

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 [Authors & Referees](#) 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 the collection of data, as this was an opportunistic study.

Data analysis All code to perform quality control and data analysis is provided in the following Github repos:

https://github.com/macarthur-lab/gnomad_qc
https://github.com/macarthur-lab/gnomad_lof
<https://github.com/konradjk/loftee>

Hail 0.2 is available at: <https://hail.is/>

Picard version 1.1431

VerifyBamID version 1.0.0

GATK nightly-2015-07-31-g3c929b0, 3.4-89-ge494930, and 3.6-0-g89b7209

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/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

All datasets are described in the manuscript or Supplementary Information, including deposition of the full dataset at <https://gnomad.broadinstitute.org>, a browser described in the Data Availability section. Data for all figures is available accordingly. There are no restrictions on the aggregate data released.

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://www.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	This study was opportunistic, and involved secondary use of all available genome and exome data. No sample size was predetermined. Nevertheless, the current sample size enables the accurate assessment of constraint against pLoF variation for over 72% of genes in the human genome (see Figure 2).
Data exclusions	Sample QC and variant QC for gnomAD are described extensively in the supplementary methods. Notably, individuals with severe pediatric disease, and known first disease relatives of those with severe pediatric disease were excluded, as previously established and described [Lek et al., 2016].
Replication	We did not attempt to reproduce any findings in a separate dataset, as no other data set of comparable size exists.
Randomization	As this was a population-based study, and not a case-control study, no randomization was performed.
Blinding	As this was a population-based study, and not a case-control study, blinding was not relevant.

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	Involvement 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
<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

Methods

n/a	Involvement 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	As an opportunistic collection of data, the participants in this study were not selected based on age, gender, or genotypic information. As described above, individuals with severe pediatric disease, and known first disease relatives of those with severe pediatric disease were excluded. The populations are provided in Supplementary Table 7, and there are 64,754 females and 76,702 males. These data were obtained primarily from case-control studies of adult-onset common diseases, including cardiovascular disease, type 2 diabetes, and psychiatric disorders.
Recruitment	As this was an opportunistic secondary use study, we did not recruit any participants.
Ethics oversight	This study was overseen by the Broad Institute's Office of Research Subject Protection and the Partners Human Research Committee, and was given a determination of Not Human Subjects Research.

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