Author's response to reviews

Title: Association between a rare SNP in the second intron of human Agouti related protein gene and increased BMI

Authors:

Ivo Kapa (ivo@biomed.lu.lv)
Ineta Kalnina (ineta@biomed.lu.lv)
Valdis Pirags (valdis.pirags@stradini.lv)
Vita Ignatovica (vita@biomed.lu.lv)
Helgi B Schioth (helgi.schioth@neuro.uu.se)
Janis Klovins (klovins@biomed.lu.lv)

Version: 3 Date: 12 February 2009

Author's response to reviews: see over
To Editors of BMC Medical Genetics

Regarding manuscript ID 1852391747232845

Thank you for your evaluation of our manuscript and the valuable comments from the referees. We are very pleased that the referees have found this work interesting and that they have carefully analysed different aspects of the study.

We enclose a revised manuscript and a detailed list of the changes made, and response to each of the comments raised by the referees.

We hope that with these changes the manuscript will now be acceptable for publication in BMC Medical Genetics.

Best regards,

______________________

Janis Klovins Ph.D.
We thank the referee for the comments that work contains interesting findings, particularly in possible microRNA research field.

Answers to the following comments:

**Referee’s comment:** “However, my major problem is because this variant is so RARE, there are no reported homozygotes in the study group, and very few actual carriers. I appreciate that the association does reach statistical significance, however, the authors will need to increase the numbers of their study. Perhaps they can collaborate with other European cohorts to do this? I realize that asking for 'more numbers' might seem unfair. But I think that if the authors believe their findings, it is a worthwhile exercise. In particular, it will be interesting to see the phenotype of a patient that is homozygous for this SNP, considering its possible role in a known miRNA.”

**Answer:** We agree that disadvantage of present study is small sample size. Since it is not possible to increase the sample size in frame of this project, we have performed permutation test to limit type I error. We applied ‘label swapping’ based permutation procedure that is implemented in the PLINK software that generates both uncorrected and corrected empirical P-values. We used corrected empirical P-values that calculate the proportion of permutation in which any of the test statistics exceeds particular observed statistic and it is more stringent than uncorrected P-values, since it controls for the probability of at least one false positive per experiment. In all cases, the permuted P-values were significant and we are confident that these results strengthen the manuscript. We also add power calculation to the present manuscript to demonstrate the actual power of the analysis performed.

We include following description in Methods: “Power calculations were performed using Quanto v.1.2.3 [18]. Our sample size provided 80% power (at α=0.05) to detect difference between genotypes in logBMI of >0.053 assuming minor allele frequency of 0.012.”

The following text is included in Results and discussion section: “The power analyses show that the present sample size had 53% power to detect changes in our logBMI of 0.04.”

We agree that it would be interesting to see the phenotype of the homozygous patients as it is noted in the manuscript. However, even with substantially larger sample size and increased BMI for the TT homozygotes, it would be difficult to statistically prove the additive model.

We also hope that publishing of these results would encourage others to perform replication studies of this SNP.
Referee #2:

We thank the Referee for the critical comments and careful inspection of the results.

Answers to the following comments:

1. Referee’s comment: “Three SNPs have been detected in this sample, and the authors choose to study the less variable of them for which no variant homozygote was observed. Why they did not analyzed also one of the two other SNP that are in complete linkage disequilibrium with each other?”

It would be important to analyze SNPs that have been reported previously to be associated with body composition such as Ala67Thr. Did the authors detect this relatively frequent variant in their sample of 95 patients?

1. Answer: The rs5030980, with MAF=0.045, was found in 95 sequenced patients and corresponds to Ala67Thr as described in the “introduction” and “results and discussion” sections. This particular association analysis was performed to evaluate the role of a rare SNP’s found in resequencing of specific cohorts of patients. We feel that analysis of rare SNP’s are often underrepresented compared to common SNP’s, while the combination of rare SNP’s may have great impact on the genetics of common diseases. Although we agree that replication studies are of great importance, the main goal of this study was to analyse SNP that was not previously studied with respect to obesity. To fully evaluate the impact of particular gene locus, we would rather consider including number of tagSNPs based on LD analysis in association study, but such analysis was left beyond the scope of the project due to financial and technical reasons.

2. Referee’s comment: “In the Abstract, what is the meaning of “carriers of the rs11575892”? All subjects carried this SNP. Also, the authors could not conclude that they reported for the first time an association of AGRP with BMI, two previous reports made it earlier, Argyropoulos (2002) with late-onset obesity, and Marks (2004) with different body composition phenotypes.”

2. Answer To avoid misunderstanding we replace “carriers of the rs11575892 polymorphism” with “carriers of the rs11575892 T allele” in the Abstract. We would like to argue that according to our knowledge it is the first time when minor allele of AGRP polymorphism is associated with increased BMI. Both the articles mentioned by the referee describe association of either major allele of rs5030980 (Ala67) with obesity or homozygotes of minor allele (Thr67) with number of leanness related phenotypes. We would also like to point that Referee #3 agrees on the fact that this is the first time when SNP in AGRP is associated with increased BMI.
3. **Referee's comment:** "In Methods, it makes no sense to use an ANOVA knowing the importance on BMI of covariate effects such as age and gender."

3. **Answer:** We agree that age and gender should be obligatory included as covariates when BMI is analysed. However we think that inclusion of uncorrected ANOVA results would give more information to evaluate effects of these covariates on association of SNP with BMI. Besides ANOVA was used also to identify non-genetic factors that are associated with BMI.

4. **Referee's comment:** "In Results page 9, stating "...carriers of rs11575892 T allele..." is not appropriate since no TT homozygotes were detected and the effect of this genotype on BMI at this stage is purely speculative, as the statement made in Page 10 ("...homozygous individuals who may have pronounced effect..."). In fact, no effect or even a lower BMI could be eventually observed in TT homozygotes. It would be more indicated to use "...rs11575892 CT heterozygotes..." for example"

4. **Answer:** Although we do not think that using phrase “...carriers of rs11575892 T allele...” is not appropriate as all heterozygotes actually are carriers of this allele, we agree that “…rs11575892 CT heterozygotes…” are less likely to be misleading and therefore we replace upper phrase with the suggested one. We fully agree that TT genotype may have no effect on phenotype. However we do not think that our statement suggesting existence of additive genetic model excludes the other possibilities.

5. **Referee's comment:** “In Page 10, splicing could also be positive in contrast to "...truncated or defective AGRP...", by producing an AGRP isoform with a greater orexigenic effect, for example”

4. **Answer:** We agree that without experimental data it is difficult to evaluate possible influence of RNA secondary structure on splicing. We therefore include following sentence in “Results and discussion”: “Alternatively the polymorphism may promote the formation of more active AGRP isoform as a result of modified splicing reaction.”

5. **Referee’s comment:** “In Page 11, the authors need to sequence more then one KB of the promoter to conclude that no variant could be found there. Usually, five kb are sequenced”

5. **Answer:** We agree that to identify possible functional mutation or SNP in LD it is necessary to sequence larger regions not only to 5’ but also 3’ direction. However we do not state in the manuscript that all possible sites were tested for presence of variants but only refer to the region included in the sequencing.

5. **Referee’s comment:** “In the Abstract, please clarify what are the "other” non-genetic factors.”
5. Answer: We include clarification in the abstract: “…and other significant non-genetic factors (presence of diseases), the BMI…”.

6. Referee’s comment: “In Page 10, what the authors mean by “controlled environmental factors”? Did they mean stratification for age, sex and other factors?.”

6. Answer: To avoid misunderstanding we include clarification in the text: “…case-control studies with stratification for environmental factors such as social background, diet and physical activity would be valuable to verify the effects of rs1157892.”

7. Referee’s comment: “In the Conclusion, the authors speculate on the usefulness of their results as a new diagnostic marker. What are the specificity and the sensitivity of their marker to detect a higher BMI? It is probably very low since the MAF of this SNP and it is doubtful that it can be use as diagnostic.”

7. Answer: We agree that by itself the rs11575892 may be unlikely to serve as diagnostic marker due to reasons mentioned by referee. We therefore introduce the following changes into text:

“This SNP could provide valuable information on the regulation of this gene and could be a component of a more complex set of diagnostic markers for susceptibility to increased body weight.”
Referee #3:

We are encouraged that the referee states that: “This is the first time that a SNP in AgRP is associated with increased BMI and the authors have properly investigated the molecular impact of the SNP on RNA structure.” and “The association analysis has also been properly performed using a robust sample size population”

Referee’s s comment: "I am not sure how the manuscript benefits from Table 3. It seems to be out of place and it is cited in the text only once in a short sentence. It could make better sense if the analysis had taken into consideration the presence of the SNP, which is the essential theme of the manuscript."

Answer: We agree that information in Table 3 is less essential, taken into account the main subject of manuscript: the role of the SNP in obesity. We therefore include this table as supplementary information to maintain the possibility for readers to access the actual levels of association of each factor with BMI.
Referee #4:

We thank the referee for the comments that work concerns interacting candidate gene that may be associated with obesity.

Answers to the following comments:

1. **Referee’s comment:** “It is not clear, why at least one additional polymorphic variant (among two SNPs in complete LD rs34123523 or rs5030980) wasn’t genotyped in the studied sample. These variants have been previously associated with obesity, but it is not known, whether the association with obesity is observed within the studied Latvian sample. Such genotyping would be important for overall evaluation of the role of AGRP in the increased BMI”

   **1. Answer:** This particular association analysis was performed in order to evaluate the role of rare SNP’s found in resequencing of a specific cohorts of patients rather than full evaluation of AGRP gene polymorphisms. We feel that analysis of rare SNP’s are often underrepresented compared to common SNP’s, while combination of rare SNP’s may have great impact on genetics of common diseases. Although we agree that replication studies are of great importance, the main goal of this study was to analyse SNP that was not previously studied with respect to obesity. To fully evaluate the impact of particular gene locus, we would rather consider including number of tagSNPs based on LD analysis in association study, but such analysis was left beyond the scope of project due to financial and technical reasons.

2. **Referee’s comment:** “As far as all observed variants within this gene are rare (only 3 heterozygotes were observed in the sequenced sample of 95 subjects), it’s not possible to exclude that in the whole sample rare allele of rs11575892 might be present in the subjects with rare variant of rs34123523 or rs5030980, i.e. some LD might be observed between these three variants”

   **2. Answer:** Although the existence of LD between these variants can not be completely ruled out, it seems unlikely that significant LD would be discovered with more extensive genotyping (also not supported by literature data). Even in case of some LD between rs11575892 and rs5030980, it would not explain the association of the rs11575892 with increased BMI as rs5030980 in associated with lean phenotype in other studies (as referred in the manuscript). It should also be noted that according to our calculation 0.2% (based on our sequencing data) of the study population would be carriers of both polymorphisms in absence of LD. In this case the presence of both variations would eventually have impact on the phenotype. From the other hand, there are many known SNPs that influence BMI and we would rather be interested in such
analysis by including the number of well known obesity SNPs as covariates or perform stratification of study group based on these SNPs. From such viewpoint there are no special reason why nearby SNP would have preference to be included in such analysis (except in the case of specific interaction effects between these SNPs).

3. Referee’s comment: “Authors performed permutation analyses to check whether the association was observed by chance. However, genotyping of the studied polymorphic variant in an additional replication sample would be also desired”

3. Answer: We agree in general that replication of results in additional population would be the best indicator of real association. Unfortunately, at the present time we do not have access to the equally designed study group. It should be noted that according to other papers AGRP may have function in regulation of obesity in age dependent manner and therefore it would be important to include age as one of the major stratification factor for replication sample. We also hope that publishing of these results would encourage others to perform replication studies of this SNP.

3. Referee’s comment: “Power calculation would be helpful to evaluate, whether the sample size of the cohort is adequate to demonstrate such association with sufficient certainty”

3. Answer: We added power calculation to the present manuscript to demonstrate the actual power of the analysis performed. We include following description in Methods: “Power calculations were performed using Quanto v.1.2.3 (). Our sample size provided 80% power (at α=0.05) to detect difference between genotypes in logBMI of >0.053 assuming minor allele frequency of 0.012.” The following text is included in Results and discussion section: “The power analyses show that present sample size had 53% power to detect changes in logBMI of 0.04 that was observed in our study.”

4. Referee’s comment: “A note about how the secondary RNA structure could be affected by other two SNPs would be useful”

4. Answer: Since rs11575892 is the main subject of the study we did not consider to include these variations in the RNA secondary structure analysis. It is also possible that rs5030980 has functional effect due to amino-acid change rather than affecting the RNA structure.

5. Referee’s comment: “Table 2 has a lot of comments in the legend, thus, is quite difficult to read, while the content is quite simple. This table could be divided into 2 parts (having separate 2 rows for each analysis) for ANOVA and GLM analyses. It would be clearer to put two sub-headers for analysis
considering BMI as continuous trait and for categorical analyses where sample was divided by median BMI”

5. **Answer:** We carefully examined suggestions of referee on change of Table 2. We introduce the header row to indicate the type of analysis performed as suggested by referee. Concerning the other changes we feel that they would require the separation of the data into two tables based on type of analysis. Since our goal was to present the data in concise form we leave the structure of the table previous, but are ready to introduce these changes if they are required.

6. **Referee’s comment:** “Table 1 and Table 3 could be merged in order to make clearer the distribution of associated traits and their loading as covariates in the association with BMI”

6. **Answer:** We have removed Table 3 from the main manuscript as suggested by Referee #3. We include this table as additional information to maintain the possibility for readers to access the actual levels of association of each factor with BMI.

7. **Referee’s comment:** “Table 1 header “Values” is not clear because the Sub-header “Percentage of patients (n)” is placed in the other column, it would be clearer to put it in the second column”

7. **Answer:** We have chosen the header “Values” since different type of data is displayed in this column. Besides categorical data there are also quantitative data included that cannot be described by “Percentage of patients”

8. **Referee’s comment:** “MC4 and MC3 receptors need to be referenced with their full name. All gene names should be in Italic characters”

8. **Answer:** “We included the full name references for MC4 and MC3 receptors as well as changed the gene names to Italic characters.”