PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Appetite Regulation Genes are Associated with Body-Mass Index in Black South African Adolescents: A Genetic Association Study
AUTHORS	Zané Lombard, Nigel J Crowther, Lize van der Merwe, Punita Pitamber, Shane A Norris and Michèle Ramsay

VERSION 1 - REVIEW

REVIEWER	Dr B.J. Hennig MRC Senior Investigator Scientist MRC ING at London School of Hygiene & Tropical Medicine UK Competing interests: none to declare
REVIEW RETURNED	12/02/2012

THE STUDY	Population substructure: The analysis of population substructure was based on the inclusion of HapMap data from YRI, CEU and CHB. Given that the study population was black South Africans, why were other African HapMap populations not included for comparative purposes, e.g. the Luhya (LWK) and/or Maasai (MKK)? Including other African populations might help to identify more subtle population substructure in the sample cohort. If LWK and MKK data was not available at the time of the selection of ancestry informative markers or the analysis (which could be noted in the manuscript) please ignore this comment . Gene/SNP selection: The authors clearly followed a well thought-through strategy in selecting candidate genes. However, the selection of candidate SNPs is not described in quite enough detail to understand the selection process. TagSNPs only appear to have been selected for LEP, NPY2R and POMC, but not the other candidate genes, why? Could LWK and MKK data have been employed in the selection of tagSNPs (if applicable the design stage)? How were the 18 ancestry informative markers selected? Suggested further reference: * Polymorphisms in the NPY2R gene show significant associations with BMI that are additive to FTO, MC4R, and NPFFR2 gene effects. Hunt SC, Hasstedt SJ, Xin Y, Dalley BK, Milash BA, Yakobson E, Gress RE, Davidson LE, Adams TD. Obesity (Silver Spring). 2011 Nov;19(11):2241-7.
RESULTS & CONCLUSIONS	Results: There is no description of the quality control measures employed to assess the genotype data, other than HWE. A statement on the average genotyping success rate across all markers and individuals would be helpful (although some of this information is shown in table 1).

	Although the genotype/allele frequencies of the 18 ancestry informative markers are not relevant to the analysis of BMI markers, it would for completeness, be useful to present this data as reference for future studies (especially those in comparable study populations), e.g. as supplementary information or in table 1. It would further be helpful to give, as a (supplementary) summary table, details of completenes/missingness of data on covariates, in particular those included in the adjusted analyses (age, sex and prepubertal stage). Discussion: The power calculation was it seems based on N=990, but the joint model only comprised 908 individuals; loss of power should therefore be discussed. It would also be helpful to state the N of individuals by SNP in tables 2 and 3. The association with MR4R rs17782313 is borderline (P=0.045), this should be commented on in relation to the other associations identified. The authors bring up linkage disequilibrium and haplotype blocks in conjunction with several of their associations. However, it is not clear whether an analysis of LD/Haplotype structure (e.g. using Haploview) in their own population was done (p13 first paragraph suggests they might have). If so then details of the method employed and results obtained should be described in the paper. It certainly would be good if the findings of this study could be discussed in the context of LD/haplotype structure in their own population. LD heat maps could be presented as supplementary figures. It is interesting to see that the cumulative effect of four risk alleles in FTO, LEP and MC4R amounted to 2.1% of the of the variation in observed (log)BMI in Bt20, compared to the study by Speliotes et al reporting an effect of only 1.5% based on 32 SNPs. Any idea why this may be? The authors suggest that it is "possible that the effects on weight of some polymorphisms may have been masked by puberty-associated changes in body fat mass." Some of the candidate genes investigated have also been shown to associate with age at menarche - has this been looked at in this study?
GENERAL COMMENTS	Overall this is a nicely-defined and well-described study, with clear objectives, and a sensible structure of the findings. In addition to the above made comments a few further suggestions for minor changes: P3 Introduction: Change "developing countries" to "low- and middle-income countries". P5 Methods: The study was carried out in a randomly selected sub-set of samples of the Bt20 cohort, but it would be good to remind the reader what proportion of the total Bt20 cohort the 990 study subjects represent. P5 Methods: Change "...(BMI increase..." to "...(changes in BMI..."

	given the results. P5 Methods: In the definition of BMI delete "(measured in kg)" and "(measured in meters)" as weight and height are defined in the previous sentence. P6 Methods: It would be helpful if the same subheadings/descriptors used in the results section were also used in the description of the statistical analysis. The sentence starting "As BMI correlated significantly with gender..." would fit better in the results section. "Each genotype was modeled as the number of minor alleles" means I believe that an additive model was employed, this could be directly stated in the methods. P11: It might be worth pointing out that FTO rs17817449 was directly assayed in the Gambian study in the context of no association in this population. P25: Does figure 1 have a title? Throughout: Change "gender" with "sex" Throughout: There is inconsisteny in the number of decimals presented (e.g. % of increase/decrease of BMI)

REVIEWER	Dr. Vimal Karani Research Associate Institute of Child Health University College London, London, UK
REVIEW RETURNED	17/02/2012

THE STUDY	-The SNP selection in the METHODS section is not clear. The authors have not reported whether the SNPs chosen from the literature were force included in the tagSNP selection. Also, the authors did not include the details of which hapmap data was used. -Correction for multiple testing should be carried out as many tests have been performed and the p values were only borderline. -English language should be improved a lot.
RESULTS & CONCLUSIONS	-The authors have not provided the HWE p value for the SNPs studied. -As the p values were marginally significant, it is very difficult to interpret the findings as they were not corrected for multiple testing. -References are not mentioned in the discussion. Further details can be seen in the attached file. -Some of the interpretations in the discussion are vague and not clear as suitable references have not been provided.
REPORTING & ETHICS	Check list has not been provided
GENERAL COMMENTS	**Comments to the author** In this paper, Lombard et al have studied the association of 44 SNPs from 6 obesity genes with BMI in a South African population (n= 990). They have reported association of four SNPs from *LEP*, *FTO* and *MC4R* genes with BMI after the adjustment for the potential confounders. Please see below for some of my comments regarding the study. 1. In the abstract, the background section is not clear. The authors should re-write it with proper usage of the English language. In the introduction, 2nd paragraph, the authors state that only one GWAS

	study has been performed in African population; but in the background section, the authors state that few studies have been performed. Are the authors trying to indicate the GWAS replication studies? 2. In the 'objective' section of the Abstract, the sentence should be re-written as '*To assess the association of genetic variants with BMI in Black South Africans, we focussed on SNPs*……' 3. In the introduction, last paragraph, the details of the Bt20 cohort (lines 48-58) should go into the Methods section. 4. Were any of the study subjects related? How was this determined? If they were related how was this taken account into analyses? 5. In the Methods, under 'Participants' section, 1st paragraph, the authors state that the study has 80% power to detect differences in BMI of beta ≥0.67 at alpha= 0.05. However, this is likely to change depending on the minor allele frequency (MAF) of the SNPs under study (more power is needed to detect a beta of 0.67 for a SNP with a MAF of 0.04 as compared to a SNP with MAF of 0.47). In this study, the MAFs range from 0.04 to 0.47. Hence, power calculation should be carried out for the range of MAFs. 6. In the Methods, 'SNP Selection' section, the authors should mention which HapMap Release data was used? Also, it is not clear whether the authors have force-included the SNPs identified from the literature in the tagSNP selection procedure. In the Abstract, 'Methods' section, the authors state '*44 SNPs previously associated with BMI (and including tagSNPs), as well as*…'. What do the authors mean by '*(and including tagSNPs)*'? The authors should clarify this issue? 7. HWE p value should be reported for all the SNPs. This can be included in the table 1 along with the allele frequencies. 8. In the Methods, 'Statistical analysis' section, the authors have tried to justify for not correcting the p values for multiple testing. May be the justification might hold true, when examining few candidate SNPs; however, in a study with 44 candidate SNPs, for every hundred test one does, five will be significant by chance (most of the significant p values observed in this study were only nominally significant). This could lead to Type 2 error. Also, the reference number 22 (second last line of the 'statistical analysis' section) does not seem to be the right reference. 9. In the Results section, under 'Gender specific effects' section, the authors state '*Prior research suggests that some loci show gender specific effects*'. This line is quite vague- authors need to giving a few valid convincing references. 10. How much of the variation in the BMI was explained by *FTO* alone, as *FTO* is the only gene that has been shown to be consistently associated with BMI in all the populations, while *LEP* and *MC4R* have been quite inconsistent in the candidate gene studies? This should be mentioned as part of the discussion section. 11. In the discussion section, 2nd paragraph is not clear enough to understand what the authors are trying to say. The authors start the paragraph stating the findings from their study and then discussing the findings from European populations and then again going back to African populations- this is not clear. The last line of the 2nd paragraph should be deleted as the interpretation wouldn't make any sense to the reader. The language should be improved. 12. In the discussion section, 3rd paragraph, the authors should give a reference in the line 48 (the line ending with '… *gene transcription*'). Also, the last line of 3rd paragraph is not clear. The authors compare their study with that of the one published in a

	German population and make a conclusion based on the finding, which sounds very irrelevant as the LD structure is quite different in European and Black populations. 13. In the discussion section, 10th paragraph, references should be given wherever appropriate. 14. Discussion section can be reduced as some of the lines are redundant. The p values reported are only marginally significant (without the correction for multiple testing); hence, discussing these nominally significant associations with biological justifications should be underplayed, especially paragraphs 4, 5 and 8. 15. The title for table 3 should be changed. Please make the title short and discuss the findings at the end of the table. 16. The last paragraph of the discussion section should be mentioned somewhere in the beginning or perhaps in the introduction. The last few lines of the discussion should be conclusive based on the study findings.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1: Dr B.J. Hennig
(MRC Senior Investigator Scientist; MRC ING at London School of Hygiene & Tropical Medicine UK)

Competing interests: none to declare

1. Population substructure: The analysis of population substructure was based on the inclusion of HapMap data from YRI, CEU and CHB. Given that the study population was black South Africans, why were other African HapMap populations not included for comparative purposes, e.g. the Luhya (LWK) and/or Maasai (MKK)? Including other African populations might help to identify more subtle population substructure in the sample cohort. If LWK and MKK data was not available at the time of the selection of ancestry informative markers or the analysis (which could be noted in the manuscript) please ignore this comment.
AIM SNPs were selected based on previously published data {Lao, 2006 #90} and local unpublished data {Schlebusch, 2011 #174}. Reference to this has been added on pg.6.

At the time of analysis, there was genotype data available for only nine of the 18 AIM SNPs in LWK and MKK. This lack of data implied that inclusion thereof would not have improved resolution. As the Bantu-expansion occurred only ~5000 years ago one wouldn't expect to see subtle African-specific substructure with relatively few markers, but would only be able to see substructure from non-African contributions.

2. Gene/SNP selection: The authors clearly followed a well thought-through strategy in selecting candidate genes. However, the selection of candidate SNPs is not described in quite enough detail to understand the selection process. TagSNPs only appear to have been selected for LEP, NPY2R and POMC, but not the other candidate genes, why? Could LWK and MKK data have been employed in the selection of tagSNPs (if applicable the design stage)? How were the 18 ancestry informative markers selected?
At the time of study design data from LWK and MKK were not available. tagSNP selection was limited by assay compatibility and plexity constraints. As a primary objective, previously associated SNPs were selected first for inclusion, followed by tagSNPs for only three of the loci, as could be accommodated in the assay we were using. See comment 1 above regarding AIMs selection.

3. Suggested further reference:
Polymorphisms in the NPY2R gene show significant associations with BMI that are additive to FTO, MC4R, and NPFFR2 gene effects. Hunt SC, Hasstedt SJ, Xin Y, Dalley BK, Milash BA, Yakobson E,

Gress RE, Davidson LE, Adams TD. Obesity (Silver Spring). 2011 Nov;19(11):2241-7.
This paper does not directly address a sex specific effect and was therefore not included in the paper.

Results:
4. There is no description of the quality control measures employed to assess the genotype data, other than HWE. A statement on the average genotyping success rate across all markers and individuals would be helpful (although some of this information is shown in table 1).
A section describing quality control measures were added (pg.6), and further information to table 2.

5. Although the genotype/allele frequencies of the 18 ancestry informative markers are not relevant to the analysis of BMI markers, it would for completeness, be useful to present this data as reference for future studies (especially those in comparable study populations), e.g. as supplementary information or in table 1.
To address this comment, the following data has been added - allele- and genotype frequencies of AIMs in the South African cohort (Supplementary tables 2) and allele frequencies of AIMs in the South African cohort (SAB) and three HapMap populations (Supplementary table 3).

6. It would further be helpful to give, as a (supplementary) summary table, details of completenes/missingness of data on covariates, in particular those included in the adjusted analyses (age, sex and prepubertal stage).
The number of individuals genotyped, as well as the number of individuals included in each model was added to table 2.

Discussion:
7. The power calculation was it seems based on N=990, but the joint model only comprised 908 individuals; loss of power should therefore be discussed.
There are 990 children in the dataset, but we miss many values of covariates. The actual tests were done on at most 865 children. Of course power calculations are tentative and based on many assumptions. We certainly had less power than we hoped for (assuming complete data), but fortunately we did have sufficient power to detect the associations reported in this manuscript. To illustrate the power we had with this study, and example of the calculation with the real data from our most significant SNP (LEP rs rs10954174)was added (pg.5)

8. It would also be helpful to state the N of individuals by SNP in tables 2 and 3.
The number of individuals genotyped, as well as the number of individuals included in each model was added to table 2.

9. The association with MR4R rs17782313 is borderline (P=0.045), this should be commented on in relation to the other associations identified.
The Discussion section pertaining to this was adjusted (pg.13)

10. The authors bring up linkage disequilibrium and haplotype blocks in conjunction with several of their associations. However, it is not clear whether an analysis of LD/Haplotype structure (e.g. using Haploview) in their own population was done (p13 first paragraph suggests they might have). If so then details of the method employed and results obtained should be described in the paper. It certainly would be good if the findings of this study could be discussed in the context of LD/haplotype structure in their own population. LD heat maps could be presented as supplementary figures.
A section describing the LD plot (now included as supplementary figure 2) was added to the Methods section (p.7) and Results section (p.9).

11. It is interesting to see that the cumulative effect of four risk alleles in FTO, LEP and MC4R

amounted to 2.1% of the of the variation in observed (log)BMI in Bt20, compared to the study by Speliotes et al reporting an effect of only 1.5% based on 32 SNPs. Any idea why this may be?
Our study is not directly comparable to the Speliotes study, as that study was performed in adults and was not adjusted for the confounders that we had to take into consideration in or adolescent cohort. Variant effects could have varying degrees of effect, depending on the point in the life-course that the analysis was done (Mei et al., 2012), which could also explain the differences in effect observed.

12. The authors suggest that it is "possible that the effects on weight of some polymorphisms may have been masked by puberty-associated changes in body fat mass." Some of the candidate genes investigated have also been shown to associate with age at menarche - has this been looked at in this study?
No, age of menarche has not been investigated in this study. This may be a focus of a follow-up study.

Overall this is a nicely-defined and well-described study, with clear objectives, and a sensible structure of the findings. In addition to the above made comments a few further suggestions for minor changes:

13. P3 Introduction: Change "developing countries" to "low- and middle-income countries".
Suggested change implemented. (pg.3)

14. P5 Methods: The study was carried out in a randomly selected sub-set of samples of the Bt20 cohort, but it would be good to remind the reader what proportion of the total Bt20 cohort the 990 study subjects represent.
Details on what proportion of the total cohort is represented in this study are added on pg. 5

15. P5 Methods: Change "...(BMI increase..." to "...(changes in BMI..." given the results.
Suggested change implemented. (pg.7)

16. P5 Methods: In the definition of BMI delete "(measured in kg)" and "(measured in meters)" as weight and height are defined in the previous sentence.
Suggested change implemented. (pg.5)

17. P6 Methods: It would be helpful if the same subheadings/descriptors used in the results section were also used in the description of the statistical analysis.
Suggested change implemented throughout the Methods and Results section.

18. The sentence starting "As BMI correlated significantly with gender..." would fit better in the results section. "Each genotype was modeled as the number of minor alleles" means I believe that an additive model was employed, this could be directly stated in the methods.
The sentence “As BMI correlated significantly with gender….” has been removed, and the section rewritten to aid clarity (pg.8).

19. P11: It might be worth pointing out that FTO rs17817449 was directly assayed in the Gambian study in the context of no association in this population.
Suggested change implemented. (pg.13)

20. P25: Does figure 1 have a title?
Yes, on page 16, under the heading Legends to Figures.

21. Throughout: Change "gender" with "sex"

Suggested change implemented throughout.

22. Throughout: There is inconsisteny in the number of decimals presented (e.g. % of increase/decrease of BMI).
Throughout the paper, the following decimals are presented: two decimals for frequencies, three decimals for P-values and one decimal for percentage change in BMI as well as summary statistics of BMI.

Reviewer 2: Dr. Vimal Karani
(Research Associate, Institute of Child Health, University College London, London, UK)

1. The SNP selection in the METHODS section is not clear.
This section in the Methods has been rewritten to aid clarity (pg. 6)

AIM SNPs were selected based on previously published data {Lao, 2006 #90} and local unpublished data {Schlebusch, 2011 #174}. Reference to this has been added on pg. 6. tagSNP selection was limited by assay compatibility and plexity constraints. As a primary objective, previously associated SNPs were selected first for inclusion, followed by tagSNPs for only three of the loci, as could be accommodated in the assay we were using.

2. The authors have not reported whether the SNPs chosen from the literature were force included in the tagSNP selection.
They were not, text added to reflect this on pg. 6.

3. Also, the authors did not include the details of which hapmap data was used.
Text added to reflect this on pg. 6.

4. Correction for multiple testing should be carried out as many tests have been performed and the p values were only borderline.
This section in the Methods was adjusted as follow:
"…Correcting for multiple testing is a contentious issue, and some approaches (such as Bonferroni correction) is considered over-conservative and one risks the rejection of true findings [25 26]. Given the strong prior information about the role of the variation tested here in obesity, we considered this a replication study, and therefore P-values below 0.05 were considered significant. For tagSNPs, tests of associations could be considered discovery rather than replication, but since these markers are correlated due to linkage disequilibrium, the Bonferroni assumption of independence is not upheld..."

This is a recognised approach to multiple testing used in previously published studies [8].

To illustrate this point further, if we were to implement Bonferroni correction for the tagSNPs, a corrected P-value threshold of 0.002 (0.05/26 SNPs) would be considered significant, which would mean only LEP rs6966536 would be excluded. The new addition of supplementary figure 2 (as requested by Reviewer 1) shows that the LEP markers are not all independant.

5. English language should be improved a lot.
Several sections of the paper have been rewritten to aid clarity.

6. The authors have not provided the HWE p value for the SNPs studied.
Table 1 was updated to include HWE P-values.

7. As the p values were marginally significant, it is very difficult to interpret the findings as they were not corrected for multiple testing.
See remarks on comment 4.

8. References are not mentioned in the discussion. Further details can be seen below. Some of the interpretations in the discussion are vague and not clear as suitable references have not been provided.
References were omitted due to the Journal's restriction on the number of references to be included. Additional references are now included.

9. In the abstract, the background section is not clear. The authors should re-write it with proper usage of the English language.
The abstract has been rewritten to aid clarity.

10. In the introduction, 2nd paragraph, the authors state that only one GWAS study has been performed in African population; but in the background section, the authors state that few studies have been performed. Are the authors trying to indicate the GWAS replication studies?
The updated abstract and introduction now makes it clear that only one GWAS in Africans focused on BMI has been published, with a few REPLICATION studies in Africans (pg.2).

11. In the 'objective' section of the Abstract, the sentence should be re-written as 'To assess the association of genetic variants with BMI in Black South Africans, we focussed on SNPs……'
Abstract was rewritten to aid clarity.

12. In the introduction, last paragraph, the details of the Bt20 cohort (lines 48-58) should go into the Methods section.
Suggested change implemented. (pg.5)

13. Were any of the study subjects related? How was this determined? If they were related how was this taken account into analyses?
Individuals were not related, which was determined based on self-reporting. Relation was not tested empirically.

14. In the Methods, under 'Participants' section, 1st paragraph, the authors state that the study has 80% power to detect differences in BMI of beta ≥0.67 at alpha= 0.05. However, this is likely to change depending on the minor allele frequency (MAF) of the SNPs under study (more power is needed to detect a beta of 0.67 for a SNP with a MAF of 0.04 as compared to a SNP with MAF of 0.47). In this study, the MAFs range from 0.04 to 0.47. Hence, power calculation should be carried out for the range of MAFs.
The reviewer's statement is absolutely true. Of course power calculations are tentative and based on many assumptions. We certainly had less power than we hoped for (assuming complete data), but fortunately we did have sufficient power to detect the associations reported in this manuscript. To illustrate the power we had with this study, and example of the calculation with the real data from our most significant SNP (LEP rs10954174) was added (pg.5)

15. In the Methods, 'SNP Selection' section, the authors should mention which HapMap Release data was used? Also, it is not clear whether the authors have force-included the SNPs identified from the literature in the tagSNP selection procedure. In the Abstract, 'Methods' section, the authors state '44 SNPs previously associated with BMI (and including tagSNPs), as well as…'. What do the authors mean by '(and including tagSNPs)'? The authors should clarify this issue?
See remarks on comments 1-3.

16. HWE p value should be reported for all the SNPs. This can be included in the table 1 along with the allele frequencies.
See remarks on comments 6.

17. In the Methods, 'Statistical analysis' section, the authors have tried to justify for not correcting the p values for multiple testing. May be the justification might hold true, when examining few candidate SNPs; however, in a study with 44 candidate SNPs, for every hundred test one does, five will be significant by chance (most of the significant p values observed in this study were only nominally significant). This could lead to Type 2 error. Also, the reference number 22 (second last line of the 'statistical analysis' section) does not seem to be the right reference.
See remarks on comments 6. Reference 22 is pertinent to this section, but was placed incorrectly. It has been replaced by an updated reference.

18. In the Results section, under 'Gender specific effects' section, the authors state 'Prior research suggests that some loci show gender specific effects'. This line is quite vague- authors need to giving a few valid convincing references.
This sentence has been removed.

19. How much of the variation in the BMI was explained by FTO alone, as FTO is the only gene that has been shown to be consistently associated with BMI in all the populations, while LEP and MC4R have been quite inconsistent in the candidate gene studies? This should be mentioned as part of the discussion section.
The 7 FTO SNPs in this study explain 0.6% of variation in log(BMI) after adjusting for age gender and sex-specific pubertal stage; and 1.4% unadjusted. This information has been added on pg.11.

20. In the discussion section, 2nd paragraph is not clear enough to understand what the authors are trying to say. The authors start the paragraph stating the findings from their study and then discussing the findings from European populations and then again going back to African populations- this is not clear. The last line of the 2nd paragraph should be deleted as the interpretation wouldn't make any sense to the reader. The language should be improved
The Discussion has been reorganized and rewritten to aid clarity.

21. In the discussion section, 3rd paragraph, the authors should give a reference in the line 48 (the line ending with '… gene transcription').
An appropriate reference has been added (pg.13)

22. Also, the last line of 3rd paragraph is not clear. The authors compare their study with that of the one published in a German population and make a conclusion based on the finding, which sounds very irrelevant as the LD structure is quite different in European and Black populations.
This conclusion has been removed.

23. In the discussion section, 10th paragraph, references should be given wherever appropriate.
Appropriate references have been added.

24. Discussion section can be reduced as some of the lines are redundant. The p values reported are only marginally significant (without the correction for multiple testing); hence, discussing these nominally significant associations with biological justifications should be underplayed, especially paragraphs 4, 5 and 8.
The Discussion has been reorganized and rewritten to aid clarity.

25. The title for table 3 should be changed. Please make the title short and discuss the findings at the end of the table.

Title amended (pg.22).

26. The last paragraph of the discussion section should be mentioned somewhere in the beginning or perhaps in the introduction. The last few lines of the discussion should be conclusive based on the study findings.
Paragraph moved to the introduction (pg. 3-4).

VERSION 2 – REVIEW

REVIEWER	Dr B.J. Hennig MRC Senior Investigator Scientist MRC ING at London School of Hygiene & Tropical Medicine UK Competing interests: none to declare
REVIEW RETURNED	30/03/2012

RESULTS & CONCLUSIONS	1. The authors added a LD heat map of the LEP gene, generated using default parameters as referenced. This is great, however, there doesn't seem to be a legend to the supplementary figure 2 and as such it is not clear whether the heatmap represents LD as D' or r-square (the latter would be preferable to account for SNPs with a low minor allele frequency). The authors say (p10) that "Evidence of LD structure in the 3'UTR of LEP can be observed, which is different to the haplotype structure observed in the HapMap European and Nigerian populations". This sentence needs amendment in my view, as it seem that the heatmap indicates three LD blocks covering the whole gene, with the largest block spanning the 3' end of the gene (and please delete the word "structure"). Finally, the authors suggest that LD pattern is different to that seen in Caucasisans and Yoruba, yet heat maps for these populations are not represented for compartative purposes - they should be included.
GENERAL COMMENTS	Thank you for addressing the reviewers comments so carefully and providing additional material as part of this manuscript. There just are a couple of further points to raise: 1. Please see above (re LD heat plots) 2. There are a few small errors: * p8, the last sentence of the 1st paragraph should read "The chromosomal location of SNPs was obtained from NCBI (genome build 37.1). * p9, 2nd paragraph should read: "...and some approaches (such as Bonferroni correction) are considered..." *p10, 2nd paragraph should read: "One AIM failed quality control measures and was subsequently excluded...."

VERSION 2 – AUTHOR RESPONSE

Suggestion 1:

A legend has been added to the file Supplementary Material.
Pg. 10 was amended as follow:
LD plot
To illustrate the unique patterns of LD in the South African population, an LD plot of the gene most significantly associated with BMI (LEP) was constructed (Supplementary Figure 2) Evidence of three

LD blocks covering the gene is observed, with the 3' end of the gene in particular exhibiting high LD.

Suggestion 2:

All small errors were corrected