SUPPLEMENTARY INFORMATION

Experimental Protocols

Patient data and sample tracking

1,133 patients with severe, undiagnosed, developmental disorders and their parents were recruited and systematically phenotyped at 24 clinical genetics centres within the UK National Health Service and the Republic of Ireland. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC). Families gave informed consent for participation and specific additional consent for publication of photographs was sought and given by a subset of families.

Patient records are created by the regional genetics services within a study-specific area of DECIPHER (https://decipher.sanger.ac.uk), and sample barcodes are scanned into this record. Clinical data (growth measurements, family history, developmental milestones, etc.) are collected using a standard restricted-term questionnaire within DECIPHER, and detailed developmental phenotypes for the proband are entered using the Human Phenotype Ontology (Köhler et al., 2014). Patient and sample information is transferred from DECIPHER into an internal laboratory information management system (LIMS). This tracks samples through reception and the laboratory pipelines, and provides other internal tools. It is implemented as a set of Java webservices on top of a Postgres database.

Sample collection and processing

Saliva samples collected in barcoded Oragene-DNA OG-500 (parent) or OG-575 (child) collection tubes (DNA Genotek Inc.) were received from all family trio members. Sample arrival was recorded in a bespoke Laboratory Information Management System (LIMS). Samples were heated to 50^{o}C for 4 hours prior to DNA extraction, with a mix by inversion after 2 hours. Extraction of DNA from saliva samples was performed on a QIAsymphony robot (Qiagen) using the QIAsymphony DSP Virus/Pathogen Midi kit (Qiagen) with a customised purification protocol. DNA was eluted in to 96 well plates, which were passed to the Sample Management core facility (Wellcome Trust Sanger Institute) for QC analysis. A median DNA yield of 6 or 12ug was achieved from the child or adult saliva samples respectively. Blood-extracted DNA samples from the child were also sent from the Regional Genetic Service Laboratories in 2D barcoded tubes (FluidX). DNA sample arrival was recorded in the LIMS and samples were racked in batches of 95 and passed to the Sample Management core facility for QC analysis.

Sample QC (WTSI core facility)

Automated volume check (BioMicroLab) and assessment of concentration via pico green assay (Beckman FX, NX-96, Molecular devices DTX reader) were performed on all samples. DNA quality was also assessed via gel electrophoresis (Beckman FX and Invitrogen E-Gel system). A Sequenom SNP panel (including gender markers) was used to allow the sample identity checks following aCGH, genotyping and sequencing. Genotyping was performed using the iPLEX™ Gold Assay (Sequenom® Inc.). Assays for all SNPs were designed using the eXTEND suite and MassARRAY Assay

Design software version 3.1 (Sequenom® Inc.). Following PCR, unincorporated dNTPs were SAP digested prior to iPLEX™ Gold allele specific extension with mass-modified ddNTPs using an iPLEX Gold reagent kit (Sequenom® Inc.). SAP digestion and extension were performed according to the manufacturer's instructions with reaction extension primer concentrations adjusted to between 0.7-1.8μM, dependent upon primer mass. Extension products were desalted and dispensed onto a SpectroCHIP using a MassARRAY Nanodispenser prior to MALDI-TOF analysis with a MassARRAY Analyzer Compact mass spectrometer. Genotypes were automatically assigned and manually confirmed using MassARRAY TyperAnalyzer software version 4.0 (Sequenom® Inc.).

Array-CGH

aCGH analysis of 1 or 1.5 μg of child DNA was performed using 2 x 1M probe custom designed microarrays (Agilent; Amadid No.s 031220/031221). DNA labelling and hybridisation was carried out using protocols based upon the Agilent Bravo automated liquid handling platform with Enzymatic Labelling protocol V2.1. Child test DNA samples were labelled with Cy-5 and reference DNA (a pool of 500 male DNA samples) with Cy-3 using Agilent reagents. Labelled samples were cleaned up using AutoScreen-96A well plates (GE Healthcare). Combined test and reference DNAs were then co-hybridised to 2 x 1M probe Agilent microarrays for 3 nights (66 hours) in an Agilent hybridisation oven at 65^{o}C. Microarrays were washed using a Little Dipper wash station and then scanned at 3 μm resolution using an Agilent scanner C 2505C. Fluorescent intensities were extracted from the scanned images using Agilent Feature Extraction v10.5.1.1.

SNP Genotyping (WTSI core facility)

200ng of DNA supplied at 50ng/ul was processed to the standard Illumina protocol (Illumina SNP-genotyping protocol) using automated hardware (Tecan Freedom Evo). Tasks were performed from Make MSA3 onwards. Samples were applied to custom Illumina beadchips (SangerDDD_OmniExPlusv1_15019773_A) before hybridisation in an Illumina Hybridisation Oven after which the process was continued as instructed. Beadchips were scanned on Illumina iScans (Sanger ID N106, N111, N125, N127).

Exome Sequencing (WTSI core facility)

Genomic DNA (approximately 1 ug) was fragmented to an average size of 150 bp and subjected to DNA library creation using established Illumina paired-end protocols. Adapter-ligated libraries were amplified and indexed via PCR. A portion of each library was used to create an equimolar pool comprising 8 indexed libraries. Each pool was hybridised to SureSelect RNA baits (Agilent Human All Exon V3 Plus with custom ELID # C0338371) and sequence targets were captured and amplified in accordance with manufacturer's recommendations. Enriched libraries were subjected to 75 base paired-end sequencing (Illlumina HiSeq 2000) following manufacturer's instructions.

SNV and InDel validation by Capillary sequencing

Primers were designed to amplifying products 400-600 bp in length centred on the site of interest. Primer3 design settings were adjusted as follows: primer length - 18 bp +/-3, GC Clamp =1, Tm 60 +/-2, using a human mispriming library. Genomic DNA from all trio members, amplified by Whole Genome Amplification (WGA) using illustra Genomiphi HY or V2 Amplification Kits (GE Healthcare), was used as template DNA in the site-specific PCR reactions. PCR reactions were carried out using Thermo-Start *Taq* DNA Polymerase (Thermo Scientific), following the manufacturers protocol. The PCR products were assessed by Agarose gel electrophoresis and submitted for sequencing to the Faculty Small Sequencing Projects (WTSI core facility). Capillary sequence traces from all trio members were aligned and viewed using an in-house designed web-based tool and scored for the presence or absence of the variant.

Digital PCR

Droplet digital PCR (ddPCR) was performed as described (Hindson et al., 2011) with some modifications. Genomic DNA was restriction digested with EcoR1 (NEB) prior to ddPCR. Droplets were made using a droplet generator (Bio-Rad), PCR amplified and loaded into a QX100 droplet reader (Bio-Rad). ddPCR data was analysed using QuantaSoft software (Bio-Rad), which calculates copy number state per well. Assays targeting the regions of interest were designed using the Universal ProbeLibrary (UPL) Assay Design tool (Roche). Test regions were labeled in FAM (UPL, Roche) and run in parallel with the RPP30 reference assay (Hindson et al., 2011), labelled in either VIC (Life Technologies) or HEX (Bio-Rad). All assays were run in triplicate and as trios (proband, mother, father) to establish inheritance. At least one assay was designed within the variant and one in a region of normal copy number adjacent to the site.

FISH

Human bacterial artificial chromosome clones and fosmid clones, for validation of selected variants by fluorescence *in situ* hybridisation (FISH), were provided by the clone-archive team at the Wellcome Trust Sanger Institute. Plasmid DNA was purified using the PhasePrep BAC DNA kit (Sigma-Aldrich) following manufacturer's protocol. The plasmid DNA was first amplified using the whole genome amplification kit (WGA2, Sigma-Aldrich) following manufacturer's recommendations, and subsequently labelled with either ChromaTide™ Texas Red®-12-dUTP (Molecular Probes/Invitrogen) or Atto 488-dUTP (Jena Bioscience) using the whole genome reamplification kit (WGA3, Sigma-Aldrich) as described before (Gribble et al., 2013). For each hybridisation, approximately 100 ng of labelled DNA from each clone and 2-4 µg of human Cot-1 DNA (Invitrogen) were precipitated down using ethanol, then resupended in hybridisation buffer containing 50% formamide, 2×SSC, 10% dextran sulphate, 0.5 M phosphate buffer, pH 7.4.

Assays Designs

Sequenom Assay

The Sequenom genotyping data consist of two plates containing SNP genotypes, W30467 and W1180, for which the combined total of SNPs is 61 (58 unique, including 4 sex markers). Plate W30467 is the standard Sanger Sample Quality Control plex including 31 SNPs while W1180 is a custom DDD specific plex that includes an additional 30 CNV tagging SNPs.

Array-CGH Platform

The DDD array comparative genomic hybridisation (aCGH) platform is composed of 2 x 1M probe Agilent arrays and has been heavily targeted towards genes and ultra-conserved elements throughout the human genome. The entire set of Gencode genes (version 17), along with some high value regulatory and mRNA coding elements have been tiled, using a minimum of 5 oligo-nucleotide probes per exon. Additionally the array maintains the presence of a high-resolution backbone with a median probe spacing of 5Kb (Amadid No.s 031220/031221).

SNP Genotyping Platform

The DDD single nucleotide polymorphism (SNP) genotyping array is a customised version of the Illumina Omni-one quad chip. Extra content has been added to standardise the genome coverage, aiming to target the largest gaps in array coverage first and additionally inserting the best quality probe within the central gap region. The array includes 811,844 mapped markers (1,734 are unmapped) and has been designed to minimise gaps between probes with a median intermarker distance of 2,378bp (SangerDDD_OmniExPlusv1_15019773_A).

Exome Plus Sequencing Platform

The exome sequencing bait region design (Exome+) is a version of the Agilent Sanger-Exome (Human All Exome 50mb Kit) with an additional 57,680 bait regions. The bait region design uses normal default parameters of; 1x coverage and less than or equal to 20bp overlap with repetitive regions. Out of the 57,680 additional bait regions, 34,825 baits are used to cover 4,322 ultra conserved regions, 10,117 are used to cover 1,664 heart enhancers and 12,296 are used to cover 622 additional enhancer regions. The total number of bait regions used in the Exome+ design is 271,063 (Agilent Human All Exon V3 Plus with custom ELID # C0338371).

Coverage

For our aCGH experiments our array platform consisted of 1,932,856 probes. In protein coding genes (GENCODE v17), every exon encompassed on average 2 probes, 11% of exons contained no probes. Every protein-coding gene encompassed on average 72 probes, less than 10% of genes encompassed fewer than 10 probes.

For our exome sequencing, targeting 58.62 Mb (271,064 baits) of which 51.64 Mb (213384 baits) consisted of exonic targets (39 Mb) and their flanking regions, 6.9 Mb (57680 baits) of regulatory regions were targeted using custom baits. The median (n=3,399) average sequencing depth (ASD = bases

sequenced/bases targeted) was 90X across the whole targeted sequence or 93X across autosomal targets only. 95% of all samples had an average sequencing depth higher than 63X. At least 90% of all targeted regions have a median ASD higher than 15X. Only 16026 baits showed a median ASD smaller than 10X, comprising 800kb of protein coding sequence. More than 85% of all probes were consistently covered (ASD >10X) across the three samples of the trio in at least 90% of the 1133 trios.

Variant Calling

Array-CGH - CNV Detection

We have developed CNsolidate, a novel change point detection system. CNsolidate makes use of 12 independent change point detection algorithms (Pique-Regi et al., 2010; Olshen et al., 2004; Price et al., 2005; Picard et al., 2005; Benelli et al., 2010; Andersson et al., 2008; Erdman et al., 2007; Barry et al., 1993) and an expert voting system to detect CNV regions (CNVRs) from aCGH data. Each change point detection algorithm is weighted based on its estimated performance (Type 1 and Type 2 error rates) across a range of data noise measures. These noise measures are drawn from each input dataset and used to calculate a combined CNV confidence measure (w-score). The w-score is a composite value based on the combined performance of all algorithms contributing to the detection of each CNVR, given the scale of all data noise values measured. CNsolidate generates an annotated single sample VCF file containing the detected CNVRs annotated with CNV confidence measures, copy number state estimates, population frequency values and gene annotations for each sample.

SNP Genotyping - Inheritance Classification

Log R ratios (LRR) and B-allele Frequencies (BAF) are calculated for each SNP for each individual using the Illumina GenomeStudio software. We have developed VICAR (Variant Inheritance Classification Algorithm in R), a novel Bayesian framework for the classification of inheritance statues using SNP genotyping data. VICAR uses the copy number state likelihoods obtained using the validate option from PennCNV (Wang et al., 2007) to determine whether a CNV identified from the aCGH data in an offspring is de novo or inherited. The inheritance classifications determined by VICAR are added as an annotation onto each CNV called in the proband using aCGH.

SNP Genotyping - UPD Detection

Uniparental disomy was detected using a trio-based strategy as previously described (King et al., 2014). For the detection of UPD events in probands UPD-informative positions were classified by inheritance (maternal, paternal) and by type (isodisomy, heterodisomy), and the number of informative genotypes was tallied for each chromosome. Chromosomes harbouring an enrichment of UPD-type proportions were flagged as potential UPD events if they were statistically unlikely using a binomial test for each chromosome, each inheritance-type combination with a multiple test adjusted p value of 0.000568.

SNP Genotyping - Mosaicism Detection

Detection of mosaic large-scale loss of heterozygosity and copy number variation events was performed with Mosaic Alteration Detection (MAD) (Gonzalez et al., 2011), using default parameter values. This tool identifies a putative mosaic event as a segment of SNP genotypes with a consistent aberration in B-allele frequency, followed by classification into loss of heterozygosity, gain, or loss based on the average log R ratio values of positions within the segment.

Exome Sequencing - SNV and INDEL Detection

All Binary Alignment/Map (BAM) files are processed using the Genome Analysis Production Informatics (GAPI) pipeline at the Wellcome Trust Sanger Institute. The GAPI pipeline is used to call single nucleotide variants (SNVs) and insertion/deletion variants (INDELs) from whole exome sequence data. Reads are mapped to the reference genome (GRCh37_hs37d5), duplicate fragments are marked using Picard (version 1.46), local realignment around indels is performed with GATK (version 1.1) and GATK is then used to recalibrate base qualities. Variants are only reported within the bait regions +/- 100bp. SNVs are called with GATK using the UnifiedGenotyper, SNVs and INDELs are called with samtools (version 0.1.16) mpileup options -d 500 -C50 -m3 -F0.002 and variants are filtered using the vcfutils.pl utility and options -p -d 4 -D 1200 from samtools. A further set of indels is called using a dedicated indel caller, Dindel (version 1.01). The GAPI pipeline produces individual single sample variant call formatted (VCF) files for each caller (samtools, GATK and Dindel) that is then combined into a merged VCF file. Merging conflicts are resolved following a caller order; [dindel, gatk, samtools]; where the primary caller (the first in the list) defines the position and genotype of the variant.

Exome Sequencing - *De novo* mutation detection

DeNovoGear version 0.2 (Ramu et al., 2013) is used to detect *de novo* mutations (SNVs and INDELs) from trio exome data (BAM files). DeNovoGear results are merged into proband VCF files and where positional conflicts occur the annotations provided by DeNovoGear are added to the existing variant.

Exome Sequencing - CNV Detection

We have developed CoNVex, a novel CNV detection algorithm for exome sequencing data. CoNVex utilises the read depth information from targeted bait regions and computes a log2 ratio by comparing the read depth at each position to a reference made up of the median read depth across a number highly correlated samples. A GC based correction is applied and the log2 ratios are converted into an error weighted score (ADM-2 scores). Copy number variable regions (CNVRs) are detected using the ADM-2 scores and the Smith-Waterman algorithm (Price et al., 2005). CoNVex generates a single sample VCF file containing the detected CNVRs for each sample.

CNVR Merging - CoNVex and CNsolidate

CNVRs detected by CoNVex and CNsolidate are merged into a single sample CNV-VCF formatted file. CNVRs sharing any overlap between the two callers are merged and the break point positions are taken from CNsolidate in preference due to the superior coverage achieved on the aCGH platform. The

CNV-VCF files are then combined with the merged SNV and INDEL VCF file to produce a single sample merged (uber-VCF) file for each proband.

Annotation

Minor Allele Frequency

To define the rarity of each SNV and INDEL, the uber-VCF is further annotated with minor allele frequency (MAF) data from a variety of different sources. A subset of these MAF annotations are used to define a maximum allele frequency (MAX_AF) for use in subsequent frequency based variant filtering. The MAF annotations used to define the MAX_AF include data from 4 different populations of the 1000 Genomes project (Abecasis et al., 2010) [AMR, ASN, AFR & EUR], the UK10K cohort, the NHLBI GO Exome Sequencing Project (ESP) and an internal DDD allele frequency generated using unaffected parents. For allele matching of SNVs we use an exact match based on a key generated from four values (chromosome, position, reference allele and alternative allele). For allele matching of INDELS we use a less stringent approach where the key is constructed using a different four values (chromosome, position, slice and direction). This key requires both INDELs to be at the same locus (chromosome and position) while the slice is computed based on the DNA sequence difference between the reference and alternative alleles and direction is either deletion or insertion. For the rarity of CNVs a number of external and internal normal CNV control data sets (CNV Consensus), comprising of merged copy number events (CNVEs) from various different studies, are used to add a population frequency estimate to each CNVR. These control CNVE sets include the 42M CNV Project Hapmap Study (Conrad et al., 2009) the Wellcome Trust Case Control Consortium (WTCCC) (Craddock et al., 2010), the 1000 Genomes Project (Abecasis et al., 2010) and the DDD normal controls (UKBS and Generation Scotland). CNVRs sharing more than 80% of their boundaries with any CNVE of the same type (deletion or duplication) observed at more than 1% population frequency in any CNVE set are classified as common.

Variant Effect Predictor

To define the functional consequence of each variant (SNVs, INDELs and CNVs) annotations from the ensembl variant effect predictor 2.6 (VEP) (McLaren et al., 2010) based on ensemble gene build 68 are added to the uber-VCF file. VEP produces a number of annotations including, SIFT and Polyphen predictions, ensemble transcripts, HGNC gene names and a prediction of the functional consequence for each variant. The transcript with the most severe consequence is selected and all associated VEP annotations are based on the effect that the variant has on that particular transcript. Furthermore, based on the functional consequence prediction form ensembl VEP variants are categorised into one of 4 different broad consequence terms:

- *Loss of function* (VEP consequences: transcript ablation, splice donor variant, splice acceptor variant, stop gained, frameshift variant)

- *Functional* (VEP consequences: stop lost, initiator codon variant, inframe insertion, inframe deletion, missense variant, coding sequence variant)
- *Silent* (VEP consequences: synonymous variant, stop retained variant
- *Other* (VEP consequences: transcript amplification, splice region variant, incomplete terminal codon variant, mature miRNA variant, 5 prime UTR variant, 3 prime UTR variant, intron variant, NMD transcript variant, non coding exon variant, nc transcript variant, upstream gene variant, downstream gene variant, TFBS ablation, FBS amplification, TF binding site variant, regulatory region variant, regulatory region ablation, regulatory region amplification, feature elongation, feature truncation, intergenic variant)

Quality Control

Sequenom

Analysis of the Sequenom data is performed at the SNP, individual, and family level. This is the first quality control (QC) step and occurs as soon as any sample arrives at the Wellcome Trust Sanger Institute. Individual samples are evaluated for DNA sample quality, call rate and average heterozygosity. Individual samples with a call rate < 0.74 or a heterozygosity value below 0.195 or above 0.756 are failed. Families are evaluated for discrepancies between the stated gender versus the genotyped gender and for the likelihood of the pedigree structure recorded in DECIPHER. The most plausible pedigree structure given the data for a family is determined using an algorithm that calculates the Bayes factor for six different possible pedigree states [Standard Trio, Father-offspring swap, Mother-offspring swap, Non-paternity, Non-maternity, All unrelated]. We use a relatively generous threshold to determine which families are classified as standard trios and all families where the log(Bayes Factor) for the standard trio is the maximum will pass the pedigree structure check. Any trio that appears to be non-standard is evaluated manually before any further sample processing is allowed to occur.

Array-CGH

For call QC of the aCGH data we apply the recommended detection quality filtering criteria from CNsolidate; comprising of a w-score above 0.468, a p-value below 0.01, and the mean log2 ratio below -0.41 for deletions and above 0.36 for duplications when passing individual CNV calls. For sample QC we use an exclusively post-calling approach and apply a robust clustering algorithm (“Aberrant”) for outlier identification and exclusion (Bellenguez et al., 2012) to the total number of passed CNVRs per sample versus the proportion of passed rare CNVRs per sample. We define rare as CNVRs that do not share greater than 80% of their boundaries with a copy number event (CNVE) of the same type, deletion or duplication, observed at more than 1% in the CNV Consensus list. Each aCGH sample is made up of 2 slides (2 x 1 million probe Agilent arrays) and any slides defined as outliers by the “Aberrant” clustering method are failed, where either slide fails for a sample the overall sample is also failed. Additionally we apply further exclusions, at the sample level, based on a low sensitivity cut-off of less than 40 QC passed detections

and a deletion / duplication ratio of greater than 10. Finally we apply a data tracking check to ensure that the CNV data is consistent with the Sequenom data linked to the same sample identifier. We make use of 28 copy number tagging SNPs to allow the aCGH data to be correlated with the SNP genotypes obtained via the Sequenom assay run at sample reception. Both slides on the aCGH platform contain dedicated probes tiling the 28 CNVRs tagged by SNPs present on the Sequenom assay. First, CNVtools (Barnes et al., 2008) is used to assign a copy number state to each SNP-tagged CNVR and the probability of observing all copy number states (aCGH data) given all observed genotypes (Sequenom data) is calculated using a Bayesian framework. The probabilities of observing individual copy number state - SNP genotype pairs are derived empirically and continually updated as more data is processed through the system.

SNP Genotyping

For sample QC on the SNP genotyping array we use the average heterozygosity and call rate for each individual and discard samples with a call rate below 0.95 and/or average heterozygosity greater than 3 standard deviations from the mean of the batch. We examine the relationships between members of a trio or duo using the kinship coefficient from KING (Manichaikul et al., 2010) for each pair within the family. We check whether individuals within a family are related in the way specified by the defined pedigree and whether the parents appear to be related. The default thresholds recommended by KING are used for defining different levels of relatedness between individuals. Finally we calculate the concordance between the genotypes from the SNP-array and the Sequenom assay using 53 SNP markers, individuals with less than 90% concordance between the SNP-array and the Sequenom assay are flagged as potential errors.

Mosaicism

As mosaicism disrupts commonly implemented SNP-genotyping sample-level QC metrics, such as standard deviations of B-allele frequency and log R ratio, we did not institute these sample-level QC metrics. Instead, we devised a downstream filtering strategy to filter at the level of detected segments. Manual inspection of a large subset of initial MAD putative detections identified four sources of technical error: 1) hypersegmentation, 2) constitutive duplications, 3) segments with absence of heterozygous genotypes (reflecting constitutive regions of homozygosity), and 4) segments with heterozygous BAFs skewed unimodally. We developed a computational strategy to mitigate these errors. First, we managed hypersegmentation by merging nearby (within 1 Mb) segments representing the same event type (loss, gain, or loss of heterozygosity). Second, we filtered constitutive duplications if the segment had a log R ratio or B deviation within the constitutive duplication centroid identified from 1,813 high-quality inherited copy number variations. Third, we identified constitutive homozygosity as segments with a very low (<5%) ratio of heterozygous to homozygous genotypes. Lastly, we avoided false segments in genomic waves by using the R language density function to calculate the peak heights of heterozygous BAF clusters, and filtered segments with a great relative difference of tallest peak to next-tallest peak height, a signal reflecting a single band of heterozygous genotypes, and thus

non-mosaic. By implementing these strategies we reduced the number of putative detections to a few dozen and were able to manually curate this number. Copy number mosaic events were validated by FISH.

Exome Sequencing - SNVs and INDELs

Low confidence variant calls from each caller are failed and are not including during any variant merging steps. Variants displaying the following criteria from samtools are failed: read depth (DP) less than 4, DP greater than 1200, mapping quality (MQ) less than or equal to 10, strand bias p value (StrandBiasPval) less than 0.0001, base quality bias p value (BaseqBiasPval) less than 1e-100, mapping quality bias p value (MapqBiasPval) less than 0 and the number of base pairs from a gap in the genome build (MinbpfromGap) less than 10. Variants displaying the following criteria from GATK are failed: DP less than 4, DP greater than 1200, variant quality (QUAL) less than 30, quality by depth (QD) less than 5, homopolymer run length (HRun) greater than 5, strand bias (SB) greater than 10, mapping quality zero (MQO) greater or equal to 4 and MQO greater than or equal to MQO/(1*DP). Variants displaying the following criteria from Dindel are failed: HRun greater than 10, and quality score less than 20.

Variants from DeNovoGear are further filtered based on the following set of pass criteria: internal minor allele frequency less than or equal to 0.01, UK10Ktwin minor allele frequency less than or equal to 0.01, segmental duplications and tandem repeat regions excluded, and variant present in child VCF but not the parental VCF files.

To ensure that each VCF file is associated to the correct sample we calculate the concordance between SNP genotypes in each VCF file and the Sequenom assay, files with less than 90% concordance are flagged as potential errors. We exclude data sets displaying high levels of sample contamination based on the extreme alternative allele fraction. For each site, we compute the allelic depth for the alternative allele divided by total depth, (AD/DP). For heterozygous SNVs if AD/DP is less than 0.15 or AD/DP is greater than 0.8 the SNV is classed as an outlier; for homozygous SNVs if AD/DP is less than 0.95 the SNV is classed as an outlier. Samples with greater than 4000 common extreme levels of heterozygous SNVs are failed due to high levels of sample contamination.

Exome Sequencing - CNVs

For sample QC of the exome CNV data we fit a lowess between the median absolute deviation and the mean number of detections for deletions and duplication separately per sample. Samples that are greater than 3x the standard deviation away from the lowess curve for either deletions or duplications are failed. To fail CNV calls we use a number of different filters aimed at increasing the specificity for rare CNVRs. The fail criteria for all autosomal CNVRs called by CoNVex are: type is DUP and convex score less than 7, size greater than 500,000 bp and convex score less than 10, size greater than 200,000 bp and convex score less than 10 and number of probes greater than or equal to 10, mean log2 ratio greater than or equal to 1.5, forward overlap with common CNVEs from the CNV consensus is equal to 0

and the internal frequency (rc50) is greater than 0.05, mean log2 ratio greater than 0 and type is DEL, mean log2 ratio less than 0 and type is DUP. For rare CNVRs an additional set of fail criteria are applied, where rarity is defined as CNVRs that do not share greater than 50% of their boundaries with a copy number event (CNVE) of the same type, deletion or duplication, observed at more than 1% in the CNV Consensus list. The rare CNVR fail criteria are: type is DUP and number of probes equals 1 and convex score less than 10, convex score less than 10 and number of probes equals 1 and size less than 500 bp, mean log2 ratio greater than -0.5 and type is DEL and convex score less than 10, mean log2 ratio less than 0.29 and type is DUP and convex score less than 10, size less than 15000 bp and convex score less than 10 and internal frequency (rc50) greater than 0.01, type is DUP and number of probes greater than 5 and convex score less than 8. Finally, for CNV calls made by CoNVex on chromosome X we apply a differential absolute size cut-of depending on how far away the sample is from the lowess curve fitted during sample QC. All CNV calls on chromosome X greater than 10,000,000 bp in size are failed and additionally, for samples defined as outliers during sample QC, all CNV calls greater than 500,000 bp in size on chromosome X are failed.

Family History and Consanguinity

Family history

Family history was assessed by a brief online questionnaire in DECIPHER completed by the patient's local clinical team. Relevant phenotypes for affected parents were also recorded using the Human Phenotype Ontology (Köhler et al., 2014).

Trio	Number (%)
Proband only affected	812 (72%)
One parent affected	93 (9%)
Both parents affected	18 (2%)
Siblings affected	157 (14%)
Second degree relatives affected	87 (8%)
Unknown	37 (3%)

Consanguinity

Self-assessed consanguinity was record online in DECIPHER by the patient's local clinical team. Identity-by-descent (IBD) was also calculated from exome data using KING (Manichaikul et al., 2010).

		IBD-defined	
		No	**Yes**
Self-declared	**No**	1068	14
	Yes	18	33

Analysis of Known Causes of Developmental Disorders

To facilitate the scalable identification and clinical review of possibly pathogenic variants in genes known to harbour variants causing developmental disorders, we established a Developmental Disorder Gene to Phenotype database, DDG2P (Wright et al. 2014). We systematically classified the confidence with which a gene-phenotype relationship has been established (e.g. Confirmed, Probable, Possible) as well as the underlying genetic mode of inheritance (e.g. dominant, recessive) and the likely mechanism by which variants cause the disorder (e.g. Loss-of-function, Activating, etc). The November 2013 version of DDG2P used in these analyses contains 1,129 genes and 1,636 gene-phenotype pairs that are sufficiently established to identify variants with a high likelihood of diagnostic relevance. Rare (minor allele frequency < 1%) variants predicted to alter the encoded protein and consistent with the known allelic requirement and mutational mechanism of that gene-phenotype pair were reviewed by a multi-disciplinary team including senior clinical geneticists. DDG2P is available for download from DECIPHER (https://decipher.sanger.ac.uk).

Mutation Rates

Modelling mutation rates

Gene-specific mutation rates for different functional classes of single nucleotide variants (missense, silent, nonsense, canonical splice site, loss of stop codon) were provided for 18,272 genes by Mark Daly and Kaitlin Samocha (Samocha, Robinson et al. 2014). These gene-specific mutation rates are based primarily on estimated triplet-specific mutation rates, thus taking into account sequence context and gene size, and accurately predict the amount of synonymous variation seen in coding sequences. We estimated the frameshift mutation rate by scaling the nonsense mutation rate by the ratio of the number of experimentally validated de novo frameshift variants (N=95) and the number of experimentally validated nonsense variants (N=102). We estimated the rate of inframe deletions and insertions by scaling the frameshift mutation rate by the ratio of inframe and frameshift indels (1/9) observed genome-wide, predominantly regions not under strong negative selection (Mills et al., 2006). The mutation rate of genes on chrX was estimated separately for the paternal and maternal germline by applying an estimate for the male driven mutation parameter (alpha) of 3.4, based on an unpublished dataset of 199 de novo mutations phased in an unbiased manner to the paternal or maternal haplotype. The loss-of-function mutation rate for each gene was estimated by summing the mutation rates of nonsense, canonical splice sites and frameshift variants. The ‘functional’ mutation rate was estimated by summing the loss-of-function mutation rate with the rate of missense variants and inframe indels.

Estimating expected numbers of mutations of different functional classes in sets of genes in a specified number of individuals

For autosomal genes the gene-specific mutation rates of different functional classes of variant were multiplied by the number of transmissions (twice the number of probands) to give the total expected number of mutations per gene given the number of probands sequenced. For X-linked genes, the gene-specific mutation rates of different functional classes of variant in paternal and maternal germlines were multiplied by the number of paternal and maternal transmissions, respectively, and then summed, to give the total expected number of mutations per gene given the number of male and female probands. For a given set of genes, the expected number of mutations of different functional classes per gene was summed to give the expected number of mutations of different functional classes across the set of genes. To estimate the significance of the observed number of mutations of a given functional class in a specific gene set, the expected number of mutations in that class was assumed to be the mean of a Poisson distribution, and the probability of drawing from that distribution a number of mutations equal or greater than the observed number of mutations was calculated.

Estimating number of genes expected to be recurrently mutated under the null hypothesis of no gene-specific mutation enrichment

We simulated the number of genes expected to be recurrently mutated by chance given a specified number of observed mutations (functional or loss-of-function only), by assigning mutations at random to genes according to their mutation rate (estimated as described above) and tabulating the number of genes with more than one mutation. Each scenario was simulated 10,000 times to estimate the shape of this distribution. The excess of recurrently mutated genes was estimated as the number of observed recurrently mutated genes minus the median from the simulated distribution. This simulation was run for four different scenarios;

- all functional mutations in all DDD families
- loss-of-functional mutations in all DDD families
- all functional mutations in undiagnosed DDD families
- loss-of-functional mutations in undiagnosed DDD families

Assessing significance of gene-specific mutation enrichment

To evaluate the statistical significance of mutation enrichment in individual genes, we estimated the expected total number of mutations in each gene (separately for functional and loss-of-function mutations) in the 1,133 DDD probands, as described above. To estimate the significance of the observed number of mutations of a given functional class in a gene, the expected number of mutations in that class was assumed to be the mean of a Poisson distribution, and the probability of drawing from that distribution a number of mutations equal or greater than the observed number of mutations was calculated. This test was performed on DDD families under four scenarios:

- all functional mutations in all DDD families
- loss-of-functional mutations in all DDD families
- all functional mutations in undiagnosed DDD families
- loss-of-functional mutations in undiagnosed DDD families

This test was also run on the meta-analysis of DDD families plus published mutations in parent-offspring trios with intellectual disability, autism,

schizophrenia, epileptic encephalopathies and congenital heart defects (Zaidi et al., 2013; Fromer et al., 2014; de Ligt et al., 2012; Rauch et al., 2012; Allen et al., 2013; Iossifov et al., 2012; Sanders et al., 2012; Neale et al., 2012; O'Roak et al., 2012). All published mutations were re-annotated for functional consequences using the same pipeline as was used for the DDD families. In these meta-analyses, only genes with mutations in the DDD trios were evaluated for statistical significance. The test was performed under four scenarios:

- all functional mutations in all DDD families + published studies
- loss-of-functional mutations in all DDD families + published studies
- all functional mutations in undiagnosed DDD families + published studies
- loss-of-functional mutations in undiagnosed DDD families + published studies

The results of the significance testing of mutations in all DDD families (in isolation and in the meta-analysis with published studies) was used for demonstrating the power of this approach by re-discovering known DD genes, but was not used for the novel gene discovery analyses. For the novel gene discovery analyses only the testing of mutations in undiagnosed DDD families was used (in isolation and in the meta-analysis with published studies).

Mutation Clustering Analysis

De novo SNV mutations were analysed for clustering within genes. De novo SNVs identified in genes were compared to simulated dispersions of mutations within each gene. De novo variants originated from two sources – 1) the DDD study, and 2) a meta-analysis including de novos from DDD and nine independent studies (Zaidi et al., 2013; Fromer et al., 2014; de Ligt et al., 2012; Rauch et al., 2012; Allen et al., 2013; Iossifov et al., 2012; Sanders et al., 2012; Neale et al., 2012; O'Roak et al., 2012). Exon coordinates and sequences for de novo containing genes were retrieved from Ensembl, using the longest transcript available that contained all the source de novo variants. Nucleotide mutation rates in trinucleotide contexts were provided by Kaitlin Samocha and Mark Daly. De novo mutations were simulated within each gene, weighted by the context specific mutation rates. Each nucleotide within the coding sequence of a gene was evaluated against the possible alternate bases. Alternate nucleotides that modified the amino acid sequence, or that altered splice sites were randomly sampled. Mutations were randomly sampled within each gene, matching the number of sampled mutations to the number of known de novos. The de novos were assessed for their tendency to cluster within close proximity to each other. The clustering distance was calculated as the geometric mean coding distance between all the possible de novo pairs. The distances for the observed de novos were compared to distributions obtained from the simulated de novos. P-values were determined as the proportion of permutations with a distance less than or equal to the known de novo distance. Each gene underwent 1 million permutations to obtain a P-value. A small subset, genes with P-values at the minimum possible value, underwent additional permutations, up to a maximum of 64 million permutations per gene test.

Combining p values for mutation clustering and mutation enrichment
Fisher's method was used to combine the significance testing of mutation enrichment and mutation clustering (described above). This combined p value was only generated for significance testing of all functional mutations and was not used for significance testing for loss-of-function mutations. The intuition behind this is that genes enriched for loss-of-function mutations will be predominantly operating by a mechanism of haploinsufficiency, which does not predict significant clustering of mutations, whereas gene enriched for other classes of functional mutations, predominantly missense mutations, could be operating by dominant negative of activating mechanisms, which are likely to be clustered at particular sites within the coding sequence of the gene.

Estimation of multiple-testing correction for gene-specific enrichment and clustering of de novo mutations
To establish a genome-wide significance threshold, we used Bonferroni correction for multiple testing. Bonferroni testing assumes that tests are independent of one another, which is clearly not the case when loss-of-function mutations are included in both the loss-of-function test of enrichment and the functional test for enrichment. However, this lack of independence testing means that a Bonferroni significance correction assuming independence will be overly conservative. In the testing of mutations in DDD families in isolation, all 18,272 genes for which mutation rates could be estimated were tested. Whereas for the meta-analysis testing of the combined dataset of published mutations and mutations in DDD families, only genes with at least one functional mutation in DDD families were tested. Genes without any functional mutations in DDD families were not tested in these meta-analyses.

When demonstrating the power of the mutation significance testing by re-discovering known genes, four tests were performed [number of tests]:

- all functional mutations in all DDD families (enrichment + clustering) [18,272]
- loss-of-functional mutations in all DDD families (enrichment only) [18,272]
- all functional mutations in all DDD families + published studies (enrichment + clustering) [980]
- loss-of-functional mutations in all DDD families + published studies (enrichment only) [980]

Thus the total number of tests for demonstrating the power of mutation significance testing was 38,504 and thus we applied a significance threshold of $0.05/38{,}504 = 1.30 \times 10^{-6}$

For novel gene discovery, in the meta-analysis only genes with at least two functional mutation in DDD families were tested, and four tests were performed [number of tests]:

- all functional mutations in undiagnosed DDD families (enrichment + clustering) [18,272]

- loss-of-functional mutations in undiagnosed DDD families (enrichment only) [18,272]
- all functional mutations in undiagnosed DDD families + published studies (enrichment + clustering) [97]
- loss-of-functional mutations in undiagnosed DDD families + published studies (enrichment only) [97]

Thus the total number of tests for novel gene discovery testing were: 36,738. Thus we applied a significance threshold of 0.05/36,738= 1.36 $x10^{-6}$.

Loss of Function Saturation analysis

To evaluate how the discovery of haploinsufficient genes might saturate with increasing sample size, we estimated the power to detect a significant enrichment of loss-of-function mutations, for each gene in the genome, across a range of sample sizes, from 1,000 to 12,000 trios. For a given gene and a given sample size, to estimate power to detect a significant enrichment of loss-of-function mutations we first calculated the integer number of mutations required to give a p value for the mutation enrichment test described above that exceeds 0.05/18,000, which represents an approximate Bonferroni correction for testing all genes in the genome. We then estimated the Poisson probability of observing at least this many loss-of-function mutations given the loss-of-function mutation rate of the gene, under the assumption that all individuals with loss-of-function mutations will be observed within the fraction of the population with severe developmental disorders. This analysis requires several assumptions: (i) there are no significant differences in mutability between haploinsufficient and haplosufficient genes in the genome, (ii) the penetrance of loss-of-function variants in haploinsufficient genes is 100%, (iii) the prevalence of severe developmental disorders is 0.5%, (iv) the most powerful way to identify haploinsufficient genes is significance testing of mutation enrichment for truncating variants, (v) loss-of-function variants in haploinsufficient genes do not result in increased rates of spontaneous miscarriage.

Untransmitted Diplotypes Control Dataset

We generated a population based control dataset of untransmitted diplotypes using the untransmitted haplotypes from the parents of the affected probands in 1,080 non-consanguineous trios. We hypothesised that if the cause of the proband's developmental disorder is genetic then it results from a variant or variants they carry or a structural rearrangement or imprinting defect within their genome. Therefore an individual who inherited the variants carried by both parents, that the proband did not inherit, is predicted to be healthy. To prevent generating untransmitted diplotypes with homozygosity by descent, consanguineous families defined by KING score > 0 (Manichaikul et al., 2010) were removed from this analysis.

An exome variant profile for the untransmitted diplotypes control was generated for each trio. The trio VCF files (mother, father, child) were merged and the following variants removed: Non 'PASS' variants, INDELS, variants

involving a multiallelic reference or alternate allele, CNVs, X and Y chromosome variants, intronic and upstream variants and variants with a QUAL score <179. The genotype in the untransmitted diplotype controls was calculated based on the genotypes of the mother, father and proband. Variants that did not fit with Mendelian inheritance were removed.

Variants were categorised according to the following classification adapted from (Purcell et al., 2014):

- *Loss of Function (LoF)*: Stop gained, transcript ablation, splice donor variant, splice acceptor variant, frameshift variant
- *Damaging (Dam)*: Functional variants predicted to be damaging by two algorithms: (SIFT = deleterious and PolyPhen = Probably damaging)
- *Functional*: Missense, inframe deletion, inframe insertion, coding sequence variant stop lost (not fulfilling the above criteria for 'Damaging')
- *Silent*: Synonymous variant

Cumulative counts of rare (MAF < 5%) homozygous and compound heterozygous loss of function and damaging functional variants were generated for the probands and untransmitted diplotype controls (see Supplementary Table S5 and Extended Data Table 3). The probands' exome variant profiles used in analysis were processed in the same way as the controls, i.e. they had had specific variant types removed as above. Probands with a likely dominant cause of their disorder (either a diagnostic *de novo* mutation or an affected parent) were identified and analysed as a separate group 'likely dominant probands' (Extended Data Table 3).

Model Organisms

Gene selection

We first identified all de novo nonsense, splice damaging, and frameshift variants classified as loss of function, as well as homozygous and compound heterozygous LOF variants from our cohort. The clinical phenotypes of the cases were assessed for likelihood of being picked up by zebrafish larval phenotyping following antisense-based knockdown. The assessment took into consideration comparative anatomy, phenotyping approaches, as well as the effective window of antisense activity (decreasing over 5 days). Our candidate gene list was further refined using a number of criteria including novelty (no informative animal model), literature background, available expression data, and burden in our cohort. A number of de novo missense variants deemed likely pathogenic were also selected for this analysis. The gene selection process identified 32 candidate genes for functional analysis.

Zebrafish morpholino and primer selection

Zebrafish orthologues of candidate human genes were identified through the ENSEMBL Compara database, ZFIN gene database, or blast searches, and refined through sequence and synteny comparisons. Morpholinos (MOs, Genetools LLC) are modified synthetic antisense oligonucleotides that bind mRNA and pre-mRNA, and knockdown protein expression through steric

hindrance of translational machinery, or disruption or normal splicing. MOs were mostly targeted to preferentially 5' exon splice donor sites, such that MO induced exon skipping should result in frame shifted transcripts, and intron retention should introduce a premature stop codon into the aberrant transcript. The specific MO sequences were selected from our target regions by the manufacturer (Genetools, LLC, Supplemental table "*MODELING Morpholinos and Primers* "). To verify the specific knockdown activity of each splice blocking MO, we designed PCR primers in the region surrounding the target sequence (Supplemental table "*MODELING Morpholinos and Primers* "). RNA was isolated from injected embryos at 2 days post injection using TRIZOL(Invitrogen), followed by reverse transcription using random hexamer primers (SuperscriptIII kit, Invitrogen) to generate cDNA. This cDNA was used in splice detecting PCRs to determine MO efficacy: loss of wildtype spliced product and/or increase in intron retention in mRNA was taken as an indication of specific activity (Bill et al., 2009). (see Extended Data Figure 8).

Replication

From the initial round of MO injections, a subset of genes with knockdown phenotypes was selected for replication. A second, non-overlapping MO was designed for these experiments (Supplemental table 6). Phenotypic analysis was performed as above on these replication MOs to assess the specificity of our screen.

Zebrafish morpholino injections

Zebrafish were maintained in accordance with UK Home Office regulations, UK Animals (Scientific Procedures) Act 1986, under project license 80/2192, which was reviewed by the Wellcome Trust Sanger Institute Ethical Review Committee. Wild-type and transgenic zebrafish embryos were obtained by natural spawning and raised at 28.5°C, as described (Westerfield 1993). All injections were performed alongside control MO within the same clutch. As co-knockdown of tp53 expression has been shown to eliminate all known morpholino toxicity (Bill et al., 2009; Gerety et al., 2011; Robu et al., 2007), gene specific MOs were co-injected with 6ng of tp53 MO (GCGCCATTGCTTTGCAAGAATTG) (Langheinrich et al., 2002). Any injected clutches that displayed the distinctive head/eye apoptosis phenotype associated with morpholino toxicity were also discarded (Bill et al., 2009; Gerety et al., 2011; Robu et al., 2007). These preventative and remedial measures ensured that clutches with phenotypes related to known MO toxicity were not analysed. All morpholinos were dissolved in water for 10 min at 65 °C, at 1 mM and stored at room temperature in glass vials. One- to four-cell embryos were microinjected with 1.8 nl of MO diluted in water. A dose curve was performed in which we established that 6ng of specific MO plus 6ng of tp53 MO would not induce off-target toxicity, and was therefore selected for all knockdown experiments. Knockdowns were repeated a minimum of two times at this dose (6ng specific MO + 6ng tp53 MO). Additional injections at lower doses of specific MO were performed when the severity of the knockdown precluded interpretation. When single knockdowns gave no phenotype, and a second zebrafish orthologue was present, double knockdown animals were phenotyped (5ng of each specific MO + 6ng tp53 MO).

Phenotyping

All MO injected embryos and their controls were observed over 3-5 days of development for abnormal phenotypes by two independent observers using light microscopy, based on published criteria (Kettleborough et al., 2013). Phenotypic findings were recorded and compared. Patient phenotype-aware observations ensured that relevant observations were made (i.e. head size measurements for potential microcephaly genes).

Microcephaly assays

For genes from patients with microcephaly, 10-20 knockdown embryos were photographed from dorsal and lateral aspects at day 2. These photos were measured for microcephalic changes using interocular distance as a surrogate for head size (Dauberet al., 2013; Golzio et al., 2012). These were repeated a minimum of 2 times. Statistical significance and p-values were determined using a two-tailed student's *t*-test. Significant changes in head size were followed up by imaging of neuronal tissue using anti-HuC/D antibody staining and confocal imaging (see below), with subsequent measurements of brain width at the level of the diencephalon to establish the effect on brain tissue.

In situ hybridisation, immunohistochemistry, and cartilage staining

Injected embryos were collected at the desired stage and fixed in 4% paraformaldehyde/PBS overnight. Fixed embryos were then stored in 100% methanol, or processed immediately for in situ hybridisation or immunohistochemistry. For in situ hybridisation, probe synthesis, hybridisation, and detection for *cmcl2* was performed as described (Chocron et al.,; Xu et al., 1994). Embryos were then re-fixed in 4% paraformaldehyde, cleared in 70% gycerol/PBS, and mounted for photography. Additional experiments were performed in *cmcl2*::GFP zebrafish (Rottbauer et al., 2006) in which the heart tissue expresses GFP. In this case embryos were fixed and photographed on a Leica M205FA fluorescent stereoscope (Leica Microsystems, UK). For antibody detection of neurons in zebrafish embryos, we used mouse anti-HuC/D (clone 16A11, A-21271, Invitrogen) as described previously (Gerety et al., 2013). Fluorescent images were captured using a Leica TCS SP2 confocal microscope (Leica Microsystems, UK). Cartilage staining was performed as described previously (Clement et al., 2008), and photographed on a Zeiss Axioplan microscope equipped with an Zeiss Axiocam digital camera.

REFERENCES

Andersson, R., et al. (2008). A segmental maximum a posteriori approach to genome-wide copy number profiling. Bioinformatics 24(6): 751-758.

Barnes, C., et al. (2008). A robust statistical method for case-control association testing with copy number variation. Nature genetics 40(10): 1245-1252.

Barry, D., Hartigan, J.A. (1993). A Bayesian Analysis for Change Point Problems. Journal of the American Statistical Association. Mar, 35(3), 309319.

Bellenguez, C., et al. (2012). A robust clustering algorithm for identifying problematic samples in genome-wide association studies. Bioinformatics 28(1): 134-135.

Benelli, M., et al. (2010). A very fast and accurate method for calling aberrations in array-CGH data. Biostatistics 11(3): 515-518.

Bill, B. R., et al. (2009). A primer for morpholino use in zebrafish. Zebrafish 6(1): 69-77.

Chocron, S., et al. (2007). Zebrafish Bmp4 regulates left-right asymmetry at two distinct developmental time points. Developmental biology 305(2): 577-588.

Clement, A., et al. (2008). Regulation of zebrafish skeletogenesis by ext2/dackel and papst1/pinscher. PLoS genetics 4(7): e1000136.

Conrad, D. F., et al. (2010). Origins and functional impact of copy number variation in the human genome. Nature 464(7289): 704-712.

Craddock, N., et al. (2010). Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature 464(7289): 713-720.

Dauber, A., et al. (2013). SCRIB and PUF60 are primary drivers of the multisystemic phenotypes of the 8q24.3 copy-number variant. American journal of human genetics 93(5): 798-811.

de Ligt, J., et al. (2012). Diagnostic exome sequencing in persons with severe intellectual disability. The New England journal of medicine 367(20): 1921-1929.

Erdman, C., Emerson, J.W.. (2007). bcp: A Package for Performing a Bayesian Analysis of Change Point Problems. R package version 1.8.4 URL http://CRAN.R-project.org/.

Fromer, M., et al. (2014). De novo mutations in schizophrenia implicate synaptic networks. Nature 506(7487): 179-184.

Gerety, S. S., Wilkinson, D. G. (2011). Morpholino artifacts provide pitfalls and reveal a novel role for pro-apoptotic genes in hindbrain boundary development. Developmental biology 350(2): 279-289.

Gerety, S. S., et al. (2013). An inducible transgene expression system for zebrafish and chick. Development 140(10): 2235-2243.

Golzio, C., et al. (2012). KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. Nature 485(7398): 363-367.

Gonzalez, J. R., et al. (2011). A fast and accurate method to detect allelic genomic imbalances underlying mosaic rearrangements using SNP array data. BMC bioinformatics 12: 166.

Gribble, S. M., (2013). Massively parallel sequencing reveals the complex structure of an irradiated human chromosome on a mouse background in the Tc1 model of Down syndrome. PloS one 8(4): e60482.

Hindson, B. J., K. D. Ness, et al. (2011). High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Analytical chemistry 83(22): 8604-8610.

Illumina SNP-genotyping protocol, http://supportres.illumina.com/documents/myillumina/0d810920-cda0-439c-95c7-dd5d440f3e13/infinium_hd_ultra_auto_euc_revb_11328108.pdf

Iossifov, I., et al. (2012). De novo gene disruptions in children on the autistic spectrum. Neuron 74(2): 285-299.

Kettleborough, R. N., et al. (2013). A systematic genome-wide analysis of zebrafish protein-coding gene function. Nature 496(7446): 494-497.

King, D., et al. (2014). A novel method for detecting Uniparental Disomy from trio genotypes identifies a significant excess in children with developmental disorders. 10.1101/gr.160465.113

Kohler, S., et al. (2014). The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. Nucleic Acids Res 42(Database issue): D966-974.

Langheinrich, U., et al. (2002). Zebrafish as a model organism for the identification and characterization of drugs and genes affecting p53 signaling. Current biology : CB 12(23): 2023-2028.

Manichaikul, A., et al. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics 26(22): 2867-2873.

McLaren, W., et al. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics 26(16): 2069-2070.

Mills, R. E., et al. (2006). An initial map of insertion and deletion (INDEL) variation in the human genome. Genome research 16(9): 1182-1190.

Neale, B. M., et al. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature 485(7397): 242-245.

O'Roak, B. J., et al. (2012). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature 485(7397): 246-250.

Olshen, A. B., et al. (2004). Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 5(4): 557-572.

Picard, F., et al. (2005). A statistical approach for array CGH data analysis. BMC bioinformatics 6: 27.

Pique-Regi, R., et al. (2010). R-Gada: a fast and flexible pipeline for copy number analysis in association studies. BMC bioinformatics 11: 380.

Price, T. S., et al. (2005). SW-ARRAY: a dynamic programming solution for the identification of copy-number changes in genomic DNA using array comparative genome hybridization data. Nucleic acids research 33(11): 3455-3464.

Purcell, S. M., et al. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. Nature 506(7487): 185-190.

Ramu, A., et al. (2013). DeNovoGear: de novo indel and point mutation discovery and phasing. Nature methods 10(10): 985-987.

Rauch, A., et al. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet 380(9854): 1674-1682.

Robu, M. E., et al. (2007). p53 activation by knockdown technologies. PLoS genetics 3(5): e78.

Rottbauer, W., et al. (2006). Cardiac myosin light chain-2: a novel essential component of thick-myofilament assembly and contractility of the heart. Circulation research 99(3): 323-331.

Samocha, K. E., et al. (2014). A framework for the interpretation of de novo mutation in human disease. Nature genetics. doi:10.1038/ng.3050

Sanders, S. J., et al. (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485(7397): 237-241.

Wang, K., et al. (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome research 17(11): 1665-1674.

Westerfield, M. The Zebrafish Book. (University of Oregon Press, 1993).

Wright CF, et al. (in press) Genetic Diagnosis of Developmental Disorders: scalable analysis of genome-wide research data, Lancet.

Xu, Q., et al. (1994). Spatially regulated expression of three receptor tyrosine kinase genes during gastrulation in the zebrafish. Development 120(2): 287-299.

Zaidi, S., et al. (2013). De novo mutations in histone-modifying genes in congenital heart disease. Nature 498(7453): 220-223.