

## Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided  
*Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g.  $F$ ,  $t$ ,  $r$ ) with confidence intervals, effect sizes, degrees of freedom and  $P$  value noted  
*Give  $P$  values as exact values whenever suitable.*
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's  $d$ , Pearson's  $r$ ), indicating how they were calculated

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection No software was used for data collection purpose

Data analysis R script to perform gene prioritization approach proposed in this manuscript is available at <https://doi.org/10.5281/zenodo.6598552>. The following software and packages were used for data analysis: PLINK 2.0 (<https://www.cog-genomics.org/plink/2.0/>), METAL (<http://csg.sph.umich.edu/abecasis/Metal/download/>), MAGMA v1.07 (<https://ctg.cncr.nl/software/magma>), FUMA (<https://fuma.ctglab.nl/>), LDscore (<https://github.com/bulik/ldsc>), LDstore v1.1 (<http://www.christianbenner.com/#>), FINEMAP v1.3.1 (<http://www.christianbenner.com/#>), PRSice v2.2.1 (<https://www.prsice.info/>), Eagle2 (<https://alkesgroup.broadinstitute.org/Eagle/downloads/>), Minimac3 (<https://genome.sph.umich.edu/wiki/Minimac3>), REGENIE v2.0.1 (<https://rgcgithub.github.io/regenie/>), MiXeR (<https://github.com/precimed/mixer>), BUHMBOX (<https://software.broadinstitute.org/mpg/buhmbox/>), R v3.6.0 (<https://www.r-project.org/>) with packages data.table v1.12.2, GenomicRegion v1.36.0, stats v3.6.3, fpc v2.2-3, coloco v3.2-1, Rtsne v0.15 and ggplot2 v3.2.0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The full GWAS summary statistics for UKB and top 10,000 SNPs for 23andMe are available at [https://ctg.cncr.nl/software/summary\\_statistics/](https://ctg.cncr.nl/software/summary_statistics/). The full GWAS

summary statistics for the 23andMe data set will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit <https://research.23andme.com/collaborate/#publication> for more information and to apply to access the data. The following publicly available datasets were used in this manuscript: GTEx v8 (<https://gtexportal.org/home/datasets>), Allen Human Brain Atlas (<http://human.brain-map.org/static/download>), scRNA-seq from Linnerson's lab (<http://linnarssonlab.org/data/>; GSE60361, GSE74672, GSE75330, GSE76381, GSE97478), DropViz (<http://dropviz.org/>), MsigDB v6.2 (<http://software.broadinstitute.org/gsea/msigdb/index.jsp>), InWeb protein-protein interaction ([https://www.intomics.com/inbio/map/api/get\\_data?file=InBio\\_Map\\_core\\_2016\\_09\\_12.tar.gz](https://www.intomics.com/inbio/map/api/get_data?file=InBio_Map_core_2016_09_12.tar.gz)), eQTLGen (<https://www.eqtlgen.org/>), PsychEncode (<http://resource.psychencode.org/>).

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For all samples the sample size consists of all individuals that remain after quality control of the data and exclusion of withdrawn subjects. Detailed information on the samples used, as well as the exclusion/inclusion criteria, are provided in the Method (Sample cohorts section).
Data exclusions	See "Sample cohorts" section in the Method. All individuals of non-European ancestry, individuals who failed quality control, or individuals who asked to withdraw from the study were removed.
Replication	Both UKB and 23andMe datasets were included in the primary meta-analysis. We explicitly examined the concordance of effects within the novel loci between these two parts of our data. In addition, we checked how our results replicated previously reported loci. This information is available in Supplementary Table 5. There were 18 out of 202 loci from the previous insomnia meta-analysis (Jansen et al. Nat. Genet. 2019) that were no longer significant in this study. These loci are most likely false positive in the previous study as they showed significantly higher P-value compared to replicated loci (details are discussed in Supplementary Note).
Randomization	Covariates were used to control for differences in age, ancestry, and sex.
Blinding	This analysis was exploratory so blinding was not applicable.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	We selected UKB 386,989 unrelated participants who had insomnia phenotype and assigned to European ancestry (208,958 females and 178,030 males). Participants are aged between 38 and 73 years old. All individuals included in the study provided informed consent.
Recruitment	Participants were recruited based on participation in biobanks
Ethics oversight	The UKB received ethical approval from the National Research Ethics Service Committee North West-Haydock), and all study procedures were in accordance with the World Medical Association for medical research.

Note that full information on the approval of the study protocol must also be provided in the manuscript.