Author's response to reviews

Title: Association of common variants in JAK2 gene with reduced risk of Metabolic Syndrome and related disorders

Authors:

Alberto Penas-Steinhardt (pufetin@gmail.com)
Mariana L Tellechea (marianatellechea78@hotmail.com)
Leonardo Gomez-Rosso Dr (mulag@hotmail.com)
Fernando Brites Dr. (fbrites@hotmail.com)
Gustavo D Frechtel Dr (gfrechtel@ffyb.uba.ar)
Edgardo Poskus Dr (eposkus@ffyb.uba.ar)

Version: 2 Date: 15 September 2011

Author's response to reviews: see over
Cover Letter

Dear Tim Sands

Executive Editor, BMC Medical Genetics

Please find the manuscript entitled "Association of common variants in JAK2 gene with reduced risk of Metabolic Syndrome and related disorders" (MS: 7788175805587496) attached for your reconsideration to be published in BMC Medical Genetics.

We really appreciate that the article has been revised by external reviewers and that they have found it of potential interest, and are really grateful for all criticisms. We have modified the article taking into account the suggestions of the referees. We provide a point-by-point response, indicating all the changes made in the manuscript or a rebuttal of each point, which you will find appended below.

We really hope the changes made will meet the expectations of the referees and that you will reconsider the publication of the revised paper after reevaluation by the referees.

Sincerely yours,

Dr Penas-Steinhhardt
Corresponding author
Response to Reviewers

Response to Reviewer #1

*Major limitations*

-The sample size is relatively small included self-reported European ancestry mainly from Spain and Italy. This implies that some degree of genetic structure can exist within the sample, affecting the genetic association study. Differences in allele frequency by means of ancestry origin should be performed, given that is not possible to perform a principal component analysis. Ancestry origin can also be included as covariable in the study.

We are aware of mayor limitations of our study design. The study was conducted in a sample of 724 young male subjects, 18–65 years of age (mean age 37). The sample size is relatively small but for each individual case-control study power estimations were performed for single-point allelic effects\(^1\), with an odds ratio of 1.5 at a nominal significance level of 0.05. Power estimation to detect effects on HTG, HW, obesity or abdominal obesity, power estimation was found to be between 81 and 99% assuming a dominant model and less than 80% assuming a recessive model for rs7849191; and between 80 and 98% assuming a recessive model and less than 80% assuming a dominant model for rs3780378. The power estimation was found to be between 85 and 99% to detect effects of rs7849191 and rs3780378 on decreased HDL-C. Unfortunatelly, power estimation was found to be less than 80% to detect effects on Metabolic Syndrome.
Since most current studies are underpowered to achieve such a stringent level of significance, replications are usually necessary for the confirmation of an association finding.

Individuals were randomly recruited at the Department of Haemotherapy of the José de San Martín Hospital of the University of Buenos Aires in the context of a cross-sectional study conducted between April 2006 and April 2008. We entirely agree that differences in allele frequency by means of ancestry origin should be performed but, we already stated that the present study provided only exploratory results and should be confirmed in a second and independent replication cohort (line 308, “This study has some limitations, the sample size was relatively small, this data should be viewed as a preliminary source for future studies including prospective studies with larger sample size and different ethnic groups“). We would appreciate that other research groups either replicate our findings in a proper and well powered sample.

We are not able to adjust for ancestry origin because individuals who have not European ancestry, from Spain and Italy (self-reported) were excluded from the study.

Finally, we should add that most molecular association studies did not be include ancestry origin as covariable.

-It seems that no quality control has been applied to genetic data i.e. duplicated samples or genotyping rates.

We already stated in the revised manuscript that (line 157) genotyping accuracy was assessed by inclusion of duplicates and negative controls. Genotyping success rate was 100% for rs7849191 and 98.7% rs3780378. Our lab randomly sampled 10% of the subjects for sequencing to corroborate the initial findings. Using these criteria consistency was 100%.
-It appears that only 707 individuals were genotyped for the rs7849191 and 724 for the rs3780378.

We already modified the revised manuscript, abstract and main document line 94, according to this observation (“The total sample includes 724 unrelated men”).

-It is not mentioned if lipid lowering medication has been taken into account.

Lipid lowering medication has been taken into account to ask for the presence of MS. We stated that each subject was assessed for the presence of MS using the American Heart Association/National Heart, Lung, and Blood Institute (AHA/NHLBI) 2005 criteria \(^2\). This statement allow for triglycerides, HDL-C levels, and blood pressure to be counted as abnormal when a person is taking drug treatment for these factors.

-No information is provided regarding the distribution of phenotypic data. The authors used ANOVA for normal-distributed and Kruskal-Wallis otherwise. In addition, linear regression analysis was used. It is not clear in table 3 which one has been applied to each analysis.

For comparison of continuous variables we conducted one-way ANOVA with Levine’s Test for equality of variances; otherwise, we used Welch tests, as stated in Statistical Analysis section of the revised manuscript. In Table 3 Leyend, it is indicated when Welch was used (“\(c\) = Welch p value”). Multiple linear regression analyses was used to adjust for possible confounding
variables, as stated in Statistical Analysis section and, in Table 3 Leyend, it is indicated when $p$ is adjusted ("$^a$ Age adjusted; $^b$ Age and BMI adjusted"). Furthermore, we already add in the revised manuscript Table 3 Leyend that we performed ANOVA for normal-distributed data and Welch otherwise, and multiple linear regression analyses was used to adjust for possible confounding variables.

- *Is the value $p=0.001$ for the df comparison at rs3780378 derived from and age and bmi adjusted analysis? (it has a $b$ superscript but in the column $p$ $b$ no significance is denoted).*

We already modified the revised manuscript, Table 3 (in the column $p$, all values has a $c$ superscript), according to this observation.

- *For Kruskal-Wallis analyses, median and range should be described rather than mean and SD.*

We entirely agree that, for Kruskal-Wallis analyses, median and range should be described rather than mean and SD, but we really believe that it is worthy for the reader to have a look to the mean and SD values. For this reason, we conducted Welch test, as stated in Statistical Analysis section and we reported Welch test $p$ values in Table 3 of the revised manuscript. We believe that reporting both median and range, and mean and SD would make Table 3 long and probably difficult to understand.
Given the strong dependency of the analyzed traits from age and BMI, I think that is more appropriate to report the adjusted linear regression models with transformed phenotypic values (i.e., log transformation) for phenotypes deviating from normality.

Multiple linear regression analyses was used to adjust for possible confounding variables. We performed linear regression analysis with log transformed values for phenotypes deviating from normality and modified the revised manuscript Table 3 p adjusted values.

- For the haplotype analysis, in order to reduce the degree of freedom, the TT haplotype, with a frequency less than 5%, could be removed from the analysis.

A Fisher exact test was used to compare haplotype frequencies. The effects of a particular haplotype load (0: no copies of the particular haplotype; 1: 1 copy; and 2: 2 copies) on continuous variables were tested using linear regression. Each haplotype was compared with haplotype TC. In order to reduce the degree of freedom, the TT haplotype (with a frequency less than 5%) was removed from the analysis. This was not previously stated in the submitted manuscript but it was already added in the corrected version.

- Minor allele frequencies for rs7849191 are bad described: for the TT genotype the frequency is 0.15 are not 0.015.

We already modified the revised manuscript according to this observation (Genotypic frequencies were as follow: rs7849191: CC 0.403 CT 0.446 and TT 0.15).
- The paper from Ge et al. (ref. 17) is referred throughout the report as Dongliang GE et al.

We already modified the revised manuscript according to this observation (In our study we genotyped only the most significant variants of a gene wide association according the study conducted by Ge et al).

-Table 1. LAP definition is absent

We indicated in Table 1 legend of the revised version of the manuscript LAP definition: (DBP, diastolic blood pressure; FPG, fasting plasma glucose; HDL-C, high-density lipoprotein cholesterol; HOMA-IR, homeostasis model assessment of insulin resistance; HW, hypertriglyceridemic waist; IR, insulin resistance; LDL-C, low-density lipoprotein cholesterol; SBP, systolic blood pressure; TC, total cholesterol; TG, triglycerides; LAP = lipid accumulation product; WC, waist circumference; HTG=Hypertriglyceridemia; HW=hypertriglyceridemic waist; IFG = impaired fasting glucose; MS = metabolic syndrome).

-Using Hapmap rel 28 and haplowlview 4.2, the selected SNPs capture a 21% of SNPs with MAF#10 at r^2>0.8. Rs3780378 captures variability at 19 SNPs.

The selection of tagSNPs for genotyping in the JAK2 gene was undertaken through the use of the International HapMAP Project (HapMap Data Rel 28 PhaseII+III, August10). The entire JAK2 gene was covered. tSNPs were selected using the pairwise algorithm based on the CEU
population (Utah residents with ancestry from northern and western Europe). A criterion of \( r^2 > 0.8 \) was used and SNPs had a minor allele frequency (MAF) of at least 0.1.

Actually two tagSNPs capture 19 of 95 (20%) of alleles on the full length JAK2 gene (142.8 kb). rs7849191 is not in LD with other SNPs and rs3780378 is in LD with other 17 SNPs (rs2230724, rs4372063, rs10815144, rs7037207, rs2149556, rs1328918, rs7034753, rs7875908, rs10115312, rs7023146, rs7857730, rs7847294, rs6476939, rs7032785, rs7043371, rs1328917, rs3780378).
Response to Reviewer #2

- Minor Essential Revisions

Following are a few typing mistakes (along with numbered lines picked from the submitted manuscript) suggested for correction before being published.

I. 143 Dongliang et all.

Correction suggested: Dongliang et al

II. 143 The LD analysis of the 142.8 kb under study

Correction suggested: The LD analysis of the 142.8 kb region under study

III. 291 The results found in this work are related to the rol of inflammation in the development

Word Spelling: The results found in this work are related to the role of inflammation in the development

IV. 306 be viewed as a preliminary source for future studies including prospective studies whit

Word Spelling: 306 be viewed as a preliminary source for future studies including prospective studies with

V. Table 2 Association between individuals JAK2 SNPs and MS and related phenotypes.

Correction suggested: Table 2 Association between individual JAK2 SNPs, Metabolic Syndrome and related phenotypes.

VI. All references need to be arranged according to the guidelines set by the journal.

For example;


Correction Suggested:


Correction Suggested:


We entirely agree with Reviewer #2. We have modified the article taking into account his suggestions indicating all the changes made in the manuscript.
