else. Please just stick with logistic regression. Second, I don't see where any interaction tests were done at all. There is a statement on page 12 (line 6) that there were no interactions among rs2789675, rs1049550, and rs2819941. Are these interactions with sex and/or age? Are these gene*gene interactions? In my original review, I ask for the following re-analysis to determine whether the three SNPs (rs2789675, rs1049550, and rs2819941)had independent effects: "… the results section states that there are three significant SNP associations, but it is not clear if these results are independent of one another. The subsequent haplotype analysis also does not clarify this question. To answer this question, the authors should present a multi-SNP model of sarcoidosis risk. Starting with rs1049550 in the model, likelihood ratio tests should be performed for each of the other two SNPs. The most significant (if likelihood ratio $p<0.05$) should then enter the model, and the remaining third SNP should be tested in the same manner. This is the only way the reader can determine whether these SNPs are merely tagging one another or have independent added value. Based on the LD values, it looks like they may be independent, but regardless of the outcome, this is a critically important analysis." Based on my reading of the current revised manuscript, I believe that the authors haven't done this analysis. I still feel that this is a critical analysis that needs to be done. To be clear, I am suggesting a forward-stepwise SNP selection procedures be conducted, and the final model be reported in the results section. 4. When this paper was re-uploaded, the tables were not formatted properly (they were all squished), which made it extremely difficult to read Table 2. It looks like the table was simply reformatted to be tall rather than wide, but I couldn't adequately check the consistency of the results given the formatting. 5. In the section on page 12 describing the haplotype results, there is no reference to the Tables that contain the haplotype results (now Tables 4 and 5). In addition to referring to the tables in the text containing the haplotype results, these tables should be Table 3 and 4, and the chest radiographic results should be changed to Table 5, as the haplotype results come before the chest radiographic results. Finally, I am not sure what happened to the results for blocks 1 and

	4? In my original review, I asked for the following: "..Also, the presentation of the haplotype results in Table 3 could be improved. First, it is not clear where the haplotype blocks and the corresponding haplotypes start and end. Clarity could be improved by putting the block number on the same row as the first haplotype and include an empty row between the blocks. Second, the first row of each block should be reserved for the reference haplotype, which are used to compute the odds ratios. Currently, the reference haplotype does not appear to be in the table." This is just a simple reformatting of the original Table 3. I would suggest the removal of the current Tables 4 and 5, and reformat as I had suggested preserving all of the haplotype results, which any reader will want to see regardless of the level of significance. 6. There are still some residual language issues to be dealt with as well as a number that were introduced in the revision.

VERSION 2 – AUTHOR RESPONSE

Reviewer #2:
Point 1: In the abstract, the last sentence of the Results (starting on page 3 line 4 of the marked version) now follows the chest x-ray case-only results. As I believe the haplotype results are from the case-control (i.e. risk) analyses, they should be moved up to follow the single SNP case control analyses. Otherwise, the reader wouldn't know which outcome the haplotype results correspond to. This re-ordering appears to be consistent with the way the results section is laid out.
Response: In abstract, this sentence has been moved up to follow the single SNP case control analyses.
Point 2: The SNP selection description (on page 8 from lines 7 to 22 of the marked version) is too confusing. First, the sentence starting on line 11 "Marker selection was done according to previous studies." makes it sound like the 15 markers were selected as part of these previous studies, which they weren't. Second, based on the allele frequencies from Table 2, only SNPs with minor allele frequencies ≥ 0.2 in the HapMap CHB were genotyped. So, it's confusing why the authors are talking about a minor allele frequency threshold of 0.05 at all. I would suggest re-structuring this paragraph to make the following points:
a. We used the CHB data from the HapMap (release 27) to select tagSNPs for ANXA11
b. We restricted our search for tagSNPs from base pair XXX to YYY.
c. Further, we limited the SNPs to tag to those with a minor allele frequency ≥ 0.2 in the CHB
d. Based on these restrictions, there were a total of ZZ SNPs.
e. Using HAPLOVIEW with an r2 threshold ≥ 0.8, 15 tagSNP were selected and used for subsequent analyses.
Response: Thank you very much! We have revised the criteria of marker selection according your points.
Point 3: In the revised Statistical Analysis section (page 10 of the marked manuscript), the sentence beginning on line 7 "Generalized linear regression was used to evaluate the interaction effects between gene and gender and age" is confusing. Logistic regression is what was used, correct? If so, there isn't any need to call logistic regression something else. Please just stick with logistic regression. Second, I don't see where any interaction tests were done at all. There is a statement on page 12 (line 6) that there were no interactions among rs2789675, rs1049550, and rs2819941. Are these interactions with sex and/or age? Are these gene*gene interactions?
Response: We have revised the description about logistic regression in statistical Analysis section.

We presented a multi-SNP model of sarcoidosis risk. We screen out the most significant model to predict the risk by forward stepwise strategy in logistic regression, and found that a model includes rs2789679 and rs1049550 is the best one (Supplemental table 2), which suggests that there is a interaction between them. A new Supplemental table 2 has been added. According to the instructions for authors of journal, the article does not exceed 4000 words, with up to five figures and tables. Than you very much!

Point 4: When this paper was re-uploaded, the tables were not formatted properly (they were all squished), which made it extremely difficult to read Table 2. It looks like the table was simply reformatted to be tall rather than wide, but I couldn't adequately check the consistency of the results given the formatting.

Response: Table 1 and Supplemental Table 1 in revised manuscript has been formatted properly (According to the instructions for authors of journal, Table 1 can't exceed the limit of 2 pages). Than you very much!

Point 5: In the section on page 12 describing the haplotype results, there is no reference to the Tables that contain the haplotype results (now Tables 4 and 5). In addition to referring to the tables in the text containing the haplotype results, these tables should be Table 3 and 4, and the chest radiographic results should be changed to Table 5, as the haplotype results come before the chest radiographic results. Finally, I am not sure what happened to the results for blocks 1 and 4?

Response: The haplotype frequencies were estimated using GENECOUNTING, which computes maximum - likelihood estimates of haplotype frequencies from unknown phase data by utilizing an expectation - maximization algorithm. The significance of any haplotypic association was evaluated using a likelihood ratio test, followed by permutation testing that compared estimated haplotype frequencies in cases and controls. About the haplotype calculation, We asked the professor of molecular genetics. We analyzed haplotype frequencies in block 2, 3 ($P = 0.003$ in block 2; $P = 0.045$ in block 3) but not block 1 and 4. Because Global $P > 0.05$ in block 1 and 4 ($P = 0.075$ in block 1; $P = 0.221$ in block 4).

Referecne

1. Curtis D, Knight J, Sham PC. Program report: GENECOUNTING support programs. Ann Hum Genet 2006;70:277-9.
2. Zhao JH, Curtis D, Sham PC. Model-free analysis and permutation tests for allelic associations. Hum Hered 2000;50:133-9.
3. Zhao JH, Lissarrague S, Essioux L, et al. GENECOUNTING: haplotype analysis with missing genotypes. Bioinformatics 2002;18:1694-5
4. Zhao JH, Sham PC. Generic number systems and haplotype analysis. Comput Methods Programs Biomed 2003;70:1-9

Point 6: There are still some residual language issues to be dealt with as well as a number that were introduced in the revision.

Response: The manuscript has beed corss-edited by native English speakers again.

VERSION 3 - REVIEW

REVIEWER	Albert M. Levin, PhD Henry Ford Health System United States of America
REVIEW RETURNED	29-Apr-2014

GENERAL COMMENTS	1. There seems to be some confusion with the term "interaction" in this paper. As described in the Methods (page 10 lines 11-13), when the main effects of age and gender are included in a model along with the effect of the SNP, there is no test of interaction being performed. One is simply adjusting for the confounding effects of

	age and gender. I think it would be fine to simply remove the sentence starting on line 11 and ending on line 12 stating “Logistic regression was used to evaluate the interaction effects between gene and gender or age.”, as the following sentence correctly states that they were included as covariates. 2. The confusion with the term “interaction” continues in the Results starting on page 11 (line 21) and ending on page 12 (line 2), where the authors present results from a multi-snp modelling approach. First, the inclusion of significant main effects does not indicate interaction. On the contrary, if the SNPs are included in the model and remain significant, this indicates independence not interaction. If the authors want to test for interaction between SNPs, these need to be specifically modelled as interaction terms and appropriately test. Second, the multi-snp forward stepwise modelling strategy needs to be described in the Methods section of the paper, including the criteria for inclusion and exclusion from the model as well as the statistical test to evaluate significant differences between models. It is insufficient to state that a particular model is best without justifying stating the statistical criterion for making this decision. Third, were age and gender included in these models? Again, this needs to appear in the methods. Third, in addition to the p-value for the final model (presented in Supplemental Table 3), the important information gained from this analysis are the SNP additive model odds ratios, 95% confidence intervals, and p-values from the multi-snp model, as the readers of this paper will want to compare the unadjusted odds ratios from Table 1 to the multi-snp adjusted odds ratios. Therefore, I would suggest replacing Supplemental Table 3 with a table that includes this information from the final best fitting model only. If at all possible, I also highly suggest that the authors consult a statistician when making the addition of the modelling methods to the paper. 3. Again, the misuse of the term “interaction” persists in the Discussion on page 13 (line 15), when the authors state they found an interaction between rs2789679 and rs1049550. As suggested above in comment #2, the results of the final model need to be presented. If the associations persist after adjustment for one another as main effects in the model, the conclusion would be that there are two independent genetic association in ANXA11 that are associated with sarcoidosis risk in the Han Chinese. If the authors want to test for interaction between the SNPs, this is an additional analysis that the authors need to add.

VERSION 3 – AUTHOR RESPONSE

Reviewer #2:

Point 1: There seems to be some confusion with the term “interaction” in this paper. As described in the Methods (page 10 lines 11-13), when the main effects of age and gender are included in a model along with the effect of the SNP, there is no test of interaction being performed. One is simply adjusting for the confounding effects of age and gender. I think it would be fine to simply remove the sentence starting on line 11 and ending on line 12 stating “Logistic regression was used to evaluate the interaction effects between gene and gender or age.”, as the following sentence correctly states that they were included as covariates.

Response 1: We have deleted the descrption “Logistic regression was used to evaluate the interaction effects between gene and gender or age.” in methods. We have added the descrption

“Models of multiple logistic regression were used to test the independence of individual allelic effect. In detail, the most significant SNP was chosen to be the conditional SNP (covariate in the regression model) when testing other significant SNPs; Meanwhile, a multi-SNP model including all significant SNPs was also performed. ” in method part. Thank you very much!

Point 2: The confusion with the term “interaction” continues in the Results starting on page 11 (line 21) and ending on page 12 (line 2), where the authors present results from a multi-snp modelling approach. First, the inclusion of significant main effects does not indicate interaction. On the contrary, if the SNPs are included in the model and remain significant, this indicates independence not interaction. If the authors want to test for interaction between SNPs, these need to be specifically modelled as interaction terms and appropriately test. Second, the multi-snp forward stepwise modelling strategy needs to be described in the Methods section of the paper, including the criteria for inclusion and exclusion from the model as well as the statistical test to evaluate significant differences between models. It is insufficient to state that a particular model is best without justifying stating the statistical criterion for making this decision. Third, were age and gender included in these models? Again, this needs to appear in the methods. Third, in addition to the p-value for the final model (presented in Supplemental Table 3), the important information gained from this analysis are the SNP additive model odds ratios, 95% confidence intervals, and p-values from the multi-snp model, as the readers of this paper will want to compare the unadjusted odds ratios from Table 1 to the multi-snp adjusted odds ratios. Therefore, I would suggest replacing Supplemental Table 3 with a table that includes this information from the final best fitting model only. If at all possible, I also highly suggest that the authors consult a statistician when making the addition of the modelling methods to the paper.

Response 2: We have revised Table 1 and Supplemental table 3. The multi-SNP model showed only rs1049550 present a significant effect on the disease phenotype ($p < 0.001$). No independent effect was found for rs2789679 or rs2819941 when adjusting for the effect of rs1049550. Both individual SNP and multi - SNP analysis supported rs1049550: T allele was an important protective factor for affecting sarcoidosis (odds ratio around 0.6). In another word, carriers of rs1049550: C allele have higher susceptibility to sarcoidosis in our Chinese Han population (Supplemental table 3). These descrption has been added in results section.

Point 3: Again, the misuse of the term “interaction” persists in the Discussion on page 13 (line 15), when the authors state they found an interaction between rs2789679 and rs1049550. As suggested above in comment #2, the results of the final model need to be presented. If the associations persist after adjustment for one another as main effects in the model, the conclusion would be that there are two independent genetic association in ANXA11 that are associated with sarcoidosis risk in the Han Chinese. If the authors want to test for interaction between the SNPs, this is an additional analysis that the authors need to add.

Response 3: We’re sorry for the misunderstanding and misuse of “interaction effect” in previous submission. We re-analyzed those three significant SNPs by multi-SNP modeling you suggested in previous comment, and clearly described it in the method part. Gender and age were not includes in the regression model since they were forced to be matched between cases and controls in our study design. As shown in revised supplementary table 3, only one SNP showed significant effect, hence no interaction analysis was conducted for any pairs of SNPs. In addition, we realized that our sample size for detecting interaction effect may be not enough, since it was usually necessary to have four times as large as that needed for main effect.

The description “Since causal SNP in ANXA11 was not known, all the SNPs significantly associated with sarcoidosis were actually indirect association through the real causal one. Though rs2789679 and rs2819941 were not in LD with rs1049550, it may not tag the causal SNP as well as rs1049550. Hence, their individual effects were eliminated when controlling for the effect of rs1049550.” has been added in discussion.

VERSION 4 - REVIEW

REVIEWER	Albert M. Levin, PhD Henry Ford Health System, Detroit, Michigan United States of America
REVIEW RETURNED	13-Jun-2014

GENERAL COMMENTS	1. The multi-snp model now clearly reveals that rs1049550 is the only independent SNP association effect after accounting for the other two marginally associated SNPs. This should be stated in the abstract. I would suggest removing the statement "The rs2789679 A allele (P = 0.00004, OR = 19 1.42, 95%CI = 1.17-1.73) and rs2819941 T allele (P = 0.0006, OR = 1.41, 95%CI = 20 1.16-1.71) were significantly more frequent in the patients with sarcoidosis compared 21 to controls.", and emphasize in the following sentence that rs1049550 was the most significant of the three, followed by another sentence stating that a multi-snp model revealed rs1049550 to be the only significant independent SNP association with sarcoidosis risk. 2. In the discussion, the authors now talk about the possibility of a single causal SNP that is possibly best tagged by rs1049550. In this context, the authors should list their SNP selection as a limitation, as it restricted the SNPs examined to only those with an allele frequency >=0.2. In other words, a broader examination of the genetic variation in ANXA11 in the Han Chinese may reveal other variants associated with disease risk.

VERSION 4 – AUTHOR RESPONSE

Reviewer #2:
Point 1: The multi-snp model now clearly reveals that rs1049550 is the only independent SNP association effect after accounting for the other two marginally associated SNPs. This should be stated in the abstract.
Response 1: We have added this sentence in abstract. Thank you very much!
Point 2: I would suggest removing the statement "The rs2789679 A allele (P = 0.00004, OR = 1.42, 95%CI = 1.17-1.73) and rs2819941 T allele (P = 0.0006, OR = 1.41, 95%CI = 1.16-1.71) were significantly more frequent in the patients with sarcoidosis compared to controls.", and emphasize in the following sentence that rs1049550 was the most significant of the three, followed by another sentence stating that a multi-snp model revealed rs1049550 to be the only significant independent SNP association with sarcoidosis risk.
Response 2: We have deleted this sentence. We have added the description " The rs1049550 was the most significant of the three. A multi-SNP model revealed rs1049550 to be the only significant independent SNP association with sarcoidosis risk." Thank you very much!
Point 3: In the discussion, the authors now talk about the possibility of a single causal SNP that is possibly best tagged by rs1049550. In this context, the authors should list their SNP selection as a limitation, as it restricted the SNPs examined to only those with an allele frequency >=0.2. In other words, a broader examination of the genetic variation in ANXA11 in the Han Chinese may reveal other variants associated with disease risk.

Response 3: We have added the description "A broader examination of the genetic variation in ANXA11 in the Han Chinese may reveal other variants associated with disease risk.". in discussion.