Author’s response to reviews

Title: Non-replication of an association of CTNNBL1 polymorphisms and obesity in a population of Central European ancestry

Authors:

Carla IG Vogel (carlavogel@yahoo.com)
Brandon Greene (brandon.greene@Staff.Uni-Marburg.de)
Andre Scherag (andre.scherag@uk-essen.de)
Timo D Müller (timo.mueller@uni-due.de)
Susann Friedel (susann.friedel@uni-duisburg-essen.de)
Harald Grallert (harald.grallert@helmholtz-muenchen.de)
Iris M Heid (heid@gsf.de)
Thomas Illig (illig@helmholtz-muenchen.de)
H-Erich Wichmann (wichmann@helmholtz-muenchen.de)
Johannes Hebebrand (Johannes.Hebebrand@lvr.de)
Anke Hinney (anke.hinney@uni-due.de)

Version: 2 Date: 27 January 2009

Author’s response to reviews: see over
We thank the reviewers for their valuable comments (below shown in italics). In the revised manuscript and in this letter we have addressed all their concerns.

Reviewer: HongWen Deng

Comments:
The CTNNBL1, FLJ42133, SH3PXD2B and SLIT3 genes were recently identified to be associated with obesity phenotypes (BMI and body fat mass) in a genome-wide association study in US Caucasians and a French population (Liu et al. Hum Mol Genet 2008). In this study, the authors aimed to replicate the association of these four genes with obesity in three independent German samples (a case-control sample consisting of 487 extremely obese children and 442 healthy lean individuals, a population-based adult sample consisting of 4261 adults, and a family-based sample consisting of 775 families). The authors claimed that no supportive evidence could be found for association of these reported genes with obesity in the German samples.

Comments:
1) Background:
a) The statement “These genes SH3PXD2B (rs13356223, rs10077897 and rs13436547), SLIT3 (rs17734503 and rs12654448) and FLJ42133 (rs7363432 and rs6095722) however, only showed association with increased BMI or fat mass in the initial sample, but could not be validated in the French sample is not correct. In Liu et al. (2008), real replication in the French sample was only performed for the most significant SNPs in the CTNNBL1 gene initially identified in the US Caucasians. Real replication effort for the SH3PXD2B, SLIT and FLJ42133 in the French sample was not pursued.

We agree with the reviewer and we have removed this sentence from the main text.

2) Methods:
b) The cases and controls were not well matched in the case-control sample. The cases were extremely obese children with of age 14.4±3.74, but the controls were healthy lean individual with age of 26.1±5.79 (who were likely to be young adults). This may cause nonreplication of the original association findings.

This was an issue raised by both reviewers. We agree with the referees that matching by age is a proper method to control for age confounding in a wide rage of epidemiological applications. For gene mapping of complex traits like obesity, however, there is the strong advantage of using normal and underweight adults that these are unlikely to become obese later in life, i.e. they should not be enriched for risk genotypes. Moreover, we checked that these individuals had normal or underweight already at the same age as the cases (Hinney et al., 2007). We added in the Methods section: “The use of lean adults that never suffered of overweight or obesity during the childhood (always were lean as accessed by interview [8]), as control group reduces the chances of misclassification compared to the use of lean children as controls that might become an overweight adult.”

c) Only 1644 probands of the KORA sample (which comprises 4261 adults) were genotyped. It is not clear why only a portion of the sample was genotyped for this replication study. It is known that, to achieve sufficient power, the replication sample should be much larger than the initial sample. A replication sample with 1644 subjects may not be sufficiently powerful to replicate the initial association findings.
The GWA for KORA was conducted only in a proportion of the population-based follow up of survey 3 (F3). From the original 3,126 individuals from KORA F3, only 1,644 individuals were genotyped using the Affymetrix® GeneChip® Human Mapping 500K Array Set (cooperation with Prof. Wichmann; Helmholtz Centre, Munich). We agree with the reviewer that KORA GWA by itself is not sufficiently powered to replicate the initial association findings (limited power of 10%) but the family study alone has a power > 80% which should be considered a well-power. This way we believe that with the combined power of the samples > 90% the study is well-powered to replicate the original findings in a Central European population (please see below in Results and Discussion (a) more details about the combined power calculation).

d) For the family sample, it is not clear how many children and parents were successfully genotyped.

775 families were genotyped in total and in 773 families we obtained complete genotypes for all family members (99.7%). We added this information in the Methods section: “…the call rate was 99.7%, with 100% concordance of duplicates.”

e) What’s the composition in terms of ethnicity for the three study samples?

The three study samples are composed of Caucasians from Germany, Central Europe. We added in the Methods section: “All individuals studied are Caucasians from Central Europe, with German ancestry.”

f) Statistics: potential population stratification was not considered in the case-control sample and the population-based KORA sample.

All individuals genotyped reside in Germany and most of which (at least based on self-reports) state that they were of German ancestry. The Case-control GWA was recruited in Marburg (located in West-Central Germany), but the participants are from all parts of Germany; the family-based study was recruited in Marburg and Essen (located in the North-West of Germany), again the participants are from all parts of Germany; the population-based KORA cohort was recruited in the region of Augsburg, Southern Germany. For Germany, empirical data support the idea that effects of population substructure on association studies under a case-control design are presumably very small in size [Steffens et al., 2006]. Moreover, we also genotyped nuclear-family samples for which population stratification issues do not arise. Taking these arguments together the authors do not think that population stratification is a major confounding factor of the analyses.

We have added in the Results and Discussion section: “Our failure to replicate the initial findings [6] also does not appear to be a result of population stratification. All recruitment was done in Germany for which population stratification effects have shown to be of minor importance [20].”

3) Results and discussion:

a) The statement “Nevertheless, the combined power, that is the probability of at least one of the three tests “re-detecting” an association, was > 90% if the estimated and the true genetic effects correspond” is problematic. The “combined power”
defined by the authors seems unreasonable. If none of the three samples is sufficiently powerful for replication purpose, it is not surprising that the initial findings could not be replicated in any of the three samples.

The family-based test (775 families with 1 or more extremely obese offspring) should be considered well-powered (> 80%) on its own. The “combined power” simply gives the probability that one or more of the tests used will correctly reject the null hypothesis when the true genetic effects are as strong as those detected by Liu et al. (2008). Put differently, assuming these effect estimated are true, the probability that all three tests would fail to replicate is < 10%.

“Combined power” =
Probability that at least one of the three – independent – studies replicates, when the true genetic effects are as strong as those detected in the original paper =
1 − probability that all three fail to replicate under these conditions =
1 − ( 1 − power of the first study) x ( 1 − power of the second study) x ( 1 − power of the third study) =
1 − ( 1 − 0.80) x ( 1 − 0.10) x ( 1 − 0.54) = 0.92

We have added in the Methods (Statistics) section: “With the “combined power” mentioned above, we refer to the probability that at least one of the tests used will correctly reject the null hypothesis when the true genetic effects are as strong as those described in [6]. This probability was > 90%, since 1 − probability that all three independent tests fail to replicate under these conditions = 1 − ( 1 − 0.80) * ( 1 − 0.10) * ( 1 − 0.54) = 0.92.”

b) It is not “surprising” that Liu et al. (2008) could not show association evidence for FTO. Many factors such as diverse linkage disequilibrium at the FTO locus between populations may result in this.

In general, we agree with the reviewer that such effects might lead to the absence of the effect or might at least contribute to the noise which complicates the detection. Thus, we have modified the statement: “While replication of association with obesity of intron 1 variants in FTO has been demonstrated robustly in almost all subsequent studies comprising obese adults and children [8-12], the study by Liu et al. [6] was an exception as none of the intron 1 FTO SNPs showed evidence for a body weight-related association. ”

4) Conclusion:
a) The claim that “no confirmation in a well-powered replication study” needs to be changed, as the study samples were not really sufficiently powerful.

The authors kindly ask the reviewer to reconsider his statement. We have shown that the family-study should be called ‘well-powered” for the replication of the initial findings of Liu et al. Moreover, the combined power was >90% to redetect/replicate the association of the CTNNB1 rs6013029 and obesity. This way the authors believe that the claim of the no confirmation in a well-powered replication study is valid. As a compromise, we added “in a population of Central European ancestry” to the title which picks up the idea raised by the reviewers comment 3.b)
Reviewer: Harald Staiger
Reviewer's report:
The manuscript presented by Vogel et al. describes lack of association of formerly reported obesity candidate SNPs in CTNNBL1 (and three other genes) with childhood/adolescence obesity in a case-control GWA study, lack of association with BMI in a population-based GWA study, and lack of overtransmission of the risk allele of CTNNBL1 SNP rs6013029 to obese children in a family study. In times where nearly every week new candidate genes for complex traits are published, it appears very important to publish well-designed replication studies including those presenting negative results. Even though this negative report is interesting and timely and appears sufficiently powered, the study also reveals serious flaws.

Major Compulsory Revisions
1. I have serious concern with the selection of the control group in the case-control GWA study. First, the controls are adults (mean age: 26 y), whereas the cases are children/adolescents (mean age: 14 y). When studying obese children/adolescents, it appears imperative to study controls of comparable age. Second, the controls have a mean BMI of 18.31 ±1.10. This suggests that the control group also includes anorectic subjects. Both points could introduce relevant bias into the study.

As suggested by both the reviewers we have now presented an explanation for the use of normal and underweight individuals as control group (see above). We added a statement in the Methods section. Further, we only used healthy individuals in this study, individuals who fulfilled criteria for anorexia nervosa were removed prior to the recent study.

2. In addition to BMI, more parameters of body adiposity and body fat distribution, such as fat mass measured by bioimpedance, waist circumference, and waist-hip ratio, should be available in the KORA cohort. In order to substantiate their negative findings, the authors should perform additional regression analyses with these parameters and include the data in the manuscript (or provide them as supplementary material).

Fat-mass, waist and waist-hip ratios are BMI-correlated anthropometric measurements, so the authors believe that performance of additional analyses would to a great extend mirror the BMI effect (no association of CTNNBL1 and obesity). Furthermore, we explicitly intended to replicate the initial findings of the study by Liu and the parameters available in the original study were BMI and fat mass, with this last not used in the replication study (French case-control sample). Assessing more phenotypes, our study could haven been easily rejected with the ‘multiplicity’ argument.

3. Since the risk alleles of the CTNNBL1 SNPs are known and, thus, the direction of the SNP effects is given, one-sided statistical tests are justified. Even though the data presented by the authors appear to point in the opposite direction, as compared to the initial study by Liu et al., the authors should add one-sided p-values to the table to further stress that the initially described SNP effects are not seen in these cohorts.

As requested by the reviewer, one-sided p-values were calculated and added to the table.
4. The authors report lack of association between SNPs in FLJ42133, SH3PXD2B, and SLIT3 with obesity as ‘data not shown’. Please, include these data in the manuscript (or provide them as supplementary material).

As requested by the reviewer we have included the data in an additional table (Table 2) in the main manuscript.

Minor Essential Revisions
How was the combined power of all three studies calculated? Please, provide information about this in the Methods section.

See the comments to the other reviewer above regarding “combined power”.

Discretionary Revisions
1. Even though the information about the power of the studies is sufficient, wouldn’t it be more informative to indicate these studies’ minimal effect sizes/odds ratios detectable with sufficient power (1-##0.8)?

We agree with the reviewer that this is a good idea; for a minor allele frequency of 10%, α=0.05 (one-sided) under a (log-)additive genetic model and the rare disease assumption which should hold for extremely obese subjects (by definition) a power of about 80% will be achieved for effect sizes of 1.33 for the family-based approach, of 1.44 for the case-control approach and of 0.15 in standardized effects for the population-based sample. Thus, except for the nuclear families all other samples are underpowered to detect realistic effect sizes of complex traits (at least at this minor allele frequency).

2. Methods section, paragraph ‘Participants and Genotyping’: standard deviations should not be presented with higher precision than the mean values.

We apologize for the mistake and we have corrected the paragraph according to the reviewer’s suggestion.