PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The European lactase persistence genotype determines lactase persistence in the Hispanic and Amerindian Chilean population: relevance for prevalence of adult type hypolactasia and lactose intolerance in Latin American populations.
AUTHORS	Morales, Eugenia; Azocar, Lorena; Maul, Ximena; Perez, Claudio; Klassen, Julieta; Nervi, Flavio; Chianale, Jose; Miquel, Juan

VERSION 1 - REVIEW

REVIEWER	***Mauro Congia*** Department of Biomedical Science and Biotechnology Cagliari University Ospedale Microcitemico Via Jenner 09121, Cagliari, Sardinia, Italy. I declare no competing interests.
REVIEW RETURNED	16-Apr-2011

GENERAL COMMENTS	Morales et al. “The European lactase persistence genotype determines lactase persistence in the Hispanic and Amerindian Chilean population: relevance for prevalence of adult type hypolactasia and lactose intolerance in Latin American populations.” determined, in a population of 51 Chilean people previously selected for gastroenterological symptoms suggestive of lactose maldigestion, the frequency of LNP and LP using the HBT with 25 g of lactose. Then, the authors established the presence in these 51 Chileans of the same LCT C>T-13910 variants associated with LNP and LP in Europe and their strong correlation with the HBT and symptoms reported by patients both during the HBT and by a self administered questionnaire. Next, they examined, among different populations that compose Chileans, a control group of 216 Hispanics and 41 Amerindians to perform a genotyping for the LNP and LP LCT C>T-13910 variant. The paper reports data already known about the association of LNP and LP LCT C>T-13910 variants and LBHT and symptoms of lactose maldigestion. However, this is still an interesting paper because it reports for the first time genetic data of these variants in Latin America and in particular in the Amerindian population. However, there are a series of critiques that need to be answered before the paper can be considered for publication. 1) Page 4 row 34. Mention also the 41 Amerindians

	2) Page 6 row 15: avoid the term “lactose intolerance” because it should be reserved only to describe the appearance of gastrointestinal symptoms due to lactose maldigestion. Indeed, many individuals with lactose maldigestion demonstrated with the BHT do not show gastrointestinal symptoms. 3) Page 12 rows 9-14. “Fifty-one patients with clinical suspicion of LNP (44 women and 7 men, age 14 - 79 y) were included in the study. Twenty-nine patients (56.8%) had a positive HBT and 22 were negative (43.2%) (Table 1).” In spite of the fact that the 51 Chilean patients were selected with the clinical suspicion of LNP they show a lower frequency of lactose malabsorption (56.8%) if compared to the frequency of reference 27 representative of the Chilean general population and very similar to the frequency of LCT-13910 CC genotype observed among the 216 Hispanics (56.9%). Have the authors an explanation for this difference? Was the proportion of Hispanics (or other Europeans) higher among the 51 patients than in the general Chilean population? 4) Page 12 row 28 and row 33. In the Results section is reported a p value of <0.0001 whilst in table 1 is reported a p value of <0.001. Please uniform the table with the text. 5) Page 15 row 8. Correct persistance with persistence 6) Page 15 row 25 “But herein called LCTC>T-13910”. The LCTC>T-13910 variant has already been used in different sections. Therefore remove the phrase “known in the literature by many names but herein called LCTC>T-13910,” or move it to the first citation of the LCTC>T-13910 in the text. 7) Page 15 rows 27-34 “has spread progressively throughout the world through population migration and mixing, increasing the state of LP in different populations given its autosomal dominant property (2)”. It seems a bit imprudent of a statement. Indeed, it has been demonstrated that the diffusion of several variants close to the LCTC>T-13910 (-13907, -13915, -13913, -13914, and -14010 variants) is driven by a strong still-ongoing adaptation of LP evolution in response to adult milk consumption in different human populations and not simply because that variant is dominant (Evidence of Still-Ongoing Convergence Evolution of the Lactase Persistence T-13910 Alleles in Humans; Enattah et al; The American Journal of Human Genetics Volume 81 September 2007). Please modify the phrase accordingly. 8) Page 15 rows 44-46. “and spread rapidly in the Hispanic populations of contemporary America, as is shown in this study as well as others”. The frequency of LP among Spanish is around 64% (reference 39) a frequency higher than that reported in the present paper (43.1%). Therefore the LCTC>T-13910 probably did not spread in the Hispanic population but more likely decreased a little from that of the original population. This data is in agreement with the next paragraph and reference 29 suggestive of a small genetic admixture between Hispanic Chilean

	population and Amerindians. 9) Page 16 rows 20-37. The whole paragraph appears obscure. Indeed, at page 10 rows 3-8 the authors state: “The sample size was selected by estimating an expected prevalence of the LCT-13910CC genotype to be at least 50% in the Hispanic population”. Therefore they were expecting a frequency of the LP determined by the LCT-13910CT or LCT-13910TT genotypes among Hispanics of 50%. Now the actual frequency of LP obtained in the group of 216 unrelated Hispanic was 43.1%, a frequency lower and not higher than what was expected (see also critiques 3 and 8). In addition at page 18 rows 6-8 the same authors state: “existence of a high prevalence of LNP state within this population” that is again in contradiction with rows 20-36 of page 16. Please revise, modify or delete the whole paragraph from row 20 to 36. 10) Please double check the calculations obtained for reference of the HBT: sensitivity, specificity, PPV, NPV and LHR of the LCTC>T-13910 variant reported in table 1 and in the text. I have obtained slightly different numbers for sensitivity and NPV. 11) Table 2. Please correct “predictive value positive and predictive value negative” with “positive predictive value” and “negative predictive value”, respectively. 12) Table 3. Please control all the p values presented in the table. I did a chi square test and found a significant difference for the category: number reporting diarrhea.

REVIEWER	***Nahum Méndez-Sánchez, MD, PhD*** Professor of Medicine and Chief of the Department of Biomedical Research Medica Sur Clinic & Foundation Mexico City, Mexico I have not any conflic of interest
REVIEW RETURNED	20-Apr-2011

GENERAL COMMENTS	Manuscript Title: The European lactase persistence genotype determines lactase persistence in the Hispanic and Amerindian Chilean population: relevance for prevalence of adult type hypolactasia and lactose intolerance in Latin American populations. Manuscript ID: bmjopen-2011-000125 The aim of this study was to investigate if the SNP influences the LP/LNP state in the Chilean population, and the prevalence of LCT C>T-13910 genotypes in a representative sample of 216 Hispanics and 43 Amerindians with correlation to digestive symptoms. The investigators found 29 patients were HBT positive with clinical suspicion of LNP. The CC genotype (LNP) was present in 89.7% of the patients with positive HBT and in only 4.7% of those with negative HBT. The prevalence of the CC genotype was 56.9% in the Hispanic population and 88.3% in Amerindians and was associated with a higher self reported clinical intolerance to dairies ingestion. I have few comments

	1. The title is too long. Also since the authors did not study other groups of Latin American countries I suggest to delete Latin American populations 2. A list of abbreviations is mandatory 3. It is important to include in the introduction or subjects-methods sections the information on the ethnical backgrounds of Chilean population. Because it well known that it is different compared to other Latin American countries.

REVIEWER	***Marcin Krawczyk, MD*** Department of Medicine II Saarland University Hospital Saarland University Germany
REVIEW RETURNED	23-Apr-2011

GENERAL COMMENTS	The genetic variant (–13,910 C>T LCT) has been previously reported to underlie the ability to digest milk by adult individuals. The current paper by Miquel et al. is first to investigate the prevalence of this variant in a South American population. The study is well designed, the included populations are well described, and the results are conclusive. Minor criticisms include: 1. Calculating positive and negative predictive values, specificity and sensitivity of genotyping against HBT should be avoided. Conversely, we propose to calculate the positive and negative predictive values, specificity, and sensitivity of HBT against genotyping. 2. In the Chilean cohort, three individuals suffer from secondary hypolactasia (i.e. are HBT positive but carry the risk allele). The authors mention that these patients suffer from lactose malabsorption. The discussion of this interesting result needs to be improved, i.e. were there cases of secondary hypolactasia in other genetic studies. 3. Departure from HWE in the cohort of 216 Hispanics remains to be explained. Were any of these individuals related (family members)? Are the authors aware of other studies concerning this polymorphism that report deviation from HWE in selected populations? 4. In the Introduction the authors should clearly explain the difference in the ethnic background between Amerindians and Hispanics. 5. In the Discussion (p. 15, line 18) the authors should also briefly mention that other lCT mutations have recently been identified in small African populations. 6. p. 4, line 50: should read 'non randomized' 7. p. 16, line 6: should read 'degree of' 8. p. 16, 2nd paragraph may be omitted. 9. p. 16, line 34: 'recently' should be deleted (refers to 2005). 10. Figure 1 is not necessary (standard technique). 11. Please include the ‘LCT’ in the abbreviations list under each table.

VERSION 1 – AUTHOR RESPONSE

Dear Mr. Sands,
Thank you for your critical evaluation of our manuscript and for the exhaustive evaluation and

contributions obtained from the reviewers, giving us the opportunity to improve the quality of our manuscript. We have made all the suggestions you requested and believe the manuscript is now significantly improved. Please find our responses to your specific questions below. The new version of the paper contains all changes in red highlight text to facilitate the revision process.

Response to authorship observation from Mr Richard Sands, Managing Editor BMJ Open:
I have revised the contribution of coauthors according to the suggestion and, decided to move two of the collaborators to the acknowledgments section. The rest of the author contributor ship makes them deserve authorship.

Response to Reviewer Mr Mauro Congia (Department of Biomedical Science and Biotechnology, Cagliari University, Italy):
1) Page 4 line 34. Mention also the 41 Amerindians.
Response: we added the number (n=43, not 41)

2) Page 6 line 15: avoid the term “lactose intolerance” because it should be reserved only to describe the appearance of gastrointestinal symptoms due to lactose maldigestion. Indeed, many individuals with lactose maldigestion demonstrated with the HBT do not show gastrointestinal symptoms.

Response: We thank the reviewer for pointing out this errors and the new manuscript includes the corrections you suggested.

3) Page 12 lines 9-14.
“Fifty-one patients with clinical suspicion of LNP (44 women and 7 men, age 14 - 79 y) were included in the study. Twenty-nine patients (56.8%) had a positive HBT and 22 were negative (43.2%) (Table 1).” In spite of the fact that the 51 Chilean patients were selected with the clinical suspicion of LNP they show a lower frequency of lactose malabsorption (56.8%) if compared to the frequency of reference 27 representative of the Chilean general population and very similar to the frequency of LCT-13910 CC genotype observed among the 216 Hispanics (56.9%). Have the authors an explanation for this difference? Was the proportion of Hispanics (or other Europeans) higher among the 51 patients than in the general Chilean population?

Response: We thank the reviewer for pointing this. The aim of this part of the clinical study (51 patients) was to view the correlation between genotyping test and HBT in Chilean subjects. So we did not check the ethnicity of individuals. This could explain your observation, and you are right in terms that it is possible that this small group differ in terms of the degree of European/Amerindian admixture compare to the general population.

4) Page 12 line 28 and line 33. In the Results section is reported a p value of <0.0001 whilst in table 1 is reported a p value of <0.001. Please uniform the table with the text.
5) Page 15 line 8. Correct persistance with persistence
6) Page 15 line 25. “But herein called LCTC>T-13910”. The LCTC>T-13910 variant has already been used in different sections. Therefore remove the phrase “known in the literature by many names but herein called LCTC>T-13910,” or move it to the first citation of the LCTC>T-13910 in the text.

Response: We thank the reviewer for pointing out this errors and the new manuscript includes all the corrections you suggested.

7) Page 15 lines 27-34. “has spread progressively throughout the world through population migration and mixing, increasing the state of LP in different populations given its autosomal dominant property (2)”. It seems a bit imprudent of a statement. Indeed, it has been demonstrated that the diffusion of

several variants close to the LCTC>T-13910 (-13907, -13915, -13913, -13914, and -14010 variants) is driven by a strong still-ongoing adaptation of LP evolution in response to adult milk consumption in different human populations and not simply because that variant is dominant (Evidence of Still-Ongoing Convergence Evolution of the Lactase Persistence T-13910 Alleles in Humans; Enattah et al; The American Journal of Human Genetics Volume 81 September 2007). Please modify the phrase accordingly.

Response: Thank you for the observation. Indeed, the diffusion of this variant is not only because it's dominant property, so we modified the phrase according your suggestion.

8) Page 15 lines 44-46. "and spread rapidly in the Hispanic populations of contemporary America, as is shown in this study as well as others". The frequency of LP among Spanish is around 64% (reference 39) a frequency higher than that reported in the present paper (43.1%). Therefore the LCTC>T-13910 probably did not spread in the Hispanic population but more likely decreased a little from that of the original population. This data is in agreement with the next paragraph and reference 29 suggestive of a small genetic admixture between Hispanic Chilean population and Amerindians.

Response: Thank you for the observation. We corrected the phrase so that we clarify that LP state spread in Hispanic descending population, By Hispanics we means Mestizos of biparental origin (Spaniard and Amerindian). We expected the frequency of LCT-13910 CT/TT genotype in our population to be lower than in Spanish population, because our pre-columbian population was LNP (LCT-13910 CC).

9) Page 16 lines 20-37. The whole paragraph appears obscure. Indeed, at page 10 rows 3-8 the authors state: "The sample size was selected by estimating an expected prevalence of the LCT-13910CC genotype to be at least 50% in the Hispanic population". Therefore they were expecting a frequency of the LP determined by the LCT-13910CT or LCT-13910TT genotypes among Hispanics of 50%. Now the actual frequency of LP obtained in the group of 216 unrelated Hispanic was 43.1%, a frequency lower and not higher than what was expected (see also critiques 3 and 8). In addition at page 18 rows 6-8 the same authors state: "existence of a high prevalence of LNP state within this population" that is again in contradiction with rows 20-36 of page 16. Please revise, modify or delete the whole paragraph from row 20 to 36.

Response: Thank you for your observation. We revised the whole paragraph and modified the text so that we clarify that the idea refers to the deviation from HWE equilibrium we found.

10) Please double check the calculations obtained for reference of the HBT: sensitivity, specificity, PPV, NPV and LHR of the LCTC>T-13910 variant reported in table 1 and in the text. I have obtained slightly different numbers for sensitivity and NPV.

Response: We thank the reviewer for pointing out this. We decided to change the data of the table N°1 according to the suggestion of Reviewer Marcin Krawczyk. Data are now reporting the values for the Hydrogen breath test (HBT) against the genetic test; you can also find the respective changes in the text. In this respect, it is probably more reasonable to show PPV, NPV and so far using as reference test the HBT, because it shows present lactose malabsorption. However, we have to bear in mind that all CC genotype will develop LNP state (lactose malabsorption and positive HBT) at some point in their lives.

11) Table 2. Please correct "predictive value positive and predictive value negative" with "positive predictive value" and "negative predictive value", respectively.

Response: The new manuscript includes the corrections you suggested.

12) Table 3. Please control all the p values presented in the table. I did a chi square test and found a significant difference for the category: number reporting diarrhea.

Response: Thank you for your observation. We double checked and found a significant difference, thank you for noting this. The other p values was corrected.

Response to Reviewer Mr Nahum Méndez-Sánchez, MD, PhD (Professor of Medicine and Chief of the Department of Biomedical Research, Medica Sur Clinic & Foundation, Mexico).
1) The title is too long. Also since the authors did not study other groups of Latin American countries I suggest to delete Latin American populations
Response: We welcomed your suggestion about the title and made it shorter, also deleted the phrase "Latin American population" as we didn`t study other groups of Latin American countries.

2) A list of abbreviations is mandatory
Response: Thank you for your observation. We included a list of abbreviations at the beginning of the manuscript.

3) It is important to include in the introduction or subjects-methods sections the information on the ethnical backgrounds of Chilean population. Because it well known that it is different compared to other Latin American countries.
Response: Thank you for your observation. We included information about ethnical background of Chilean population since we have previously estimate in these cohorts a 40% and 80% Amerindian admixture in the Hispanic and Mapuche Chilean population, respectively. This was done by using an Amerindian Admixture Index (AAI) based on ABO blood group distribution. Additionally we added new references about Chilean ethnical background. It is interesting that the AAI of our Hispanic population (representative of >70% of the contemporary Chilean population) is similar to the AAI reported for Mexican American population in different papers.

Response to Mr Marcin Krawczyk, MD (Department of Medicine II, Saarland University Hospital, Germany).
1) Calculating positive and negative predictive values, specificity and sensitivity of genotyping against HBT should be avoided. Conversely, we propose to calculate the positive and negative predictive values, specificity, and sensitivity of HBT against genotyping.
Response: Thank you for your suggestion. We recalculated these values for the HBT as compared to the genetic test. We found a good correlation between both tests and are similar to those recently reported by other groups including yours (Krawczyk M, et al. Concordance of genetic and breath tests for lactose intolerance in a tertiary referral centre. J Gastrointestin Liver Dis 2008;17:135-139).

2) In the Chilean cohort, three individuals suffer from secondary hypolactasia (i.e. are HBT positive but carry the risk allele). The authors mention that these patients suffer from lactose malabsorption. The discussion of this interesting result needs to be improved, i.e. were there cases of secondary hypolactasia in other genetic studies.
Response: Thank you for this interesting observation. As we mentioned, it was not possible to recall these 3 patients to carry out complementary studies to rule out other causes of lactose malabsorption, such as intestinal parasitic infections, seronegative Celiac disease or other conditions as has been shown in other recent studies. We cited a study in which a similar situation was explained by Crohn disease and IBS.

3) Departure from HWE in the cohort of 216 Hispanics remains to be explained. Were any of these individuals related (family members)? Are the authors aware of other studies concerning this polymorphism that report deviation from HWE in selected populations?
Response: Thank you for your observation. Hispanic individuals were unrelated as described in

method section. Deviation from HWE could be explained by selection bias or by a higher frequency of LP in the Spanish population that founded the Chilean colony; it could also be reflective of positive selection for LP allele carriers. Yes, and we found and cited in our paper other studies reporting deviation of HWE in mixed populations.

4) In the Introduction the authors should clearly explain the difference in the ethnic background between Amerindians and Hispanics.
Response: Thank you for your suggestion. We included information about ethnical background of Chilean population since we had previously demonstrated in these cohorts a gradient of Amerindian admixture between Hispanic and Mapuche population. Additionally we added new references about the Chilean ethnical background.

5) In the Discussion (p. 15, line 18) the authors should also briefly mention that other LCT mutations have recently been identified in small African populations.
Response: We thank the reviewer for pointing out this point. In the new manuscript we clarify that the mentioned mutation is not the only responsible and that we mention one of the SNPs involved.

6) p. 4, line 50: should read 'non randomized'
7) p. 16, line 6: should read 'degree of'
8) p. 16, 2nd paragraph may be omitted.
9) p. 16, line 34: 'recently' should be deleted (refers to 2005).
10) Figure 1 is not necessary (standard technique).
Response: We thank the reviewer for pointing out these errors and the new manuscript includes all corrections you suggested.

11) Please include the ‘LCT’ in the abbreviations list under each table.
Response: Thank for your suggestion. As other reviewer suggested we added a list of abbreviations in the first page of the new manuscript.

Sincerely yours,

Juan Fco. Miquel P., MD.

VERSION 2 - REVIEW

REVIEWER	***Nahum Méndez-Sánchez***
REVIEW RETURNED	14-May-2011

GENERAL COMMENTS	Reviewer completed checklist only. No further comments were made

REVIEWER	***Marcin Krawczyk***
REVIEW RETURNED	16-May-2011

GENERAL COMMENTS	The authors satisfactorily addressed the issues and improved their manuscript according to the suggestions. I have no further comments.

REVIEWER	***Mauro Congia***
REVIEW RETURNED	20-May-2011

GENERAL COMMENTS	Reviewer completed checklist only. No further comments were made

VERSION 2 – AUTHOR RESPONSE

Thank you for your response. We resolved the two issues you mentioned in your letter. As you suggested the study design is included in the title. Also, we confirm that EM, JC and LA meet all three ICMJE criteria for authorship, so we amended the contributorship statement as you suggested. The new version of the paper contains all changes in red highlight text to facilitate the revision process. Also we correct the reference format and wrote the funding of our paper.