Reviewers' comments:

Reviewer #1 (Remarks to the Author):

Justice et al. analyzed GWAS data for the detection of novel variants for obesity-related traits. Different to previous approaches smoking behavior, that is known to influence body weight regulation, was taken into account. GWAS data of more than 50,000 current smokers and more than 190,000 non-smokers were used for the analyses. A total of 23 novel loci plus 9 loci with convincing evidence for SNP - smoking interaction were detected. A large number of in silico analyzes (eQTL, expression, epigenetic, PubMed, biological pathway analyzes to name a few) were performed for all newly derived genomic regions. Interesting and partially unexpected new interactions and biological pathways were detected.

Analyses are up to date taking into account a large variety of publically available data bases pertaining to function, biological pathways and epigentics.

The conclusions are robust, valid and reliable.

Discussion is thoughtful and comprises all relevant aspects.

The approach elegantly shows that taking environmental factors into account might be very useful to detect genetic loci that could not be identified by analysis of the trait under study alone.

Reviewer #2 (Remarks to the Author):

A. Summary of the key results
This study used GWAS data, including approximately 51,000 current smokers and 190,000 nonsmokers, mostly of European descent, to identify novel loci influencing general and central adiposity, as measured by various combinations of weight and height, and waist and hip circumference. In all the authors identified 23 novel genetic loci related to these adiposity indices, and nine loci where the genetic effects appeared to be modified by smoking status.
The authors noted that main effects of genetic loci may sometimes be masked by the presence of opposite effects among smokers and non-smokers. Further, some genetic effects may present as main effects, but actually carry stronger influence in smokers than in non-smokers (or vice-versa).

B. Originality and interest: if not novel, please give references
 I was very impressed by this comprehensive piece of work which should increase our understanding of the relationship between adiposity and smoking.

C. Data & methodology: validity of approach, quality of data, quality of presentation
The authors used four "approaches" to modelling in their GWAS analysis - (1) only considering main effects, (2) considering main effects in the presence of interaction, (3) considering interaction effects across the whole genome, (4) considering interaction effects only in variants where main effects have already been established
D. Appropriate use of statistics and treatment of uncertainties
I cannot really comment on the biological plausibility of the findings, but I have four main statistical comments which to some extent the authors have already addressed

1. In clinical trials, the presence of subgroup effects, whereby the effect of a treatment may be modified by another factor such as sex or age group, may occur in two different ways, often called qualitative and quantitative interactions. Qualitative interactions occur when the effect of the treatment in one subgroup is in the opposite direction to the other. Pocock et al 2002 (Statistics in Medicine) suggest these are rare and usually implausible. Quantitative interactions occurs when the treatment effect is stronger in one subgroup than another. In searching for such interactions, there are special issues over both Type 1 and Type 2 errors. There is considerably less power to detect interactions than main effects. Also the hunt for interactions may among a plethora of possibilities may lead to false positive findings.

The study of Justice et al seems to find both types of interaction. They frequently refer to reference 9 when Taylor et al, studying a single SNP related to smoking quantity in the CHRNA5-CHRNA3-CHRNB4 locus, showed an unexpected positive effect of the minor allele on BMI among never-smokers, together with the expected inverse effect among smokers. Justice et al appear to reproduce this finding but also show analogous effects among other gene loci. The question is - how plausible are such qualitative interactions, and can we be sure they are not simply chance?

To be fair, Figure 3 helpfully demonstrates considerable power to detect qualitative interactions (if they actually exist), especially for approach 3. Neither approach appears at all powerful to detect quantitative interactions, however. Does this suggests that findings of apparent quantitative interactions are most likely to be chance? The same may be true also of qualitative interactions if their existence is actually very rare.

When carrying out a GWAS on main effects (approaches 1 and 2), the authors use $p<5*10-8$ and they do the same when searching for interactions in approach 3. Did they consider a double correction when searching for both main effects and interactions?

2. The investigators decided to lump ex-smokers with never smokers into a general "non-smoker" category. While this has the virtue of simplicity, it may miss some nuances of the effects. Taylor et al showed that while opposite effects were seen for current smokers and never-smokers, findings for ex-smokers fell neatly in between (as might be expected, given the in between exposure to smoking that this group would have experienced). I was especially curious as to the effects of R2 of the lumping by Justice et al: this would have induced heterogeneity of exposure to smoking, possibly thus reducing the R2 for gene vs adiposity trait. The smokers meanwhile are also a heterogeneous group, with a mixture of light and heavy smokers. I think this bears particularly on some results of the enrichment analysis (lines 679-691), where no difference in beta coefficients is discernible between smokers and non-smokers, but differing R2's are reported.
I also suggest further analysis on novel hits, to see if effects for ex-smokers fall in between those for never-smokers and current smokers, for some of the studies for which the lead authors have rapid access to individual data.

3. Three trait indices are used all of which are complex combinations of anthropometric measures. Even well-known BMI involves a ratio of weight and height2. There may well be a residual relationship of this index with height. The authors respond by asserting that no known loci for say height are contiguous with the hits found by the present study. I think this is only an indirect defence (dependent on arbitrary significance levels for GWAS's conducted on height). Especially for WCadjBMI, and especially WHRadjBMI (which is arguably a double adjustment for general adiposity), one wonders if the indices represent artefacts. For WCadjBMI, a sensitivity analysis might involve say log(WC) as dependent variable with log(weight) and height as covariates. I know it takes a long time to run sensitivity analysis when you are dependent on hundreds of co-investigators of the primary studies to

run these. Perhaps the authors have individual data quickly available for a few of the bigger studies to see if novel hits still exhibit directionally similar betas when such a sensitivity analysis is carried out?

4. I was curious at the novel use of four analytic approaches. Approaches 1 and 2 are both looking for main effects, but approach 2 allows for interactions while approach 1 does not. This must mean that if interactions do actually exist, then approach 1 results in undue precision for the main effect. The table at the foot of Figure 1 partly bears this out: fewer loci and fewer novel loci are detected by approach 2 than approach 1. Does approach 1 give some false positive findings for main effects if interactions are present? It would be interesting to see a cross tabulation of effects detected by the two approaches. I imagine that any hits found by approach 2 but not approach 1 is likely to be chance alone.

At the same time, that table at the foot of Figure 1 suggests very few extra hits found in approaches 3 and 4 compared with approaches 1 and 2. This is probably because of the lack of power to detect interactions compared with main effects (see my point 1 above). Also, approach 4 is capable of detecting only quantitative interactions (the sort that are considered more plausible in a clinical trial context), but not qualitative ones.

E. Conclusions: robustness, validity, reliability
Need to check whether some novel hits are really plausible or simply false positives, in view of comments under D above.

F. Suggested improvements: experiments, data for possible revision
See my suggestion of sensitivity analysis for the adiposity indices, especially WCadjBMI (comment D3), and for dividing non-smokers into never and current (comment D2).

G. References: appropriate credit to previous work?
Am not qualified to appraise this fully

H. Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions
The abstract is nice to read. But the term GXSMK is undefined - I knew what the authors meant but it may not be obvious to all. Also the last sentence, when referring to "genetic susceptibility" - does the directions of effect depend on whether you are discussing the minor or major allele for each individual SNP?

REVIEWERS' COMMENTS:

Reviewer #2 (Remarks to the Author):

I appreciate the thorough further work carried out by the authors to address my previous concerns about Type i error rates.

I have no further comments.

Response to Reviewer Comments: "Genome-Wide Meta-Analysis of 241,258 Adults Accounting for Smoking Behavior Identifies Novel Loci for Obesity Traits"

We thank the two reviewers for their detailed comments, which we feel have led to changes that greatly enhanced our manuscript. We have taken great care to revise our work in line with all of the reviewer's concerns. Below you will find a detailed response for each point raised and references to any related text corrections or alterations (highlighted in yellow in the attached revised manuscript). To allow for these adjustments to the text and to stay within the manuscript page limits, we have removed some text, tables, and references, as highlighted in the revised draft. Please note that all line numbers referenced below refer to the line numbers contained in the revised draft, as line numbers are not used in the final draft.

REVIEWER #1:

Justice et al. analyzed GWAS data for the detection of novel variants for obesity-related traits. Different to previous approaches smoking behavior, that is known to influence body weight regulation, was taken into account. GWAS data of more than 50,000 current smokers and more than 190,000 non-smokers were used for the analyses. A total of 23 novel loci plus 9 loci with convincing evidence for SNP - smoking interaction were detected. A large number of in silico analyzes (eQTL, expression, epigenetic, PubMed, biological pathway analyzes to name a few) were performed for all newly derived genomic regions. Interesting and partially unexpected new interactions and biological pathways were detected.

Analyses are up to date taking into account a large variety of publically available data bases pertaining to function, biological pathways and epigentics.

The conclusions are robust, valid and reliable.

Discussion is thoughtful and comprises all relevant aspects.

The approach elegantly shows that taking environmental factors into account might be very useful to detect genetic loci that could not be identified by analysis of the trait under study alone.

RESPONSE:
We thank the reviewer for this appreciation of the novelty of our analytical approaches and results. We hope that our responses to other concerns raised will only strengthen your opinion and appreciation for the proposed manuscript.

REVIEWER #2:

1. Summary of the key results
This study used GWAS data, including approximately 51,000 current smokers and 190,000 nonsmokers, mostly of European descent, to identify novel loci influencing general and central adiposity, as measured by various combinations of weight and height, and waist and hip circumference. In all the authors identified 23 novel genetic loci related to these adiposity indices, and nine loci where the genetic effects appeared to be modified by smoking status.

The authors noted that main effects of genetic loci may sometimes be masked by the presence of opposite effects among smokers and non-smokers. Further, some genetic effects may present as main effects, but actually carry stronger influence in smokers than in non-smokers (or vice-versa).

B. Originality and interest: if not novel, please give references

I was very impressed by this comprehensive piece of work which should increase our understanding of the relationship between adiposity and smoking.

C. Data & methodology: validity of approach, quality of data, quality of presentation
The authors used four "approaches" to modelling in their GWAS analysis - (1) only considering main effects, (2) considering main effects in the presence of interaction, (3) considering interaction effects across the whole genome, (4) considering interaction effects only in variants where main effects have already been established

RESPONSE:
We appreciate that the reviewer has recognized the main strengths and findings of our manuscript. We hope that the responses to the below criticisms and concerns have improved confidence in the results and have only acted to improve your overall evaluation of the manuscript.

D. Appropriate use of statistics and treatment of uncertainties
I cannot really comment on the biological plausibility of the findings, but I have four main statistical comments which to some extent the authors have already addressed

1. *In clinical trials, the presence of subgroup effects, whereby the effect of a treatment may be modified by another factor such as sex or age group, may occur in two different ways, often called qualitative and quantitative interactions. Qualitative interactions occur when the effect of the treatment in one subgroup is in the opposite direction to the other. Pocock et al 2002 (Statistics in Medicine) suggest these are rare and usually implausible. Quantitative interactions occurs when the treatment effect is stronger in one subgroup than another. In searching for such interactions, there are special issues over both Type 1 and Type 2 errors. There is considerably less power to detect interactions than main effects. Also the hunt for interactions may among a plethora of possibilities may lead to false positive findings.*

The study of Justice et al seems to find both types of interaction. They frequently refer to reference 9 when Taylor et al, studying a single SNP related to smoking quantity in the CHRNA5-CHRNA3-CHRNB4 locus, showed an unexpected positive effect of the minor allele on BMI among never-smokers, together with the expected inverse effect among smokers. Justice et al appear to reproduce this finding but also show analogous effects among other gene loci. The question is - how plausible are such qualitative interactions, and can we be sure they are not simply chance?

To be fair, Figure 3 helpfully demonstrates considerable power to detect qualitative interactions (if they actually exist), especially for approach 3. Neither approach appears at all powerful to detect quantitative interactions, however. Does this suggests that findings of apparent quantitative interactions are most likely to be chance? The same may be true also of qualitative interactions if their existence is actually very rare.

When carrying out a GWAS on main effects (approaches 1 and 2), the authors use $p<5*10-8$ and they do the same when searching for interactions in approach 3. Did they consider a double correction when searching for both main effects and interactions?

RESPONSE:
We thank the reviewer for these important comments that led us to include simulations to evaluate type 1 error for the various approaches and also to extend our power analyses. We want to note that under the assumption of true biological interaction, the main effect will range between the two stratified effects, which may lead to a loss of power for Approach 1 to find main effects, yet constant type 1 error rates. Also, Aschard et al. 2010 demonstrated increased power under valid type 1 error for Approach 2 given some interaction exists. Therefore,

any hit found by Approach 2 alone should be due to larger power of the joint test that accounts for interaction rather than increased type 1 error rate. In order to illustrate the validity of the approaches with regards to type 1 error, we conducted simulations. We assumed 50,000 smokers and 180,000 nonsmokers and generated 10,000 independent simulated stratum-specific effect sizes under the stratum-specific null hypotheses of "no stratum-specific effects". We applied the four approaches to the simulated stratum-specific association results and inferred type 1 error of each approach by visually examining QQ plots (new **Supplementary Fig. 7**) and by calculating type 1 error rates (new **Supplementary Table 17**). We observed uninflated QQ plots as well as low type 1 error rates (ranging from 4.53% to 5.24%) at the 5% nominal level for any simulation scenario and across approaches. Since we are employing conservative Bonferroni-corrected significance levels for each approach, this suggests that any positive finding is equally likely to be chance – irrespective of the utilized approach.

As noted by the reviewer, the power of Approaches 3 and 4 to identify interaction effects differs by type of interaction. In order to provide more details, we have extended our analytical power computations to additional scenarios of varying effect sizes (added small and large effects sizes to the already included medium effect size considerations, see newly extended **Figure 3**). Indeed, power to find quantitative interaction is generally lower than finding pure (effect in one subgroup, zero effect in the other) or qualitative (opposite effects) interaction – a circumstance that mirrors our observation of mostly pure and qualitative interactions. Yet, we want to note to the reviewer that approach 4 shows substantial power to find quantitative effects when large genetic effect sizes are involved. For example, the power of approach 4 to identify interaction comprising a large genetic effect in non-smokers (R^2_{NONSMK}=0.34%, comparable to the realistic *FTO* effect on BMI) that is half (R^2_{SMK}=0.17%) or quarter (R^2_{SMK}=0.09%) as large in smokers, was >99% or 53.3%, respectively (new **Figure 3F**).

We sought additional evidence for validation for the top signals identified in the GIANT meta-analysis in an independent study from the UK Biobank (UKBB). We found consistent direction of effect for each of our novel loci and those with evidence of GxSMK interaction. We also meta-analyzed the GIANT and UKBB results and found that the majority of loci remained significant at the conservative Bonferroni-corrected significance levels (P<5E-8) for the respective Approach. We have added these validation results to the main text (See below).

We are confident that these additional analyses illustrate that our models are statistically sound, are not the results of inflated Type I error. While our findings are compelling, especially in light of follow-up analyses presented in our paper, only repeated replication and functional assessments of genes underlying these signals can ultimately confirm these association results.

To add further confidence to the future application of our chosen Approaches, we have also added the following text to describe these findings.

We chose to apply two analytical approaches to detect genetic main effects on adiposity traits. To illustrate the validity and utility of applying both approaches, we performed simulations to infer type 1 error (**Online Methods**, **Supplementary Table 17** and **Supplementary Fig 7**) and analytical power computations (**Figure 3**) for Approaches 1 and 2, and a heat map providing the cross-tabulation of P values for Approach 1 and 2 along with Approach 3 examining interaction only (**Supplementary Fig 8**). We demonstrate that the two approaches yield valid type 1 error rates and that Approach 1 can be more powerful to find associations given zero or negligible quantitative interactions, whereas Approach 2 is more efficient in finding associations given some interaction effects.

Validation of Novel Loci

We pursued validation of the associations of our novel and interaction SNPs in an independent study sample of up to 119,644 European adults from the UK Biobank study (**Tables 1-4**, **Supplementary Table 25, Supplementary Fig 9**). We found consistent directions of effects in smoking strata (for Approach 2 and 3) and in the SNPadjSMK model (Approach 1) for each locus examined (**Supplementary Fig. 13**). For BMI, 3 SNPs were not GWS (P>5E-8) following meta-analysis with our GIANT results, including one of the two variants near *EPAH3* for Approach 1; rs1809420 identified in Approach 4, which lies within a known locus near *ADAMTS7*, remained significant for interaction, but not for SNPadjSMK; and rs336396 near *INPP4B* identified in Approach 3. For WCadjBMI, 3 SNPs were not GWS (P>5E-8) following meta-analysis with our GIANT results, including rs1545348 near *RAI14* for Approach 1; rs4141488 near *GRIN2A* identified in Approach 3; and rs6012558 near *PRNP* identified in Approach 3. For WHRadjBMI, only 1 SNP from Approach 4 was not significant following meta-analysis with our GIANT results, rs12608504 near *JUND*, remained GWS for SNPadjSMK, but was only nominally significant for interaction (P_{int}=0.013).

In the Discussion:

There are several challenges in the ability to validate genetic associations that account for environmental exposure. In addition to exposure harmonization and potential bias due to adjustment for smoking exposure, differences in trait distribution, environmental exposure frequency, ancestry-specific LD patterns and allele frequency across studies may lead to difficulties in replication, especially for gene-by-environment studies[71,72]. Further, the problem of "winner's curse" (inflated discovery effects estimates) require larger sample sizes for adequate power in replication studies[73]. Despite these challenges, we were able to detect consistent direction of effect in an independent sample for all novel loci. Some results that did not remain GWS in the GIANT + UKBB meta-analysis still had results that were just under the threshold for significance, suggesting that a larger sample may be needed to confirm these results, and thus the associations near *INPP4B*, *GRIN2A*, *RAI14*, *PRNP*, and *JUND* should be interpreted with caution.

…

Eighteen of these remained significant in our validation with the UK Biobank sample.

Additionally, we edited the abstract as follows:

We identify 23 novel genetic loci, and nine loci with convincing evidence of gene-smoking interaction (GxSMK) on obesity-related traits. We show consistent direction of effect for all identified loci and significance for 18/23 novel and for 5/9 interaction loci in an independent study sample.

2. The investigators decided to lump ex-smokers with never smokers into a general "non-smoker" category. While this has the virtue of simplicity, it may miss some nuances of the effects. Taylor et al showed that while opposite effects were seen for current smokers and never-smokers, findings for ex-smokers fell neatly in between (as might be expected, given the in between exposure to smoking that this group would have experienced). I was especially curious as to the effects of R2 of the lumping by Justice et al: this would have induced heterogeneity of exposure to smoking, possibly thus reducing the R2 for gene vs adiposity trait. The smokers meanwhile are also a heterogeneous group, with a mixture of light and heavy smokers. I think this bears particularly on some results of the enrichment analysis (lines 679-691), where no difference in beta coefficients is discernible between smokers and non-smokers, but differing R2's are reported.

I also suggest further analysis on novel hits, to see if effects for ex-smokers fall in between those for never-smokers and current smokers, for some of the studies for which the lead authors have rapid access to individual data.

RESPONSE:

We agree with the reviewer that it is quite possible that our means of harmonizing the smoking variable introduced heterogeneity into our effect estimates. However, by allowing for this definition of our smoking exposure, we have greatly increased the sample size. Where there are loci with large differences in effects between current and former smokers, there may yet be additional loci to be discovered. In order to determine if our new loci deviated strongly from the expectation of intermediate effects in former smokers, we have examined the association results in some of our larger available studies (ARIC, FRAMHS, FAMHS, KORA, and WHI) with available data on current vs. former smoking status. Additional large datasets would be needed to fully address the question of whether or not smoking intensity plays a role in these associations. Beta estimates from former smokers exhibited intermediate effects between current and never smokers for 20 of the 34 SNPs tested. For the *CHRNB4* locus identified in the Taylor et al. paper, we show similar evidence of direction of effects with opposite direction between current smokers and never smokers and intermediate effect estimates for former smokers. We have included a table below summarizing these meta-analyses results. Given our small stratum-specific sample sizes, and thus lack of power to detect significant associations, these results should be interpreted with caution and are thus not included in our manuscript.

Response to Comment 2. Association results for all top novel and interaction loci stratified by current smokers, former smokers, and never smokers, main effects adjusted for smoking, and heterogeneity test to asses interaction between smoking strata. We show the results for each locus based on the stratum in which it was identified (e.g. men, women, combined sexes).

dbSNPID	Gene	App	CHR	POS(hg19)	A1	A2	Current Smokers					Former Smokers					Never Smokers					Main and Interaction Effects						Former Smokers Interm.
							EAF	BETA	SE	P	N	EAF	BETA	SE	P	N	EAF	BETA	SE	P	N	EAF	$BETA_{adjSMK}$	SE_{adjsmk}	P_{adjsmk}	P_{het}	N	
BMI																												
rs2481665	INADL	1:EC	1	62594677	t	c	0.585	0.012	0.017	4.7E-01	7,966	0.597	0.025	0.015	9.2E-02	10,559	0.598	0.007	0.013	5.7E-01	15,985	0.594	0.014	0.008	9.4E-02	6.5E-01	34,510	
rs10929925	SOX11	1,2	2	6155557	a	c	0.458	-0.016	0.016	3.2E-01	7,603	0.459	-0.021	0.015	1.7E-01	9,397	0.461	-0.028	0.012	2.2E-02	14,384	0.460	-0.023	0.008	5.4E-01	8.3E-01	31,384	Y
rs6794880	SRRM1P2	1	3	84451512	a	g	0.838	0.014	0.024	5.7E-01	6,654	0.841	0.009	0.023	6.9E-01	7,934	0.840	0.043	0.018	1.9E-02	12,352	0.840	0.025	0.012	3.9E-02	4.3E-01	26,940	
rs2173039	EPAH3	1:EW	3	89142175	c	g	0.258	0.015	0.025	5.4E-01	4,377	0.246	0.034	0.023	1.5E-01	5,589	0.249	0.036	0.017	3.0E-02	10,953	0.250	0.031	0.012	9.8E-03	7.7E-01	20,919	Y
rs12629427	EPAH3	1:AW	3	89145340	t	c	0.742	-0.014	0.025	5.7E-01	4,376	0.754	-0.032	0.023	1.7E-01	5,589	0.751	-0.036	0.017	3.3E-02	10,951	0.750	-0.030	0.012	1.2E-02	7.7E-01	20,916	Y
rs13069244	CCDC39	2	3	180441172	a	g	0.076	0.025	0.032	4.4E-01	7,637	0.074	0.058	0.029	4.8E-02	9,977	0.074	0.051	0.025	4.0E-02	15,723	0.074	0.046	0.016	4.4E-03	7.4E-01	33,337	Y
rs336396	INPP4B	3	4	143062811	t	c	0.141	0.082	0.027	2.2E-03	5,932	0.138	0.008	0.025	7.5E-01	7,349	0.140	-0.034	0.021	1.0E-01	11,214	0.139	0.009	0.014	5.2E-01	**2.8E-03**	24,495	Y
rs12902602	CHRNB4	2,3	15	78967401	a	g	0.611	0.076	0.017	8.8E-06	7,971	0.625	0.029	0.015	6.0E-02	10,562	0.625	-0.008	0.013	5.6E-01	15,994	0.621	0.025	0.009	3.7E-03	**4.7E-04**	34,527	Y
rs1809420	ADAMTS7	1,4:AM	15	79056769	t	c	0.598	0.057	0.027	3.6E-02	3,194	0.588	0.017	0.022	4.2E-01	4,604	0.589	0.009	0.022	6.8E-01	4,629	0.591	0.024	0.014	7.4E-02	3.7E-01	12,427	Y
WCadjBMI																												
rs17396340	KIF1B	1,2	1	10286176	a	g	0.123	0.005	0.027	8.6E-01	7,390	0.122	-0.022	0.022	3.1E-01	10,442	0.122	0.030	0.019	1.0E-01	15,433	0.122	0.008	0.013	5.5E-01	1.9E-01	33,265	
rs6743226	HDLBP	1,2	2	242236972	t	c	0.496	-0.021	0.017	2.1E-01	7,387	0.507	-0.015	0.014	2.8E-01	10,438	0.511	-0.009	0.012	4.6E-01	15,429	0.506	-0.014	0.008	8.9E-02	8.3E-01	33,254	Y
rs4378999	DOCK3	1	3	51208646	a	t	0.870	-0.065	0.032	4.1E-02	6,074	0.866	-0.032	0.027	2.3E-01	7,814	0.871	-0.061	0.023	7.7E-03	11,792	0.869	-0.053	0.015	5.7E-04	6.5E-01	25,680	Y
rs7697556	ADAMTS3	1,2	4	73515313	c	g	0.515	0.005	0.017	7.7E-01	7,390	0.524	0.009	0.015	5.4E-01	10,441	0.528	0.026	0.012	3.7E-02	15,431	0.524	0.016	0.008	5.8E-02	5.4E-01	33,262	Y
rs1545348	RAI14	1:EM	5	34718343	t	g	0.744	0.051	0.029	7.8E-02	3,338	0.742	0.032	0.024	1.8E-01	4,902	0.744	0.012	0.024	6.2E-01	4,883	0.743	0.029	0.015	4.4E-02	5.8E-01	13,123	Y
rs10269774	CDK6	1	7	92253972	a	g	0.356	0.043	0.019	2.6E-02	6,074	0.350	0.013	0.017	4.6E-01	7,814	0.354	0.009	0.014	5.4E-01	11,792	0.353	0.018	0.010	5.5E-02	3.3E-01	25,680	Y
rs6470765	GSDMC	1	8	130736697	a	c	0.719	0.041	0.022	6.2E-02	6,074	0.742	0.000	0.020	9.9E-01	7,814	0.726	0.043	0.016	7.5E-03	11,792	0.729	0.030	0.011	6.3E-03	2.0E-01	25,680	
rs9408815	TMEM38B	2	9	108890521	c	g	0.755	0.029	0.023	2.1E-01	6,074	0.752	0.055	0.020	5.9E-03	7,814	0.752	0.031	0.017	6.2E-02	11,792	0.753	0.038	0.011	6.6E-04	5.9E-01	25,680	
rs9409082	TMEM38B	1	9	108901049	t	c	0.231	-0.036	0.024	1.3E-01	6,074	0.236	-0.049	0.021	1.8E-02	7,814	0.234	-0.022	0.017	2.1E-01	11,792	0.234	-0.034	0.012	3.5E-03	6.0E-01	25,680	
rs4141488	GRIN2A	3	16	9629067	t	c	0.492	0.016	0.019	4.1E-01	6,074	0.498	-0.011	0.017	5.2E-01	7,814	0.491	-0.030	0.014	3.5E-02	11,792	0.493	-0.013	0.009	1.7E-01	1.6E-01	25,680	Y
rs6076699	PRNP	2,3:EW	20	4566688	a	g	0.904	0.091	0.064	1.6E-01	2,844	0.917	-0.055	0.067	4.1E-01	3,119	0.910	-0.016	0.045	7.3E-01	7,321	0.910	0.002	0.032	9.4E-01	2.5E-01	13,284	
rs6012558	ARFGEF2	1	20	47531286	a	g	0.421	0.014	0.017	4.3E-01	7,390	0.428	0.004	0.015	7.9E-01	10,442	0.420	0.010	0.013	4.5E-01	14,467	0.423	0.009	0.008	3.0E-01	9.0E-01	32,299	
WHRadjBMI																												
rs765751	LYPLA1	4	1	219669226	t	c	0.381	0.004	0.018	8.4E-01	6,786	0.388	-0.008	0.015	6.0E-01	9,380	0.384	-0.025	0.013	5.2E-02	13,252	0.385	-0.013	0.009	1.3E-01	4.0E-01	29,418	Y
rs670752	BBX	1:AW	3	107312980	a	g	0.300	-0.002	0.025	9.3E-01	3,747	0.300	0.027	0.022	2.2E-01	4,923	0.298	0.016	0.016	3.4E-01	9,432	0.299	0.015	0.012	2.0E-01	6.8E-01	18,102	Y
rs30000	MAP3K1	4:EM	5	55803533	a	g	0.689	0.019	0.030	5.3E-01	3,039	0.700	-0.022	0.024	3.7E-01	4,457	0.702	-0.053	0.026	4.0E-02	3,820	0.698	-0.022	0.015	1.5E-01	1.9E-01	11,316	Y
rs459193	MAP3K1	4:AM	5	55806751	a	t	0.277	-0.011	0.029	7.1E-01	3,039	0.270	0.025	0.024	3.0E-01	4,457	0.271	0.053	0.026	4.0E-02	3,820	0.272	0.025	0.015	9.9E-02	2.6E-01	11,316	Y
rs1049281	HLA-C	1,2	6	31236567	t	c	0.326	-0.065	0.022	3.1E-03	4,748	0.323	-0.026	0.019	1.8E-01	6,167	0.325	-0.020	0.017	2.3E-01	8,474	0.325	-0.033	0.011	2.5E-03	2.4E-01	19,389	Y
rs589428	EHMT2	1:EC	6	31848220	t	g	0.298	-0.020	0.021	3.1E-01	6,568	0.292	-0.055	0.018	1.8E-03	8,574	0.293	-0.018	0.015	2.4E-01	12,287	0.294	-0.030	0.010	2.4E-03	2.3E-01	27,429	
rs7766106	RSPO3	4	6	127455138	t	c	0.457	0.045	0.018	1.1E-02	6,786	0.450	0.041	0.015	5.6E-03	9,375	0.448	0.047	0.013	2.0E-04	13,248	0.451	0.045	0.008	1.3E-07	9.5E-01	29,409	
rs1856293	EYA4	2:EC	6	1334480940	a	c	0.493	0.015	0.020	4.6E-01	5,470	0.503	0.007	0.017	6.8E-01	6,752	0.492	-0.030	0.015	4.5E-02	9,612	0.496	-0.007	0.010	4.9E-01	1.2E-01	21,834	Y
rs2001945	TRIB1	1:AW	8	126477978	a	c	0.570	-0.004	0.024	8.6E-01	3,747	0.617	-0.015	0.022	5.1E-01	4,923	0.588	-0.029	0.016	6.6E-02	9,433	0.592	-0.020	0.011	8.2E-02	6.7E-01	18,103	Y
rs754133	HOXC4-HOXC6	4:EM	12	54418920	a	g	0.339	-0.026	0.029	3.4E-01	3,039	0.340	0.014	0.022	5.4E-01	4,457	0.344	0.040	0.025	1.1E-01	3,820	0.341	0.012	0.014	4.1E-01	2.0E-01	11,316	Y
rs2071449	HOXC4-HOXC6	4:AM	12	54428011	a	c	0.354	-0.024	0.029	4.1E-01	2,706	0.349	0.007	0.023	7.8E-01	4,157	0.352	0.042	0.026	1.0E-01	3,550	0.351	0.010	0.015	4.9E-01	2.3E-01	10,413	Y
rs17065323	SMIM2	1:EC	13	44627788	a	c	0.130	0.121	0.054	2.4E-02	2,630	0.135	-0.009	0.040	8.3E-01	5,740	0.135	0.022	0.034	5.1E-01	9,350	0.132	0.030	0.023	1.9E-01	1.5E-01	17,720	
rs12608504	JUND	4:AM	19	18389135	a	g	0.367	-0.019	0.027	4.9E-01	3,038	0.367	-0.004	0.023	8.7E-01	4,456	0.365	0.037	0.024	1.2E-01	3,819	0.366	0.006	0.014	6.5E-01	2.6E-01	11,313	Y

3. *Three trait indices are used all of which are complex combinations of anthropometric measures. Even well-known BMI involves a ratio of weight and height[2]. There may well be a residual relationship of this index with height. The authors respond by asserting that no known loci for say height are contiguous with the hits found by the present study. I think this is only an indirect defence (dependent on arbitrary significance levels for GWAS's conducted on height). Especially for WCadjBMI, and especially WHRadjBMI (which is arguably a double adjustment for general adiposity), one wonders if the indices represent artefacts. For WCadjBMI, a sensitivity analysis might involve say log(WC) as dependent variable with log(weight) and height as covariates. I know it takes a long time to run sensitivity analysis when you are dependent on hundreds of co-investigators of the primary studies to run these. Perhaps the authors have individual data quickly available for a few of the bigger studies to see if novel hits still exhibit directionally similar betas when such a sensitivity analysis is carried out?*

RESPONSE:
We apologize for any confusion caused by the wording of our conditional analyses on known height variants. In fact, we intended to identify cases in which our novel loci from primary meta-analyses were dependent or independent of known height loci. These are highly correlated anthropometric traits which are known to share underlying biological pathways and shared SNP-trait associations (Shungin et al. 2015); therefore we do expect to find that some new loci associated with WC and WHR are known height loci.

However, as the reviewer points out, associations may exist with other loci in our results that have not been reported in the GWAS Catalog. Accordingly, we conducted a lookup of all of our novel variants (lookups for known loci were reported in cross-trait association analysis in Locke et al. and Shungin et al.) in results from a previous GWAS of Height (Wood et al. 2014). Overall, there are few additional variants that may be associated with height, but not previously reported in GWAS examining height (2 for WHRadjBMI and 2 for WCadjBMI with P<0.002[0.05/24 SNPs]) or in previous cross-trait association efforts.

We have added these lookup results to our **Supplementary Table 18**, along with other previous GWAS lookups, and added the following text to describe these results.

==Given the high phenotypic correlation between WC and WHR with height, and established shared genetic associations that overlap across all three of our adiposity traits and height [1, 2, 68] we would expect additional cross-trait associations between our novel loci and height that may have been missed due to the strict threshold requirements for reporting associations in the GWAS Catalog. Therefore, we conducted a look-up of all of our novel loci in a recent large GWAS meta-analysis of height[68] to identify additional overlapping association signals (**Supplementary Table 18**). No novel BMI loci were significantly associated with height (P<0.002[0.05/24] SNPs). However, there are additional variants that may be associated with height, but not previously reported in GWAS examining height, including 2 for WHRadjBMI near *EYA4* and *TRIB1*, and 2 for WCadjBMI near *KIF1B* and *HDLBP* (P<0.002).==

4. *I was curious at the novel use of four analytic approaches. Approaches 1 and 2 are both looking for main effects, but approach 2 allows for interactions while approach 1 does not. This must mean that if interactions do actually exist, then approach 1 results in undue precision for the main effect. The table at the foot of Figure 1 partly bears this out: fewer loci and fewer novel loci are detected by approach 2 than approach 1. Does approach 1 give some false positive findings for main effects if interactions are present? It would be interesting to see a cross tabulation of effects detected by the two approaches. I imagine that any hits found by approach 2 but not approach 1 is likely to be chance alone.*

At the same time, that table at the foot of Figure 1 suggests very few extra hits found in approaches 3 and 4

compared with approaches 1 and 2. This is probably because of the lack of power to detect interactions compared with main effects (see my point 1 above). Also, approach 4 is capable of detecting only quantitative interactions (the sort that are considered more plausible in a clinical trial context), but not qualitative ones.

RESPONSE:

We thank the reviewer for these valid comments and want to note that under the assumption of true biological interaction, the main effect will range between the two stratified effects. For example, opposite effects between strata may fully cancel out, resulting in a zero main effect. As a consequence and for given interaction, we expect lower power with Approach 1 to find main effects, yet constant type 1 error rates. We do not expect false positive findings for Approach 1 given interaction. Moreover, we want to note that previous work has demonstrated increased power under valid type 1 error for Approach 2 given some interaction (Aschard et al 2010). Therefore, hits found by Approach 2 alone are likely due to larger power of the joint test to account for interaction rather than increased type 1 error rate. To further illustrate this and to improve clarity, we added simulations to infer type 1 error of Approaches 1 and 2 (**Supplementary Table 17** and **Supplementary Fig. 7**), analytical power computations for Approaches 1 and 2 (**Figure 3**), and a heat map providing the cross-tabulation of Approaches 1 and 2 along with Approach 3 (**Supplementary Fig. 8**). We demonstrate that the two approaches yield valid type 1 error rates and that Approach 1 can be more powerful to find associations given zero or very tiny quantitative interactions, whereas Approach 2 is more efficient in finding associations given some underlying interaction effects.

Please see additions to text noted above in the response to concern 1.

E. Conclusions: robustness, validity, reliability
Need to check whether some novel hits are really plausible or simply false positives, in view of comments under D above.

Please see our responses as noted above.

F. Suggested improvements: experiments, data for possible revision
See my suggestion of sensitivity analysis for the adiposity indices, especially WCadjBMI (comment D3), and for dividing non-smokers into never and current (comment D2).

Please see our responses as noted above

G. References: appropriate credit to previous work?
Am not qualified to appraise this fully.

H. Clarity and context: lucidity of abstract/summary, appropriateness of abstract, introduction and conclusions

The abstract is nice to read. But the term GXSMK is undefined - I knew what the authors meant but it may not be obvious to all. Also the last sentence, when referring to "genetic susceptibility" - does the directions of effect depend on whether you are discussing the minor or major allele for each individual SNP?

RESPONSE:

We thank the reviewer for bringing this to our attention. We have edited the abstract to clearly define GXSMK. Additionally, our results are all oriented on the obesity-increasing allele in the smokers. We understand that the interpretation of findings (increasing vs. decreasing) is always relative. In the case, we have chosen this orientation for its epidemiological significance. We have edited the abstract as follows:

"smoking may alter the genetic susceptibility to overall adiposity and body fat distribution."

Response to Reviewer Comments: "Genome-Wide Meta-Analysis of 241,258 Adults Accounting for Smoking Behavior Identifies Novel Loci for Obesity Traits"

We would like to thank the reviewers for their willingness to review our revised manuscript and we are glad that they have found our revisions have addressed all of their concerns.

REVIEWER #1:

None provided.

No comments.

REVIEWER #2:

I appreciate the thorough further work carried out by the authors to address my previous concerns about Type i error rates.

I have no further comments.

No comments.