naturereserach

Corresponding author(s):   Matthew Hurles

☐ Initial submission   ☐ Revised version   ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

▶ Experimental design

1. **Sample size**

   Describe how sample size was determined.

   | No statistical methods were used to predetermine sample size. |

2. **Data exclusions**

   Describe any data exclusions.

   | Contaminated samples or other samples failing quality control were compiled into a 'blacklist' that was excluded a priori from all analyses. |

3. **Replication**

   Describe whether the experimental findings were reliably reproduced.

   | The experimental findings have not been replicated in an independent cohort. |

4. **Randomization**

   Describe how samples/organisms/participants were allocated into experimental groups.

   | There was no randomization of experimental groups. |

5. **Blinding**

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   | Investigators were not blinded group allocation. |

   Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. **Statistical parameters**

   For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒	A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	A statement indicating how many times each experiment was replicated
☐	☒	The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒	A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒	The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒	A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒	Clearly defined error bars

   See the web collection on statistics for biologists for further resources and guidance.

▶ Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

Burrows-Wheeler Aligner (version 0.59) was used for short-read sequence mapping. Variant calling was performed using GATK and de novo mutation calling was performed using DeNovoGear. Subsequent analyses based on this data used custom R scripts, all of which have been deposited in a github repository with links included in the manuscript. In R, Bioconductor version 3.1 was used to install and import phastCons100way.UCSC.hg19 v3.6.0 and GenomicRanges v1.30.0.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

▶ Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

There were no unique materials used in this analysis.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

There were no antibodies used in this analysis.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used.

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used, so no cell line authentication was required.

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used, so no mycoplasma contamination testing was needed.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.

No eukaryotic cell lines were used.

▶ Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

No research animals were used in this study.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Patients with severe, undiagnosed DDs and their parents were recruited and systematically phenotyped at 24 clinical genetics centres within the United Kingdom National Health Service and the Republic of Ireland. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South Research Ethics Committee and GEN/284/12, granted by the Republic of Ireland Research Ethics Committee). Families gave informed consent for participation.