

The contribution of X-linked coding variation to severe developmental disorders

Hilary C. Martin^{1,*}, Eugene J. Gardner¹, Kaitlin E. Samocha¹, Joanna Kaplanis¹, Nadia Akawi^{1,2}, Alejandro Sifrim^{1,3}, Ruth Y. Eberhardt¹, Ana Lisa Taylor Tavares^{1,4,5}, Matthew D. C. Neville¹, Mari E. K. Niemi^{1,5}, Giuseppe Gallone^{1,7}, Jeremy McRae^{1,8}, Caroline F. Wright⁹, David R. FitzPatrick¹⁰, Helen V. Firth^{1,5}, Matthew E. Hurles¹, on behalf of the Deciphering Developmental Disorders study

1. Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK
2. Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK.
3. Department of Human Genetics, University of Leuven, Leuven, Belgium
4. Department of Clinical Genetics, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK.
5. Genomics England, Queen Mary University of London, London, EC1M 6BQ, UK
6. Institute for Molecular Medicine Finland, University of Helsinki, Tukholmankatu 8, Helsinki, FI-00014, Finland.
7. Max Planck Institute for Molecular Genetics, Ihnestraße 63, 14195, Berlin
8. Illumina Inc., 5200 Illumina Way, San Diego, CA, 92122, USA
9. Institute of Biomedical & Clinical Science, University of Exeter Medical School, Exeter, EX2 5DW, UK
10. MRC Human Genetics Unit, MRC IGMM, University of Edinburgh, Western General Hospital, Edinburgh, EH4 2XU, UK

* Correspondence to hcm@sanger.ac.uk

Abstract

Over 130 X-linked genes have been robustly associated with developmental disorders (DDs), and X-linked causes have been hypothesised to underlie the higher DD rates in males. We evaluated the burden of X-linked coding variation in 11,046 DD patients, and found a similar rate of X-linked causes in males and females (6.0% and 6.9%, respectively), indicating that such variants do not account for the 1.4-fold male bias. We developed an improved strategy to detect novel X-linked DDs and identified 23 significant genes, all of which were previously known, consistent with our inference that the vast majority of the X-linked burden is in known DD-associated genes. Importantly, we estimated that, in male probands, only 13% of inherited rare missense variants in known DD-associated genes are likely to be pathogenic. Our results demonstrate that statistical analysis of large datasets can refine our understanding of modes of inheritance for individual X-linked disorders.

Introduction

Several attributes of X-chromosomal biology render it unique among chromosomes, and have profoundly influenced the landscape of X-linked monogenic disorders. The hemizyosity of the X chromosome in males results in a distinctive male-specific pattern of segregation in pedigrees for X-linked recessive disorders, which has facilitated the recognition of such disorders and catalysed the identification of the underlying associated genes¹. By contrast, X-linked dominant disorders do not result in such characteristic segregation patterns in pedigrees, and are expected predominantly in females due to the considerably lower mutation rate of the maternally-inherited X chromosome in males².

X-linked recessive disorders can be caused by both *de novo* mutations (DNMs) and maternally inherited variants. In 1935, Haldane showed that the relative contribution of these two classes is a function of the reproductive fitness of a disorder and the mutation rate in the paternal and maternal germline³. Specifically, the fraction of male X-linked recessive cases due to DNMs should be $\frac{m\mu}{2\mu+v}$, where m is the reproductive loss in affected males, μ is the mutation rate in eggs and v in the mutation rate in sperm. For developmental disorders (DDs) which are reproductively lethal (i.e. $m=1$), if the maternal and paternal germline mutations rates were equal, one-third of pathogenic variants would be expected to be *de novo*. However, the paternal mutation rate of single nucleotide variants (SNVs), the predominant class of pathogenic variant, is ~ 3.5 times higher than the maternal rate⁴, which would be expected to lower the proportion of pathogenic SNVs that arise *de novo* to $\sim 18\%$ (see Methods).

While most X-linked disorders exhibit a profound sex-bias, suggestive of the underlying mode of inheritance, it is frequently observed that both sexes can manifest the same disorder. There are several possible explanations, which are not mutually exclusive. Skewing of X chromosome inactivation in females (which normally achieves dosage compensation) provides a mechanism for some female carriers of pathogenic variations in X-linked recessive genes to manifest disease of varying severity levels, although extreme skewing is rare on a population level^{5,6}; this mechanism has previously been inferred to occur in 7.6% of female patients with intellectual disability⁷. One alternative explanation is an incompletely penetrant dominant phenotype associated with a variant that is fully penetrant when hemizygous; a special case of this is semi-dominance, in which heterozygous females are affected but hemizygous males are seldom observed due to lethality. Another explanation is that a disorder is truly dominant such that the hemizygous and heterozygous phenotypes are identical. These complexities have led some to recommend that the field refer

collectively to ‘X-linked disorders’, avoiding explicit classification based on their individual modes of inheritance^{8,9}. Nonetheless, X-linked Mendelian disease genes are often classified as ‘X-linked recessive’ (XLR) or ‘X-linked dominant’ (XLD) eg. by OMIM.

The majority of X-linked monogenic disorders that have been identified are DDs, especially neurodevelopmental disorders such as intellectual disability (ID). The highly-curated DDG2P (Developmental Disorder Gene-to-Phenotype) database¹⁰ contains over 130 DD-associated genes on the X chromosome, most of which are observed predominantly in males and are presumed X-linked recessive disorders. Analyses of new large datasets of population variation have called into question a few of these gene associations¹¹, although most remain robust. Importantly, mutations in the same gene can cause more than one condition: 39 of the 132 X-linked genes in DDG2P are associated with more than one syndrome, many of them with different modes of inheritance and different mechanisms (e.g. loss-of-function versus activating). It has been suggested that the 1.3-fold male bias in the incidence of ID can be largely attributed to the male-biased contribution of X-linked disorders^{12,13}, although this has not been formally demonstrated.

The recent availability of exome sequencing, large cohorts of both cases and controls, and a fine-grained understanding of the germline mutation rate¹⁴, have together empowered ‘burden’ analyses which can quantify the absolute and relative contributions of different classes of inherited and *de novo* variation to particular disorders^{15–18}. Here we analyse exome sequencing data from 11,046 families in the Deciphering Developmental Disorders (DDD) study to estimate the relative contribution of X-linked causes of DDs to both male and female probands.

Results

Comparison of male versus female phenotypes in DDD

There are 40% more male than female probands in the DDD study (7,843 males versus 5,619 females), similar to the bias reported in other ID/DD cohorts¹¹, but lower than the four-fold male bias reported in autism cohorts¹⁹. We compared the phenotypes of male versus female probands in the study to explore whether phenotypic differences might be contributing to recruitment bias in males. Males tended to be recruited ~4.8 months earlier than females (linear regression $p=0.0004$), so we controlled for age at assessment in the following tests of phenotypic differences. Males had slightly more affected organ systems than females, although this was only nominally significant (mean difference = 0.06; linear regression $p=0.049$; Supplementary Fig. 1). There were significant differences in the prevalence of several phenotypic features between the sexes (Supplementary Table 1). For example, after correcting for age at assessment, males were 2.4-times more likely to have an abnormality of the genitourinary system (logistic regression $p=1.3\times 10^{-48}$), 2.1-times more likely to have autistic behaviour ($p=8\times 10^{-41}$), and 2.0-times more likely to show hyperactivity ($p=7\times 10^{-10}$). However, none of these differences were large enough in magnitude to suggest that they made a major contribution to recruitment bias. Males in DDD were significantly taller (linear regression $p=8\times 10^{-8}$) and had greater occipital frontal circumference at recruitment ($p=4\times 10^{-17}$) than females (measured relative to the sex- and age-adjusted distributions in the general population) (Supplementary Table 2). Curiously, we observed that males walked on average 1.2 months earlier than females in the study (linear regression $p=2\times 10^{-7}$) (Supplementary Table 2), although in the general population, they tend to walk about two weeks later than females on average²⁰. This may correspond with the observation that females in the cohort are slightly more likely than males to have severe ID/DD (12.4% versus 10.9%; logistic regression $p=0.02$) but equally likely to have mild or moderate ID/DD (Supplementary Table 2). These comparisons suggest that although there are

some significant differences in average phenotypes between the sexes in DDD, these are small in magnitude and males and females are broadly similar in clinical presentation.

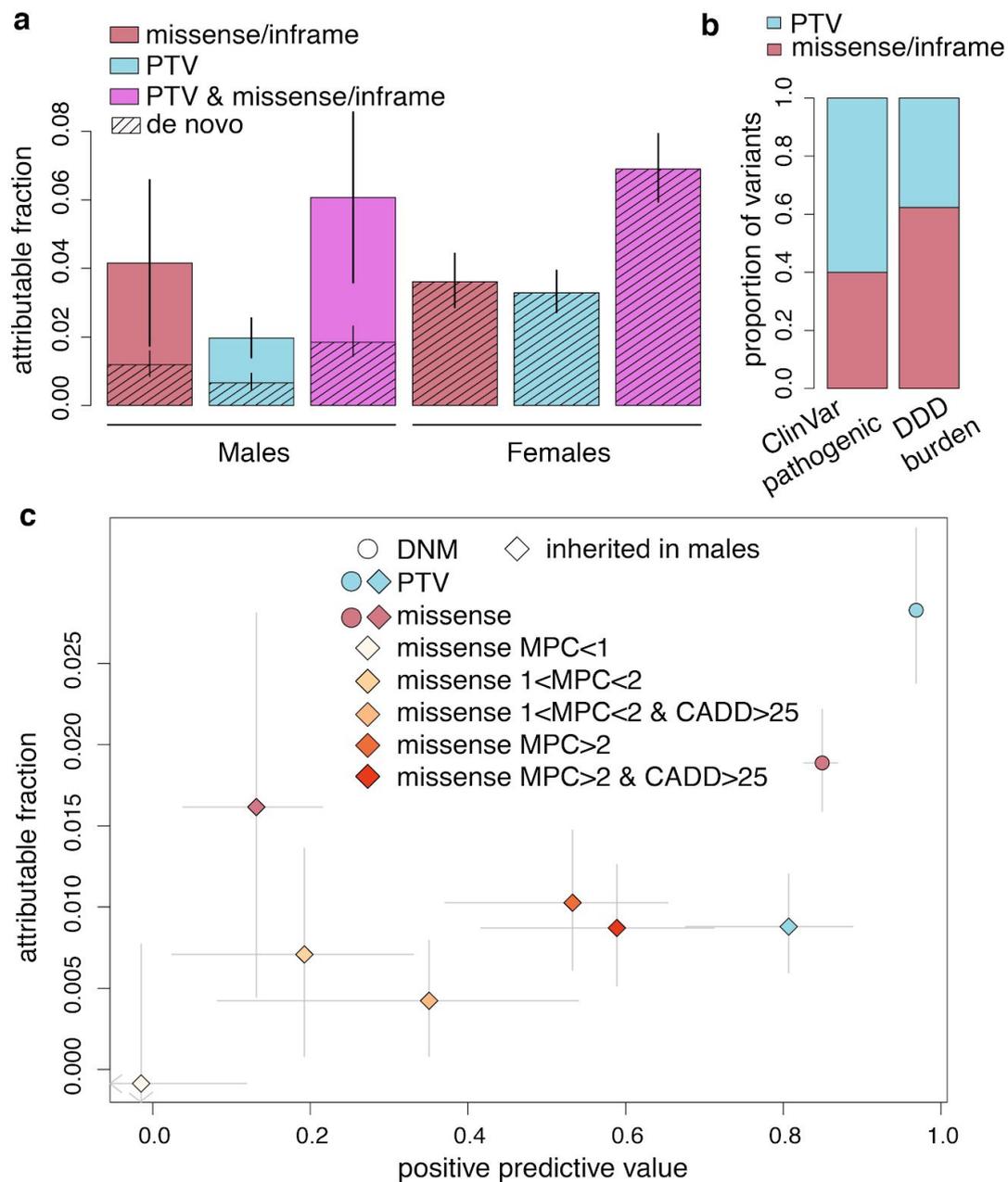
X chromosome burden analysis

We hypothesised that the higher number of males in DDD might be due to variation on the X chromosome, since males, being haploid, might be more vulnerable to pathogenic variants on this chromosome and could be affected by both *de novo* and maternally inherited variants. To avoid biases that would be introduced by considering only diagnoses in known DD genes, we carried out a sex-specific burden analysis to estimate the fraction of patients attributable to rare or *de novo* coding variants in *all* genes in the nonpseudoautosomal regions of the X chromosome, assuming a monogenic model with full penetrance. We focused on 7,138 independent male probands (5,138 in family trios) and 3,908 independent female probands in trios, and considered variants with minor allele frequency <0.1% and with no hemizygotes in the gnomAD resource of population variation²¹. In male probands, we performed a case/control analysis, comparing probands to 8,551 unaffected DDD fathers. This analysis implicitly includes both inherited variants and DNMs. For the purposes of some analyses, we also conducted an enrichment analysis of DNMs alone in the male trio probands, with quality control optimised to detect DNMs (Supplementary Table 3), and compared to a null mutation model¹⁴. In female trio probands, we performed a DNM enrichment analysis (Supplementary Table 3), assuming that there would be very few inherited pathogenic variants on the X chromosome in females since the vast majority of parents are unaffected.

Overall, we estimated from the burden analysis that 6.0% of males (95% confidence interval [3.6%-8.6%]) and 6.9% of females [5.9%-7.9%] had a pathogenic X-linked protein-truncating or missense/inframe variant (Fig. 1A). This implies that monogenic X-linked coding causes of DDs are not the cause of the male bias in DDD. In females, 90% [83%-97%] of the burden was in DD-associated genes, versus 63% [44%-100%] in males (95% confidence intervals from bootstrapping shown in Supplementary Fig. 2A,B). In trio males, 41% [23-100%] of the burden was *de novo* and the rest inherited (Supplementary Fig. 2C).

The relative contribution of missense/inframe variants was higher in males than females, but not significantly so (Fig. 1A). Overall, 38.0% of the X-linked exonic burden was driven by protein-truncating variants (PTVs), with the rest being missense/inframe (Fig. 1B). In contrast, in a set of 4,318 variants in X-linked DD-associated genes reported as “pathogenic” or “likely pathogenic” in ClinVar, 60.0% were PTVs and 40.0% were missense/inframe variants. This is significantly different from our burden analysis of DD-associated X-linked genes (Fisher’s exact test $p=2 \times 10^{-20}$; Fig. 1B), and likely reflects the fact that PTVs are easier to interpret and hence more likely to be considered pathogenic or likely pathogenic by clinical geneticists and genetic diagnostic laboratories.

Figure 1: Results from burden analysis of rare and *de novo* coding variants on the X chromosome. a) Fraction of males and females attributable to rare inherited and *de novo* coding variants on the X chromosome. Note that in the males, the overall attributable fraction was estimated from the case/control analysis of all male probands (7,138 cases versus 8,551 controls), whereas that for DNMs was estimated only in the 5,138 male trio probands. b) Relative fraction of PTV versus missense/inframe variants amongst ClinVar likely pathogenic or pathogenic variants in X-linked DDG2P genes, versus the fraction inferred in the burden analysis in DDD. c) Estimated attributable fraction versus positive predictive value for DNMs and inherited variants in males in X-linked DDG2P genes. Inherited missense variants are split according to CADD²² and MPC²³ scores. In (a) and (c), error bars show 95% confidence intervals calculated as described in the Methods.



We next estimated the fraction of the observed rare inherited or *de novo* variants in known X-linked DD genes were actually likely to be pathogenic (i.e. the positive predictive value, PPV) (Fig. 1C). These PPVs can assist accurate diagnostic interpretation by providing prior probabilities of pathogenicity for different classes of variation. For *de novo* PTVs and *de novo* missense mutations and inherited rare PTVs in males, the PPV was >80%. However, the PPV for inherited rare missense variants (MAF<0.001) that were not observed as hemizygotes in gnomAD was estimated to be only 13.2% [3.9%-21.5%], indicating that there is a substantial risk of incorrect diagnosis and hence clinical mismanagement. In line with this, of variants passing these filters that have been reported in DECIPHER and rated by clinicians, 56/272 (20.6% [15.9-25.9%]) were classed as pathogenic or likely pathogenic, and 47.8% as ‘uncertain’ [41.7-53.9%]. If we did not apply the additional filter requiring zero hemizygotes in gnomAD in our burden analysis, the PPV was not significantly different from 0, although it increased to 15.6% [5.4%-24.6%] if we reduced the MAF filter from 0.1% to 0.005% (Supplementary Fig. 3). We were able to increase the PPV to ~60% by applying more stringent filters on CADD²² and MPC²³ scores to the missense variants (Fig. 1C), indicating that such filters are likely to aid clinical genetics practice by reducing rates of incorrect assignment of pathogenicity. However, Fig. 1C also indicates that this improvement in specificity is counterbalanced by some reduction in sensitivity. In line with this, of 56 inherited X-linked variants that have been reported in DECIPHER and rated pathogenic or likely pathogenic by clinicians, 52 (93%) had MPC>1 and 38 (68%) had MPC>2, whereas the PPVs for variants with these filters were 31% and 53% respectively. These results indicate the value of this population-based burden analysis for informing improvements in clinical practice.

Contribution of *de novo* versus inherited variants in X-linked recessive genes in males

Based on Haldane’s theory³ and under certain assumptions, we anticipated that ~18% of pathogenic variants in XLR genes in males would be *de novo* (see Methods). However, using the excess number of variants (observed-expected), we found that this fraction was ~36%. However, even within this large dataset, the bootstrap confidence intervals around this estimate remain large ([19-83%]) (Supplementary Fig. 2D). Using an approximation to a binomial distribution, we estimate that the probability of observing our data under the null hypothesis of Haldane’s theory is about 0.0001 (two-sided p-value from a two-sample test for equality of proportions). Together with our results from bootstrapping (Supplementary Fig. 2D), this suggests that our data are unlikely to be consistent with Haldane’s theory. There are at least three non-mutually exclusive explanations for this. Firstly, the DDD study may be biased away from classic inherited X-linked families because these are easier to diagnose through the usual clinical means and because patients who had been previously recruited to the UK-wide GOLD study focused on X-linked ID¹ tended not to be recruited to DDD. Secondly, it may be that our assumption about the ratio of male to female mutation rates differs on the X chromosome compared to the autosomes; however, the male:female mutation rate ratio would have to be substantially lower on the X chromosome than the autosomes to be consistent with the observed 36% DNMs, which seems unlikely. Finally, Haldane’s theory did not incorporate the possibility of a reduced number of offspring in heterozygous carrier females, which would be expected to increase the proportion of DNMs in XLR genes. We note that this need not be due to any physiological phenotype in carrier females, and could include the effect of women choosing not to have more children after having one or more affected sons. This phenomenon is well recognised in current clinical practice and can result in halving the number of offspring of women known to be at risk of being carriers²⁴.

We tested the effect of rare PTVs in XLR genes in females in UK BioBank (N=13 carriers) and found that carrier females had a nominally significant reduced number of children (average

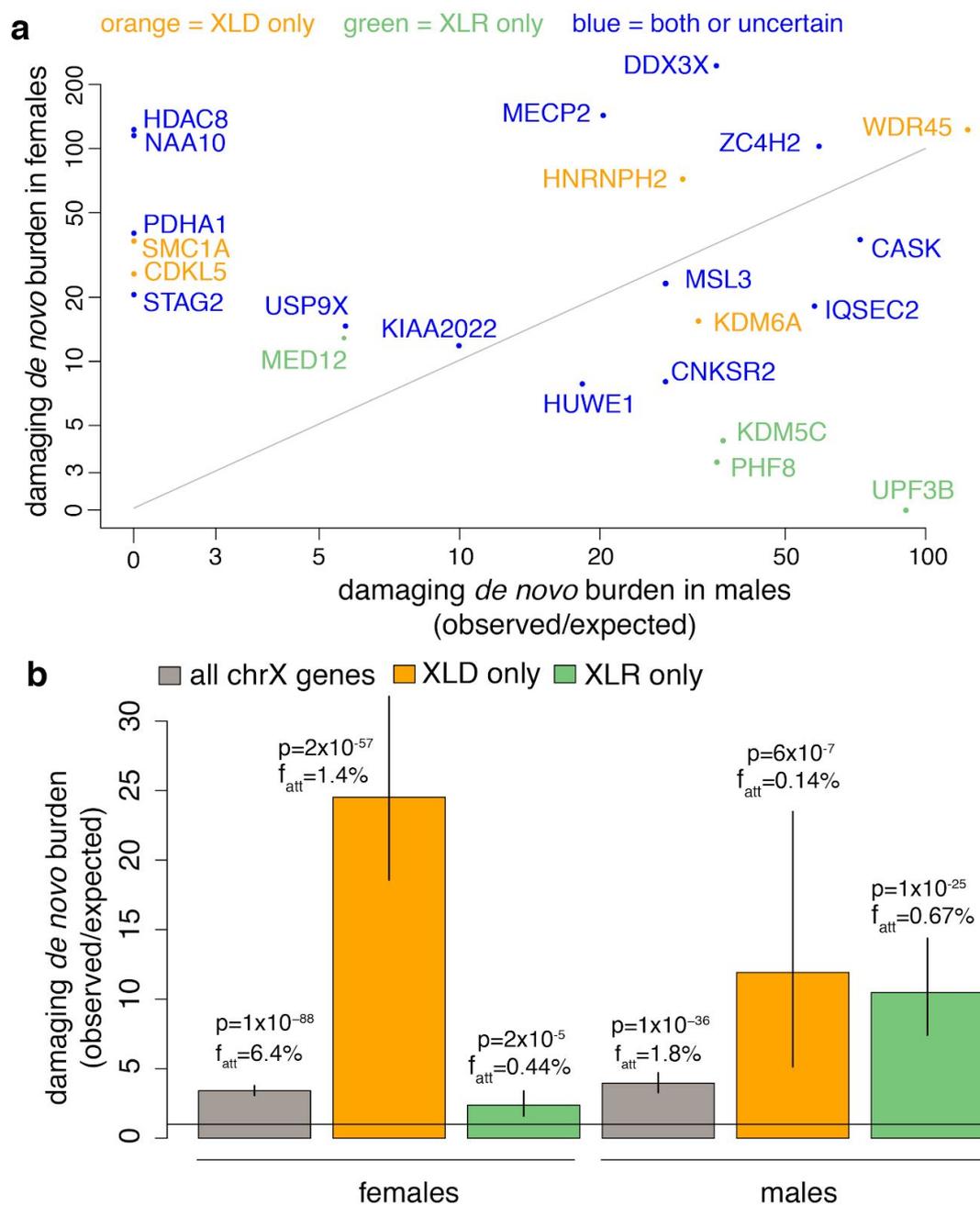
uncorrected values: 1.31 for carriers, 1.76 for non-carriers; ratio t-test $p=0.037$), with a fertility ratio of carriers to noncarriers of 0.742 [0.503-0.981] (Supplementary Fig. 4). Shermer *et al.*²⁵ built on Haldane's theory and determined that the expected fraction of male diagnoses in XLR genes that are *de novo* should be $\frac{\mu(2f+m-mf)}{2\mu+v-vf}$, where f is the reproductive loss in female carriers, m the reproductive loss in affected males (assumed to be 100% here), μ is the mutation rate in eggs and v in the mutation rate in sperm (assumed to be equal to 3.5μ here). Thus, the level of reduced fertility observed in female carriers in UK Biobank implies we should expect 27.3% of the burden in X-linked recessive genes to be *de novo*, and this fraction could be as high as 40.2% considering the lower bound of the fertility ratio. This fraction overlaps the 95% confidence interval we estimate from bootstrapping ([19-83%]; Supplementary Fig. 2D). Hence, reduced numbers of offspring in female carriers may be contributing to this fraction being higher than expected under Haldane's theory.

Gene discovery and delineation of inheritance mechanisms

Our previous efforts at gene discovery on the X chromosome involved testing for *de novo* enrichment in males and females combined¹⁶. This is best powered to detect XLD genes but ignores the substantial contribution from inherited variants in males, so will be underpowered to find new XLR genes. In contrast, many previous gene discovery studies focused on males from obviously X-linked pedigrees so will have missed *de novo* causes of X-linked disorders¹. Hence, we implemented three different tests to optimise power to detect XLR, XLD, and X-linked semi-dominant (male lethal) genes (see Methods). This identified 23 genes that passed Bonferroni correction. These genes were all already known to be DD-associated, reflecting the fact that, in our burden analysis, 78% of the excess was in these known genes, meaning that only an additional ~1.4% of DDD probands (~157 probands) have a diagnostic variant in an X-linked gene not currently associated with DDs. Of these 23 genes, 19 passed Bonferroni correction in the combined analysis of both sexes (our old method), one (*STAG2*) was only significant in the female-only test, and three were only significant when incorporating both *de novo* and inherited variants in males using the Transmission and *De novo* Association Test (TADA)²⁶ (Supplementary Table 4).

We observed that a subset of genes were significantly enriched for DNMs in females only (e.g. *HDAC8*, *NAA10*, *PDHA1*, *SMC1A*, *CDKL5*, *STAG2*), a subset only in males (*UPF3B*, *KDM5C*), and some in both sexes (*IQSEC2*, *CASK*, *WDR45*) (Fig. 2A). The patterns of enrichment we observed were largely consistent with the inheritance modes previously reported for these genes. In principle, this kind of large-scale data analysis should allow us to explore modes of X-linked inheritance in a less biased way than previous small-scale case reports in the literature. We can see from burden analysis (Fig. 2B; positive predictive values shown in Supplementary Fig. 5) that genes annotated as XLD versus XLR clearly have different patterns of DNM enrichment, indicating that there is meaningful heterogeneity among X-linked genes. However, the fact that we still see enrichment of DNMs in nominally XLR genes in females indicates that these classifications in the literature are not perfect.

Figure 2: Sex-specific *de novo* burden analysis. (a) Burden of damaging DNMs (PTV + missense/inframe) in females versus males, per gene. Shown are the 23 X-linked genes that passed multiple-testing correction. The text colour indicates whether the gene was classed in the consensus of DDG2P and OMIM (see Methods) as X-linked dominant only (orange), X-linked recessive only (green) or both/uncertain (blue). P-values for the genes under different tests are shown in Supplementary Table 4. (b) Burden of damaging (PTV + missense/inframe) DNMs for males and females in the indicated gene sets. p : p-value from upper-tailed Poisson test. f_{att} : attributable fraction for DNMs in this gene set. Error bars show 95% confidence intervals calculated as described in the Methods.



MED12 presents a good illustration of the challenges in trying to classify inheritance modes for X-linked genes. It had been previously reported to cause X-linked recessive FG, Lujan and Ohdo syndromes as well as non-syndromic intellectual disability²⁷⁻²⁹, and DDG2P and OMIM class it as XLR. However, heterozygous females have previously been reported with mild and, in some cases, severe phenotypes, with the severity not obviously being correlated with the degree of skewed X-inactivation³⁰⁻³³. We observed eight damaging DNMs in females and one in a male. The phenotypes of our patients and those reported for other *MED12* patients in the literature are largely consistent. For five of the eight female *MED12* patients, the clinician reported that the contribution of the *MED12* mutation was ‘uncertain’, presumably partly because it is reported to be recessive in OMIM. Our results suggest that *MED12* should likely be reclassified as causing both X-linked dominant and X-linked recessive conditions, and be considered to cause a spectrum of related disorders rather than distinct syndromes. Further work on the functional consequences of different *MED12* mutations and the degree of X-inactivation in female brain tissue would be required to understand why some mutations appear to be recessive and others dominant.

The *MED12* example above illustrates the power of these kinds of large-scale data analyses to identify patterns of sex-biased DNMs that are inconsistent with current classification of inheritance mode, but ideally we would like to be able to assign all genes to modes of inheritance with high confidence. Genes that exhibit a substantial female bias in observed DNMs (Fig. 2B) could be semi-dominant and lethal in males, or dominant (cause equivalent disease in males and heterozygous females). The low mutability of the maternally transmitted X chromosome in males results in a low expected number of DNMs in genes pathogenic in males. It may well be that we simply have not observed any males with DNMs in those genes by chance due to limited sample size. To distinguish these possibilities, we need to model the expected number of DNMs in females versus males, given the null mutation model, coverage, ploidy and sample size (see Methods). Under this model we would expect ~78% of DNMs on the X chromosome to occur in females, although they only make up 43% of the trios used for the DNM enrichment tests. When we accounted for this, we identified two genes with a nominally significant female bias (Supplementary Table 5), *MECP2* (binomial $p=0.02$; 24/25=96% of DNMs observed in females) and *DDX3X* (binomial $p=0.002$; 48/50=96% of DNMs observed in females). This provides statistical evidence that both of these genes exhibit a semi-dominant mode of inheritance. These two genes were recently reported to show significant female bias for DNMs³⁴, although that work did not correct for the different mutation rates in males and females and hence overestimated the significance of the female bias. *MECP2* is already known to cause female-limited Rett syndrome, and males with verified PTV mutations have not been observed, although missense mutations have been reported in males with severe epileptic encephalopathy. In line with this, in both *MECP2* and *DDX3X*, the only DNMs in males in our cohort were missense or inframe. *MECP2* and *DDX3X* are among the genes with the most DNMs in our cohort, and therefore have the most power for assigning inheritance mode. Confident assignment of all X-linked genes to one or other inheritance mode would require a larger sample size than is available in this study.

Role of polygenic background

Clinicians recruiting patients to the DDD study were asked to indicate at recruitment whether they suspected the patient may have an X-linked cause. Male patients with suspected X-linked inheritance (N=271) were enriched 2.7-fold for higher inherited X-linked coding burden (attributable fraction = 16.2% [6.3%-27.1%] versus 6.1% [3.6%-8.6%] for all males) (binomial $p=2\times 10^{-9}$). Based on recent work showing that common variants also contribute to risk of rare developmental disorders³⁵, we hypothesised that polygenic background could be contributing to the presence of multiple affected

males in families, leading clinicians to incorrectly suspect X-linked inheritance. Using imputed genotype chip data on a subset of the cohort, we tested for a difference in polygenic scores for relevant traits between 216 males suspected to have X-linked inheritance versus 3439 who were not, having excluded those with a potentially diagnostic X-linked variant. Specifically, we assessed polygenic scores for educational attainment³⁶, intelligence³⁷, schizophrenia³⁸, and severe neurodevelopmental disorders (NDD) derived from our own GWAS³⁵. None were significantly different after correcting for four tests (linear regression correcting for 10 genetic principal components; $p > 0.05/4$; Supplementary Table 6). This analysis will require more powerful polygenic scores and a larger sample size to clarify the contribution of polygenic background to the clustering of affected individuals in families.

Discussion

Here we analysed the burden of *de novo* and inherited rare coding variants on the X chromosome and estimated that these explain ~6% of both male and female probands in 11,046 families in the DDD study. We found that about three-quarters of this burden was in known DD genes. These proportions are similar to findings in other exome-sequencing studies of similar cohorts³⁹⁻⁴¹. For example, in one of the largest comparable studies, 34/938 (3.6% [2.5%-5.0%]) male probands and 29/728 (4.0% [2.7%-5.7%]) female probands with neurological disorders had a diagnostic X-linked variant⁴¹, versus 3.8% and 5.4% respectively in DDD (inferred from burden analysis in known DD-associated genes). Importantly, our results show that, when rare monogenic causes in as-yet-undiscovered X-linked genes are also accounted for, they still cannot explain the male bias within our cohort. There is a common perception amongst clinicians that X-linked causes of DDs are much less likely in females than males (except for e.g. Rett Syndrome due to *MECP2* mutations)⁴², but our burden analyses do not support this. Our results imply that the discovery of the remaining X-linked DD genes may allow us to diagnose another ~1.4% of our cohort. In contrast, we recently estimated that discovery of the remaining autosomal dominant DD genes would allow us to diagnose about another ~23% of the cohort with pathogenic DNMs⁴³. This likely reflects the fact that it was easier to find X-linked DD genes than autosomal dominant DD genes in the linkage era.

We found that ~41% of the burden on the X chromosome in trio males was *de novo*, and this fraction was ~36% in XLR genes, higher than the ~18% expected under Haldane's theory. To our knowledge, this number has previously only been estimated for individual genes, and only based on likely diagnostic variants rather than in an unbiased burden analysis (e.g.^{44,45}). It is important to emphasise that the fractions we have estimated in DDD are likely higher than they would be in an unbiased sample of DD patients. DDD may be biased away from inherited X-linked causes due to our ascertainment strategy, which excluded patients who already had a genetic diagnosis. X-linked inherited causes may be easier to diagnose, since an X-linked inheritance pattern in a family reduces the search space. Additionally, the earlier GOLD study had already recruited several hundred of these families in the UK, so they may have been under-recruited to DDD. Reduced reproduction in heterozygous carrier females might also be contributing to the higher-than-expected contribution of DNMs in X-linked recessive genes; this could include increased pre-reproductive mortality in females with skewed X-inactivation as well as phenotypically normal mothers choosing not to have more children after having an affected son. We observed ~26% lower reproductive success (fertility ratio 0.742 [0.503-0.981]) in a small sample (N=13) of carrier females in UK Biobank compared to non-carriers. Larger sizes of relatively unbiased population samples are needed to confirm this apparently decreased reproductive success, and to quantify its influence on the ratio of *de novo* to inherited pathogenic variants in XLR genes.

Our results have important diagnostic implications. We demonstrated that, while the vast majority of observed *de novo* and inherited PTVs in males in known X-linked DD genes are pathogenic, inherited missense variants in these genes have only a low PPV (Fig. 1C; Supplementary Fig. 4). This implies that it is challenging to accurately diagnose males with rare inherited missense variants in X-linked DD genes, particularly since the high recurrence risk for such variants presents a legitimate concern and thus may increase motivation for clinicians to take action. Additional simple manoeuvres such as testing the apparently unaffected males in the family for an inherited missense variant identified in an index male can be very helpful in excluding causality but are rarely deployed outside of clinical genetics services. On the other hand, our comparison with ClinVar (Fig. 1B) suggests that, in fact, current clinical interpretation is likely highly conservative and missing a large fraction of pathogenic missense variants in X-linked DD genes in males, since PTVs are relatively enriched in the ClinVar likely pathogenic/pathogenic variants compared to our burden analysis. Indeed, assuming that all diagnostic PTVs are being identified, we estimate from this analysis that ~58% of diagnostic missense/inframe variants in known X-linked DDG2P genes are not being classed as pathogenic. Incorporating CADD and MPC scores during variant interpretation may improve specificity in clinical diagnosis, but reduce sensitivity (Fig. 1C).

We developed a new strategy for finding X-linked Mendelian disease genes that considers several different inheritance modes and incorporates both *de novo* and inherited variation in males, incorporating and building on the TADA method²⁶. We showed that this strategy identified 21% (23 versus 19) more X-linked DD-associated genes at genome-wide significance than our previous approach. However, all genes that passed genome-wide significance were already known, reflecting low power which is due to the fact that only ~1.4% of the cohort has an X-linked diagnosis in an as-yet-undiscovered gene. Although our approach was designed to distinguish XLD from XLR genes, it is clear that more data will be needed from larger cohorts for a data-based classification of inheritance modes. Larger sample sizes may also reveal more X-linked genes with a sex bias other than just *DDX3X* and *MECP2*³⁴ which appear to be semi-dominant; we emphasise that, in testing for these, it will be important to account for the differential rates of DNMs on the X chromosome in the two sexes, otherwise the degree of female bias will be over-estimated (see Methods).

Our study has several limitations. Firstly, there are wide confidence intervals around several of the key parameters we have estimated (attributable fraction, fraction of *de novo* versus inherited causes, positive predictive value etc.), despite our relatively large sample size. An even larger sample size would increase the precision of these estimates. Secondly, there may be ascertainment bias in the DDD study away from very recognisable disorders and families with a clear XLR inheritance pattern. This means our estimates of important parameters will not necessarily hold for other cohorts. Large cohorts in which exome-sequencing has been applied as a first-line test will allow less biased estimates. Finally, our attempts at new gene discovery and classification of inheritance modes were limited not only by sample size, but also by lack of data on X-inactivation in females and the phenotypes of carrier mothers. Furthermore, we anticipate that extra data will reinforce that many X-linked genes can show both dominant and recessive inheritance depending on the severity of the variant, including, for some genes, different phenotypic features associated with hemizygous versus heterozygous variants (e.g. *NAA10*⁴⁶). Hence, future gene discovery efforts will need to develop unbiased ways of discriminating between different disorders caused by mutations in the same gene on the basis of phenotype or functional evidence.

It is notable that within our cohort, less than a quarter of males who were suspected by clinicians to have an X-linked condition were inferred to have a pathogenic rare or *de novo* X-linked coding

variant in any gene. This may imply that there are other factors contributing to the recurrence of a DD in multiple males from the same family. Under the hypothesis that males have a lower liability threshold than females for NDDs, it seems plausible that an enrichment of multiple deleterious variants across the frequency spectrum might push multiple males but not females in a family over this threshold, creating the appearance of X-linked Mendelian inheritance. We did not see a difference in polygenic scores (with MAF>5% variants) for relevant traits between males without likely diagnostic X-linked variants who were suspected to have an X-linked disorder versus those who were not. However, given that existing polygenic scores for intelligence explain only ~4-5% of variance in IQ in out-of-sample prediction⁴⁷⁻⁴⁹, we anticipate that we may see a difference with a larger sample size and more informative polygenic scores, potentially including rarer variants that have been shown to explain a substantial proportion of variance in intelligence⁵⁰.

In conclusion, our work shows that monogenic causes on the X chromosome cannot account for the male bias in developmental disorders. Analyses of variants across the full frequency spectrum in large cohorts may reveal a contribution of more common variants to the sex bias. This work provides a robust statistical framework for analyses of the X chromosome in large Mendelian disease cohorts, which will aid in future gene discovery and inform improvements in clinical practice.

Methods

Family recruitment

Individuals with severe, undiagnosed developmental disorders were recruited to the DDD project by 24 clinical genetics centres within the United Kingdom National Health Service and the Republic of Ireland as described previously⁵¹. They had to have at least one of the following phenotypes:

1. Neurodevelopmental disorder – for example developmental delay and/or learning disability (of a level requiring or likely to require a statement of special educational needs), epileptic encephalopathy or cerebral palsy
2. Congenital anomalies – multiple congenital anomalies (two or more major anomalies) or a single major anomaly together with a neurodevelopmental disorder, aberrant growth, dysmorphic features or unusual behaviour
3. Abnormal growth parameters (height, weight, head circumference (OFC)) – two or more parameters >3SD above or below the mean or a single parameter >4SD above or below the mean (except for obesity where the threshold for isolated obesity is >4.5SD together with a strong suspicion of a genetic aetiology)
4. Unusual behavioural phenotype in conjunction with one or more of the above features or extreme behavioural phenotype strongly suspected to have a genetic basis (including classical autism)
5. Genetic disorder of significant impact for which the molecular basis is currently unknown with: (i) several affected family members or (ii) one other affected family member with a rare, consistent and distinctive phenotype or (iii) a single case that is associated with a severe phenotype.

Families gave informed consent to participate, and the study was approved by the UK Research Ethics Committee (10/H0305/83, granted by the Cambridge South Research Ethics Committee and GEN/284/12, granted by the Republic of Ireland Research Ethics Committee). As described previously, DNA was collected from saliva samples obtained from the probands and their parents, and from blood obtained from the probands⁵². The individuals analysed in this paper include those analysed in previous publications^{16,43,51,53,54}.

Clinical features

As described previously⁵⁴, the patients were systematically phenotyped using Human Phenotype Ontology (HPO) terms⁵⁵, and growth measurements, developmental milestones, family history (including whether X-linked inheritance was suspected) etc. were collected within DECIPHER⁵⁶. For the summary of phenotypes and comparison between the sexes (Supplementary Fig. 1, Supplementary Table 1), we followed the procedure in⁵⁴ when counting organ systems affected, to avoid double-counting HPO terms that fall under multiple organ systems. For the comparison of age at walking and talking in Supplementary Table 2, we excluded probands who had not yet reached these milestones; the sample sizes are shown in the table. We used linear regression correcting for age at assessment to compare quantitative phenotypes between sexes, and logistic regression correcting for age at assessment to compare the frequency of binary phenotypes between sexes.

Exome sequencing, variant annotation and variant quality control

Exome sequencing, alignment and calling of single-nucleotide variants and small insertions and deletions was carried out as previously described¹⁶. Variants were annotated with Ensembl Variant Effect Predictor⁵⁷ based on Ensembl gene build 88, using the LOFTEE plugin. We analyzed three categories of variant based on the predicted consequence: (1) protein-truncating variants (PTVs) classed as “high confidence” loss-of-function variants by LOFTEE (including the annotations splice donor, splice acceptor, stop gained, frameshift, initiator codon and conserved exon terminus variant); (2) missense variants and inframe indels; (3) synonymous variants. We assigned each variant the worst consequence across all the transcripts for a gene. Missense variants were annotated with CADD v1.3²² and MPC²³ scores. All variants were annotated with MAF data from four different populations of the 1000 Genomes Project⁵⁸ (American, Asian, African and European), two populations from the NHLBI GO Exome Sequencing Project (European Americans and African Americans) and six populations from the Genome Aggregation Consortium (gnomAD) release 2.0.2 (African, East Asian, non-Finnish European, Finnish, South Asian, Latino), and internal allele frequencies from the European and South Asian unaffected DDD parents.

For the case-control analysis of chrX in males, we used the following filters:

- Genotypes were set to missing if they had genotype quality (GQ)<20, depth (DP)<7, or were heterozygous.
- Variants were removed if they met any of the following criteria:
 - were in the pseudoautosomal regions (chrX:60001-2699520 and chrX:154931044-155260560 in GCh37)
 - had a strand bias p-value < 0.001
 - had >50% missing calls (after the genotype-level filtering) within the samples that underwent Agilent SureSelect Human All Exon V5 capture or within those that underwent Agilent SureSelect Human All Exon V3 capture
 - had minor allele frequency (MAF)>0.001 in any gnomAD population or in the unaffected European or South Asian parents from DDD. For calculating the PPV for lower MAF variants (MAF<0.0005, 0.0001, 0.00005), we only considered the gnomAD POPMAX and the frequency in European DDD parents, since the set of South Asian DDD parents in DDD was too small.
 - had any hemizygotes in gnomAD

For calling DNMs on chrX, we ran DeNovoGear⁵⁹ as previously described¹⁶, but with a different set of hard filters to account for the lower coverage in males and to maximise sensitivity and specificity.

We examined all candidate DNMs in males and a large subset of those in females manually in IGV, and used this to settle on the following set of filters:

- We removed DNMs in the pseudoautosomal regions.
- The variant had to be called heterozygous or (for males) hemizygous in the child in the original GATK calls, and called homozygous reference in the parents.
- For male probands, we required the following: in the child, alternate allele depth >2 and overall depth > 2; in the mother, depth > 5.
- For female probands, we required the following: in the child, alternate allele depth > 2, overall depth > 7; in the mother, depth > 5; in the father, depth > 1.
- For single nucleotide variants, we required $p > 10^{-3}$ on a Fisher's exact test for strand bias, pooling across trios (or mother-child pairs, for male probands) where a DNM was called at the same site by DeNovoGear.
- For female probands, we removed indels <5bp if they had variant allele fraction <0.3 or MAF > 0, since these were vastly over-represented and seemed to be a common error mode.
- We did a two-sided binomial test on the number of alternate reads at the candidate site in mothers, assuming the proportion of these should be 0.5 if heterozygous, and then discarded sites where the p-value from this test (called p_{het}) was > 0.01, since these indicated that the mother was likely to be truly heterozygous and not mosaic.
- We did a binomial test to evaluate whether the fraction of alternate reads was greater than the expected error rate of 0.2% (p_{error}), and then flagged variants as mosaic if they had lower-tailed $p_{\text{het}} < 0.01$ and upper-tailed $p_{\text{error}} < 0.01$. We then removed variants in segmental duplications if both the child and at least one parent (mother for female probands) were flagged as mosaic.
- We set a cutoff for the posterior probability of being a DNM from DeNovoGear to $pp_{\text{DNM}} > 0.00247679$. This value was chosen because the observed number of synonymous DNMs in females that passed this pp_{DNM} cutoff as well as the hard filters above was very close to the expected number calculated using null mutation rate determined as described below. We chose to calibrate the pp_{DNM} threshold in females since the numbers in males were very small.

We subsequently removed DNMs that had MAF > 0.001 in any gnomAD population or in the unaffected European or South Asian parents from DDD, or that had any hemizygotes in gnomAD.

Calculation of expected mutation rates on chrX

We estimated the expected number of DNMs per gene in different functional classes using the method in ¹⁴. We adjusted for the reduced sensitivity to detect DNMs due to limited coverage following the method in ⁶⁰, with some minor adaptations. For this, we first calculated the median depth per exon in 250 samples on the Agilent V5 capture, and then took the mean of these. To determine the depth-uncorrected expected number of variants per exon, we took the exons with mean median depth ≥ 30 and regressed the number of rare (MAF < 0.001) synonymous variants on the probability of a synonymous mutation. For males and females separately, we plotted the ratio of observed to depth-uncorrected expected synonymous variants against the depth in bins of 2X (for up to 40X in males and 80X in females) and fitted a logarithmic curve. We then used this formula to predict the depth-corrected expected number of variants for all exons:

$$\text{depth-adjusted expected count} = \begin{cases} 0 & \text{if depth} < 1 \\ \text{expected count} \times (0.2778 \ln(\text{depth}) + 0.0279) & \text{if } 1 < \text{depth} < 30 \text{ in males} \\ \text{expected count} \times (0.2464 \ln(\text{depth}) + 0.035) & \text{if } 1 < \text{depth} < 50 \text{ in females} \\ \text{expected count} & \text{if depth} > 30 \text{ in males or } > 50 \text{ in females} \end{cases}$$

In calculating the expected counts of DNMs in the non-pseudoautosomal regions of chrX in males and females, we followed the method previously described in ⁵² to account for the different inheritance pattern of this chromosome and the different mutation rates in the male and female germline. Specifically, we determined the scaling factors as follows:

$$f_{females} = t_{F>F} + t_{M>F} = n_{female}\lambda_{female} + n_{female}\lambda_{male}$$
$$f_{males} = t_{F>M} = n_{male}\lambda_{female}$$

where $t_{F>F}$ and $t_{M>F}$ are the number of transmissions from females and males to female probands respectively, $t_{F>M}$ is the number of transmissions from females to male probands, n_{female} and n_{male} are the numbers of female and male probands, and λ_{female} and λ_{male} are adjustment factors to account for the higher mutation rate in males, given by:

$$\lambda_{female} = \frac{2}{1+\alpha} \text{ and } \lambda_{male} = \frac{2}{1+\frac{1}{\alpha}}$$

where $\alpha = 3.4$ is the ratio of the mutation rate in fathers to mothers in DDD, determined using 199 phased DNMs ⁵². These scaling factors $f_{females}$ and f_{males} were multiplied by the depth-adjusted mutation rates to determine the expected number of DNMs as $n_{female}\mu_{female}$ and $n_{male}\mu_{male}$.

Sample Quality Control

We removed probands with sex chromosome aneuploidies (N=47) or whose chromosomal sex did not match the sex recorded by the clinician (N=49). We also removed six probands with an implausibly high number of *de novo* calls, likely to be spurious, and 330 samples with >20% missing genotypes on chrX after the genotype QC described above. We used KING⁶¹ to estimate relatedness between individuals, applying it to variants with MAF>0.01 with <5% missingness after genotype QC. Then, for unaffected fathers (who were used as controls) and for probands separately, we removed one person from each pair of individuals inferred to be third-degree relatives or closer (doing this in such a way as to minimise the number of individuals removed); this led to the removal of 586 male probands, 211 female probands and 108 unaffected fathers. In total, 1,311 unique individuals were removed. The final analysis after QC was conducted on 7,138 male probands (5,138 in trios), 3,908 female probands in trios, and 8,551 unaffected fathers.

Burden analysis

We conducted sex-specific burden analyses to test for an enrichment of certain classes of X-linked variants in probands and to estimate the fraction of probands attributable to such variants.

1. In males, we conducted a case/control analysis comparing male probands to the unaffected fathers. This incorporates both maternally inherited variants and DNMs that passed the genotype and variant filtering for the case/control analysis (but did not necessarily pass the filtering of mutations called by DeNovoGear). In the case-control analysis, we calculated the rate of a particular class of variant per person in the unaffected fathers and used this to calculate the expected number of variants in the male probands by simply multiplying this rate by the number of probands.
2. In both females and males, we conducted a DNMs enrichment analysis comparing the observed number of DNMs to the expected number, as calculated above.

For both burden analyses, we restricted to variants with maximum MAF<0.001 and with no hemizygotes in gnomAD, to try to restrict to the most damaging subset of variants. We also examined synonymous, PTV and missense/inframe variants (the latter filtered in various ways) separately. Synonymous variants were used as a control: we confirmed that the number of observed synonymous variants was not significantly different from expectation, to ensure that the test was well-calibrated.

We tested for enrichment assuming a Poisson distribution using an upper-tailed test. The metrics reported in the main text were calculated as follows:

$$\begin{aligned} \text{burden} &= \frac{\# \text{ observed variants}}{\# \text{ expected variants}} \\ \text{attributable fraction} &= \frac{\text{excess}}{\# \text{ probands}} = \frac{\# \text{ observed variants} - \# \text{ expected variants}}{\# \text{ probands}} \\ \text{positive predictive value (PPV)} &= \frac{\text{excess}}{\# \text{ observed variants}} = \frac{\# \text{ observed variants} - \# \text{ expected variants}}{\# \text{ observed variants}} \end{aligned}$$

For DNMs, we treated the expected rate as a fixed quantity, calculated 95% confidence intervals on the number of expected DNMs using the `poisson.test()` function in R, and then substituted these into the above formulae to calculate confidence intervals on those metrics. For the case/control analysis in males, the expected rate of variants is calculated based on the number observed in fathers, so is a random variable. Thus, we used the `moverci()` function in R to calculate confidence intervals on the difference or ratio of two Poisson rates, and substituted these back into the above formula as appropriate. Specifically, the confidence interval for each metric was calculated as follows:

- burden: `moverci(Vprobands, Nprobands, Vfathers, Nfathers, distrib = "poi", contrast="RR")` (i.e. a rate ratio) where N is the number of individuals and V is the number of observed variants.
- attributable fraction: `moverci(Vprobands, Nprobands, Vfathers, Nfathers, distrib = "poi", contrast="RD")` (i.e. a difference in rates)
- PPV: `1-moverci(Vfathers, Nfathers, Vprobands, Nprobands, distrib = "poi", contrast="RR")`

To calculate a confidence interval on the fraction of pathogenic variants that were in known genes and the fraction of pathogenic variants in XLR genes in male trio probands that was *de novo*, we bootstrapped probands. Specifically:

- For females, we bootstrapped probands 1000 times, each time recalculating the excess number ($\text{excess} = \# \text{ observed} - \# \text{ expected}$) of DNMs in all genes and in known X-linked DD genes, then determined the 2.5th and 97.5th percentile of the ratio $\frac{\text{excess}_{\text{DNM, known genes}}}{\text{excess}_{\text{DNM, all genes}}}$ across the 1000 iterations (Supplementary Fig. 2B).
- For males, we bootstrapped probands and fathers each separately 1000 times, each time recalculating the excess number of variants in probands in all genes and in known X-linked DD genes, then determined the 2.5th and 97.5th percentile of the ratio $\frac{\text{excess}_{\text{vars, known genes}}}{\text{excess}_{\text{vars, all genes}}}$ across the 1000 iterations (Supplementary Fig. 2A). We repeated this procedure with just the male trio probands, and used the same 1000 sets of bootstrapped trios to also calculate $\text{excess}_{\text{vars, all genes}}$, $\text{excess}_{\text{vars, XLR genes}}$, $\text{excess}_{\text{DNM, all genes}}$ and $\text{excess}_{\text{DNM, XLR genes}}$, then determined the 5th and 95th percentile of the ratios $\frac{\text{excess}_{\text{DNM, all genes}}}{\text{excess}_{\text{vars, all genes}}}$ (Supplementary Fig. 2C) and $\frac{\text{excess}_{\text{DNM, known genes}}}{\text{excess}_{\text{vars, known genes}}}$ (Supplementary Fig. 2D). In Supplementary Figures 2A and 2C, the 97.5th percentile of the ratio was greater than 1, so in the text we report the upper bound of those confidence intervals to be 100%.

Per-gene enrichment tests

For each gene, we tested for a significant burden of DNMs using a Poisson test, as previously described⁵². For males and females combined (the old test) and for females alone, we tested PTVs alone and PTVs and missense/inframe variants combined, then took the lowest p-value for each proband set.

We also applied the Transmission and De Novo Association test (TADA)²⁶ to each gene in males to test for enrichment of *de novo* and inherited variants combined. The inherited counts were determined from the male probands and their fathers respectively. We removed from these counts the *de novos* that passed our DeNovoGear filtering, since these were counted already in the TADA *de novo* mutation model; we also tried removing from the inherited counts variants that did not pass the DeNovoGear filtering but which had 0 alternate reads in mothers, and counting these as DNMs, but in practice this made little difference to the results. Since there are ~15% of genes on chrX already implicated in DDs (in the DDG2P list) but it was unclear how the case ascertainment in DDD might have created biases against these, we tried varying π , the prior on the fraction of risk genes, from 0.05 to 0.25. Other prior parameters were calculated accordingly, following the procedure described in the TADA user guide (http://www.compgen.pitt.edu/TADA/TADA_guide.html). These are shown in Supplementary Table 7. For all runs, as parameters in the prior for the allele frequencies, we set $\nu = 100$ and $\rho = 0.618$ for PTVs, and $\nu = 100$ and $\rho = 11.729$ for missense/inframe variants. We ran TADA separately on PTVs alone and on PTVs and missense/inframe variants combined.

We calculated q-values using the Bayesian.FDR() function in TADA, and p-values using a sampling approach via the TADAnull() and bayesFactor.pvalue() functions.

Under the old testing strategy (PTVs alone and PTVs+missense/inframe variants combined for males and females combined), we corrected for $2 \times 19,685$ genes, giving a genome-wide significance threshold of $p < 0.05 / (2 \times 19,685) = 1.27 \times 10^{-6}$. Under our new testing strategy, overall we applied six tests to each of 804 X-linked genes: (A) PTVs alone and (B) PTVs+missense/inframe variants combined for each of (1) DNMs in females alone (Poisson), (2) DNMs in males and females combined (Poisson), and (3) DNMs and inherited variants in males (TADA). Since we would typically apply two separate DNM enrichment tests to each autosomal gene (PTVs alone and PTVs+missense/inframe variants), in total, we corrected for $(6 \times 804 + 2 \times (19,685 - 804))$ tests, giving a genome-wide significance threshold of $p < 0.05 / 40780 = 1.17 \times 10^{-6}$.

Testing for sex bias in DNMs per gene

For each gene, we calculated the fraction of expected PTV+missense/inframe DNMs that should be in males as $n_{male} \mu_{male} / (n_{male} \mu_{male} + n_{female} \mu_{female})$, using the mutation rates calculated as described above accounting for coverage and ploidy. We then compared the fraction of observed DNMs that were in males to this expected fraction using a lower-tailed binomial test (under the hypothesis that the gene would be depleted for DNMs in males due to lethality).

Gene list definitions

To define the list of known X-linked DDG2P genes, we took the intersection of confirmed or probable DDG2P genes on the X chromosome and OMIM genes with a disease annotation. To define “X-linked recessive” (called ‘hemizygous’ in DDG2P) and “X-linked dominant” genes, we compared the inheritance annotations between DDG2P and OMIM and took the consensus. Hence, “X-linked recessive” genes were those annotated *only* as hemizygous in DDG2P and *only* as X-linked recessive in OMIM, and similarly for X-linked dominant genes. There were 12 genes classified as exclusively X-linked dominant and 63 as exclusively X-linked recessive. Genes annotated as both X-linked dominant and X-linked recessive, or annotated simply as “X-linked” in either DDG2P and OMIM, have been coloured in blue in Fig. 2, and excluded from analysis of X-linked recessive genes and from Supplementary Figures 2D, 4 and 5.

Investigating the *de novo* versus inherited contribution in X-linked recessive genes

Under Haldane's theory, the fraction of male X-linked recessive cases due to DNMs, $p_{de\ novo}$, should be $\frac{m\mu}{2\mu+v}$, where m is the reproductive loss in affected males, μ is the mutation rate in eggs and v in the mutation rate in sperm. If we assume $m = 1$ and that $v = 3.5\mu$ (based on a previous estimate⁴), we should expect $p_{de\ novo}$ to be about $\frac{1}{5.5} \approx 18.2\%$.

We focused on X-linked recessive genes from DDG2P that were not also annotated as X-linked dominant, and included only PTV and missense SNVs because the male to female mutation rate ratio was calculated for SNVs. We estimated (based on the excess=observed-expected) that 68.5 [31.2-102.9] male trio cases could be attributed to pathogenic variants in these genes, of which 21.8 [15.5-24.8] were *de novo*. Assuming the data approximate a binomial distribution, we tested the consistency with Haldane's theory using a two-sample test for equality of proportions; specifically, we used `prop.test(22,69,p=1/5.5)` in R. We also assessed consistency with Haldane's theory through bootstrapping as described in the "Burden analysis" section above.

Sherman *et al.* modified Haldane's theory to account for reduced reproductive fitness in carrier females²⁵. Under their theory, $p_{de\ novo}$ should be $\frac{\mu(2f+m-mf)}{2\mu+v-vf}$, where f is the reproductive loss in carrier females. We again assume $m = 1$ and $v = 3.5\mu$. Any value of $p_{de\ novo}$ between 0.22 and 0.44 should be consistent ($p\text{-value} > 0.05$) with the excess values we observed. This corresponds to a value for f between ~12% and ~56%. We next turned to the UK Biobank exome data to estimate f directly.

UK Biobank analysis

To assess the fitness consequences of heterozygous variants in X-linked recessive genes in females, we turned to UK Biobank (UKBB). We downloaded variant calls for all 49,959 samples subjected to WES as part of the UKBB study⁶². We next annotated all variants with Variant Effect Predictor (v97)⁵⁷, extracted variants in exclusively X-linked recessive genes, and retained only the most severe consequence in canonical transcripts. We note that none of these X-linked recessive genes are affected by the recently-reported problem with this UKBB exome data release which is related to mapping errors⁶³.

We then removed PTV and missense variants with $CADD \leq 25$ ⁶⁴, missense variants with an MPC²³ score ≤ 2 , low-confidence PTV variants using the LOFTEE plugin for VEP⁶⁵, and all variants with a gnomAD²¹ non-Finish European minimum allele frequency ≥ 0.001 . Only UKBB variants identified in a single individual were retained for downstream analysis. To further ensure that we did not include any deleterious variants of potentially incomplete penetrance, we also filtered any PTV and missense variants which were also found in any male individuals in UKBB. We then removed all related, non-white British ancestry individuals and all males, leaving a final total of 18,632 female samples. PTV, missense, and synonymous variants passing the above criteria were then counted for each remaining individual (Supplementary Fig. 4A). All 13 likely PTV variants in X-linked recessive genes we identified were confirmed via manual inspection using the Integrative Genomics Viewer⁶⁶.

To determine the total number of live births for each female in UKBB, we accessed and downloaded field 2734. To determine fluid intelligence scores, we accessed and downloaded field 20016. Only data obtained via in-centre testing were retained for further analysis. Age and pre-computed ancestry PCs⁶⁷ were obtained from UKBB fields 21022 and 22009, respectively. To independently determine the effect of PTV, missense, or synonymous variants on each phenotype, we used a simple linear model (via the `glm` function in R with family “gaussian”) in the form:

$$phenotype \sim N_{VAR} + age + age^2 + PC1 + \dots + PC10$$

where *phenotype* is either number of live births or normalised fluid intelligence, and N_{VAR} is one of total PTV, missense, or synonymous variants in each individual in all confirmed X-linked recessive DD genes (Supplementary Fig. 4B, C). To determine the ratio of number of live births between X-linked recessive PTV carrier and non-carrier females, corrected for age and ancestry PCs, we used the function `ttestratio` from the R package `mratios` v1.4.0 with default settings.

Polygenic score analysis

We restricted the analysis to 4,168 male probands with a neurodevelopmental disorder who had been genotyped on the CoreExome array, had European ancestry and passed our quality control in³⁵. We excluded 513 males who had an X-linked variant in a DDG2P gene reported to DECIPHER that had not yet been classed as ‘benign’ or ‘likely benign’ by clinicians. Polygenic scores for educational attainment³⁶, intelligence³⁷ and schizophrenia³⁸ were calculated as previously described (see parameters in Extended Data Table 2 of³⁵). For calculating the NDD polygenic score, we repeated the GWAS in the same dataset as described in³⁵ but without sex as a covariate, then used $p < 1$ and $r^2 < 0.1$ when pruning the SNPs. For the comparisons of polygenic scores between males who were versus were not suspected by clinicians to have an X-linked diagnosis, we ran a linear regression of polygenic score on suspected group status plus 10 genetic principal components.

References

1. Tarpey, P. S. *et al.* A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat. Genet.* **41**, 535–543 (2009).
2. Jónsson, H. *et al.* Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature* **549**, 519–522 (2017).
3. Haldane, J. B. S. The rate of spontaneous mutation of a human gene. *J. Genet.* **31**, 317 (1935).
4. Rahbari, R. *et al.* Timing, rates and spectra of human germline mutation. *Nat. Genet.* **48**, 126–133 (2016).
5. Amos-Landgraf, J. M. *et al.* X Chromosome–Inactivation Patterns of 1,005 Phenotypically Unaffected Females. *Am. J. Hum. Genet.* **79**, 493–499 (2006).
6. Tukiainen, T. *et al.* Landscape of X chromosome inactivation across human tissues.

doi:10.1101/073957.

7. Fieremans, N. *et al.* Identification of Intellectual Disability Genes in Female Patients with a Skewed X-Inactivation Pattern. *Human Mutation* vol. 37 804–811 (2016).
8. Lubs, H. A., Stevenson, R. E. & Schwartz, C. E. Fragile X and X-linked intellectual disability: four decades of discovery. *Am. J. Hum. Genet.* **90**, 579–590 (2012).
9. Dobyns, W. B. *et al.* Inheritance of most X-linked traits is not dominant or recessive, just X-linked. *American Journal of Medical Genetics* vol. 129A 136–143 (2004).
10. Thormann, A. *et al.* Flexible and scalable diagnostic filtering of genomic variants using G2P with Ensembl VEP. *Nat. Commun.* **10**, 2373 (2019).
11. Piton, A., Redin, C. & Mandel, J.-L. XLID-causing mutations and associated genes challenged in light of data from large-scale human exome sequencing. *Am. J. Hum. Genet.* **93**, 368–383 (2013).
12. Piton, A. *et al.* Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia. *Mol. Psychiatry* **16**, 867–880 (2011).
13. Raymond, F. L. X linked mental retardation: a clinical guide. *J. Med. Genet.* **43**, 193–200 (2006).
14. Samocha, K. E. *et al.* A framework for the interpretation of de novo mutation in human disease. *Nat. Genet.* **46**, 944–950 (2014).
15. Sifrim, A. *et al.* Distinct genetic architectures for syndromic and nonsyndromic congenital heart defects identified by exome sequencing. *Nat. Genet.* **48**, 1060–1065 (2016).
16. Deciphering Developmental Disorders Study. Prevalence and architecture of de novo mutations in developmental disorders. *Nature* **542**, 433–438 (2017).
17. Singh, T. *et al.* The contribution of rare variants to risk of schizophrenia in individuals with and without intellectual disability. *Nat. Genet.* **49**, 1167–1173 (2017).
18. Satterstrom, F. K. *et al.* Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* **180**, 568–584.e23 (2020).
19. Fombonne, E. Epidemiological surveys of autism and other pervasive developmental disorders: an update. *J. Autism Dev. Disord.* **33**, 365–382 (2003).

20. Smith, M. E., Lecker, G., Dunlap, J. W. & Cureton, E. E. The Effects of Race, Sex, and Environment on the Age at Which Children Walk. *The Pedagogical Seminary and Journal of Genetic Psychology* **38**, 489–498 (1930).
21. Karczewski, K. J., Francioli, L. C., Tiao, G. & Cummings, B. B. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *BioRxiv* (2019).
22. Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
23. Samocha, K. E., Kosmicki, J. A. & Karczewski, K. J. Regional missense constraint improves variant deleteriousness prediction. *BioRxiv* (2017).
24. Turner, G. *et al.* Restoring reproductive confidence in families with X-linked mental retardation by finding the causal mutation. *Clin. Genet.* **73**, 188–190 (2008).
25. Sherman, S. L., Morton, N. E., Jacobs, P. A. & Turner, G. The marker (X) syndrome: a cytogenetic and genetic analysis. *Ann. Hum. Genet.* **48**, 21–37 (1984).
26. He, X. *et al.* Integrated Model of De Novo and Inherited Genetic Variants Yields Greater Power to Identify Risk Genes. *PLoS Genetics* vol. 9 e1003671 (2013).
27. Vulto-van Silfhout, A. T. *et al.* Mutations in MED12 cause X-linked Ohdo syndrome. *Am. J. Hum. Genet.* **92**, 401–406 (2013).
28. Schwartz, C. E. *et al.* The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. *J. Med. Genet.* **44**, 472–477 (2007).
29. Risheg, H. *et al.* A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. *Nat. Genet.* **39**, 451–453 (2007).
30. Bouazzi, H., Lesca, G., Trujillo, C., Alwasiyah, M. K. & Munnich, A. Nonsyndromic X-linked intellectual deficiency in three brothers with a novel MED12 missense mutation [c.5922G>T (p.Glu1974His)]. *Clin Case Rep* **3**, 604–609 (2015).
31. Lesca, G. *et al.* Clinical and neurocognitive characterization of a family with a novel MED12

- gene frameshift mutation. *Am. J. Med. Genet. A* **161A**, 3063–3071 (2013).
32. Prontera, P. *et al.* A novel MED12 mutation: Evidence for a fourth phenotype. *Am. J. Med. Genet. A* **170**, 2377–2382 (2016).
 33. Rubinato, E. *et al.* MED12 missense mutation in a three-generation family. Clinical characterization of MED12-related disorders and literature review. *Eur. J. Med. Genet.* 103768 (2019).
 34. Turner, T. N. *et al.* Sex-Based Analysis of De Novo Variants in Neurodevelopmental Disorders. *Am. J. Hum. Genet.* (2019) doi:10.1016/j.ajhg.2019.11.003.
 35. Niemi, M. E. K. *et al.* Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature* **562**, 268–271 (2018).
 36. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
 37. Sniekers, S. *et al.* Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat. Genet.* **49**, 1107–1112 (2017).
 38. Pardiñas, A. F. *et al.* Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* **50**, 381–389 (2018).
 39. Yamamoto, T. *et al.* Genomic backgrounds of Japanese patients with undiagnosed neurodevelopmental disorders. *Brain Dev.* **41**, 776–782 (2019).
 40. Fernández-Marmiesse, A. *et al.* Rare Variants in 48 Genes Account for 42% of Cases of Epilepsy With or Without Neurodevelopmental Delay in 246 Pediatric Patients. *Frontiers in Neuroscience* vol. 13 (2019).
 41. Yang, Y. *et al.* Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* **312**, 1870–1879 (2014).
 42. De Luca, C., Race, V., Keldermans, L., Bauters, M. & Van Esch, H. Challenges in molecular diagnosis of X-linked Intellectual disability. *Br. Med. Bull.* (2020) doi:10.1093/bmb/ldz039.
 43. Kaplanis, J. *et al.* Integrating healthcare and research genetic data empowers the discovery of 49

- novel developmental disorders. *bioRxiv* doi:10.1101/797787.
44. Garcia, S. *et al.* Identification of de novo mutations of Duchénnè/Becker muscular dystrophies in southern Spain. *Int. J. Med. Sci.* **11**, 988–993 (2014).
 45. Leslie, N. D. Haldane was right: de novo mutations in androgen insensitivity syndrome. *The Journal of pediatrics* vol. 132 917–918 (1998).
 46. Wu, Y. & Lyon, G. J. NAA10-related syndrome. *Exp. Mol. Med.* **50**, 85 (2018).
 47. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat. Commun.* **9**, 2098 (2018).
 48. Karavani, E. *et al.* Screening Human Embryos for Polygenic Traits Has Limited Utility. *Cell* **179**, 1424–1435.e8 (2019).
 49. Savage, J. E. *et al.* Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nat. Genet.* **50**, 912–919 (2018).
 50. Hill, W. D. *et al.* Genomic analysis of family data reveals additional genetic effects on intelligence and personality. *Mol. Psychiatry* **23**, 2347–2362 (2018).
 51. Wright, C. F. *et al.* Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet* **385**, 1305–1314 (2015).
 52. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. *Nature* **519**, 223–228 (2015).
 53. Akawi, N. *et al.* Discovery of four recessive developmental disorders using probabilistic genotype and phenotype matching among 4,125 families. *Nat. Genet.* **47**, 1363–1369 (2015).
 54. Martin, H. C. *et al.* Quantifying the contribution of recessive coding variation to developmental disorders. *Science* **362**, 1161–1164 (2018).
 55. Kohler, S. *et al.* Clinical diagnostics in human genetics with semantic similarity searches in ontologies. *Am. J. Hum. Genet.* **85**, 457–464 (2009).
 56. Bragin, E. *et al.* DECIPHER: database for the interpretation of phenotype-linked plausibly pathogenic sequence and copy-number variation. *Nucleic Acids Res.* **42**, D993–D1000 (2014).

57. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
58. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
59. Ramu, A. *et al.* DeNovoGear: de novo indel and point mutation discovery and phasing. *Nat. Methods* **10**, 985–987 (2013).
60. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
61. Manichaikul, A. *et al.* Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
62. Van Hout, C. V., Tachmazidou, I. & Backman, J. D. Whole exome sequencing and characterization of coding variation in 49,960 individuals in the UK Biobank. *BioRxiv* (2019).
63. Jia, T., Munson, B., Allen, H. L., Ideker, T. & Majithia, A. R. Thousands of missing variants in the UK Biobank are recoverable by genome realignment. *bioRxiv* (2019).
64. Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J. & Kircher, M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res.* **47**, D886–D894 (2019).
65. MacArthur, D. G. *et al.* A systematic survey of loss-of-function variants in human protein-coding genes. *Science* **335**, 823–828 (2012).
66. Robinson, J. T. *et al.* Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
67. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

Data availability

The full variant call files used in this study are accessible in the European Genome-Phenome Archive as dataset EGAD00001004389, and a file of phenotypic and family descriptions under EGAD00001004388. The *de novo* mutations used in the analysis are in Supplementary Table 3.

Acknowledgments

We thank the DDD families for participating, the DDD clinicians for recruiting patients, the Sanger Sample Management and Sequencing pipelines teams for generating the data, the Sanger Human Genome Informatics team for helping to process the exome data, and Nicola Whiffin for helpful

comments on the manuscript. The study was approved by the UK Research Ethics Committee (10/H0305/83, granted by the Cambridge South Research Ethics Committee and GEN/284/12, granted by the Republic of Ireland Research Ethics Committee). This work has been conducted using the UK Biobank Resource under Application Number 44165.

Funding: The DDD study presents independent research commissioned by the Health Innovation Challenge Fund (grant HICF-1009-003), a parallel funding partnership between the Wellcome Trust and the UK Department of Health, and the Wellcome Trust Sanger Institute (grant WT098051). The views expressed in this publication are those of the author(s) and not necessarily those of the Wellcome Trust or the UK Department of Health. 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). The research team acknowledges the support of the National Institutes for Health Research, through the Comprehensive Clinical Research Network. This study makes use of DECIPHER (<http://decipher.sanger.ac.uk>), which is funded by the Wellcome Trust.

Author contributions

H.C.M. analysed the DDD data and drafted the manuscript together with M.E.H who conceived and supervised the study. J.K. and K.E.S. assisted with the filtering of DNMs in the DDD study and modelling of mutation rates. A.S., N.A., J.M. contributed to analyses of the DDD exome data. R.Y.E. and G.G. conducted quality control on the DDD exome data. E.J.G. and M.D.C.N. analysed the UK Biobank exome data. A.L.T.T. examined clinical features of patients. M.E.K.N processed the chip genotype data and constructed polygenic scores. C.F.W., D.R.F. and H.V.F. provided clinical and analytical supervision.

Competing Interests Statement

M.E.H. is a co-founder of, consultant to, and holds shares in Congenica Ltd., a genetics diagnostics company. J.F.M is an employee of Illumina Inc. A.L.T.T. is an employee of Genomics England.