Supplemental Methods

Patient #1

Conventional karyotyping (46,XX) and an Agilent (Santa Clara, CA, USA) 180K probe array-CGH test returned normal results. We then performed singleton gene panel sequencing using a custom designed Agilent SureSelect enrichment kit for 1292 neurodevelopmental delay-associated genes on an Illumina NextSeq500 (San Diego, CA, USA) Next-Generation-Sequencing (NGS) platform at a mean sequence coverage of 317x and 99.5% of the target region sequenced at ≥10x. Variants were called using the QIAGEN CLC Biomedical Workbench (Hilden, Germany), and annotated and filtered in QIAGEN Ingenuity Variant Analysis. The gene panel dataset was filtered for ultra-rare (allele frequency <0.1%) homozygous variants, possible compound-heterozygous variants and heterozygous variants with a predicted truncating effect on gene function or variants with a previously described disease association as listed in the Human Gene Mutation Database (HGMD Professional 2018.2). In the context of the patient's phenotype, only the putative null variant c.2051_2052dup; p.(Glu685*) in *KMT2E* remained as a plausible disease associated variant. This variant was subsequently Sanger sequenced in the patient and both parents and was confirmed *de novo*. The variant was thus classified as "Likely pathogenic" with the ACMG[1] criteria PS2 (*de novo*), PM2 (absent from controls) and PP3 (deleterious *in silico* prediction). After publication of the case series by O'Donnell-Luria et al.[2], the variant could be reclassified as "Pathogenic" using the additional criterion PVS1 (null variant in established gene with known loss-of-function pathomechanism).

In the remaining genes of the gene panel, no other variants with a clear association to the neurodevelopmental phenotype were detected. NGS-based copy number variation analysis of the gene panel dataset did not reveal any deleterious CNVs.

Patient #6

Singleton exome capture and sequencing were performed by Integragen SA from 1µg of genomic DNA using the TWIG kit. The resulting libraries were sequenced on Illumina NovaSeq6000. Resulting reads were aligned to the human genome reference sequence (GRCh37 assembly) using Burrows-Wheeler aligner (BWA; version 0.7.15). Duplicate reads were marked using Picard MarkDuplicates (version 2.4.1). Aligned reads were then processed using GATK BaseRecalibrator and PrintReads (Genome Analysis Toolkit; version 3.8) to recalibrate base quality scores, following GATK Best Practices recommendations. Quality control was performed on the BAM file by calculating depth of coverage with GATK DepthOfCoverage (mean sequence coverage of 108x and 97.2% of the target region sequenced at ≥10x). SNVs and indels were identified from BAM file using GATK HaplotypeCaller and annotated using SnpEff (version 4.3). CNVs were identified using XHMM. A heterozygous 5bp duplication leading to a frameshift was identified in the *KMT2E* gene and confirmed by Sanger sequencing. The patient's symptomatic father carries this variant.

Patient #7

Trio exome capturing was carried out by using Agilent SureSelect Target Enrichment Clinical Research Exome V2. Sequencing (paired-end 150bp) was performed on the Illumina HiSeq4000 platform. Data was demultiplexed by Illumina Software CASAVA. Reads are mapped to the genome (build hg19/GRCh37) with the program BWA (reference: http://bio-bwa.sourceforge.net/). Variants are detected with the Genome Analysis Toolkit (reference: http://www.broadinstitute.org/gatk/). Subsequently, variants were filtered with the Alissa Interpret software package (Agilent Technologies) and were further selected based on

three inheritance models (de novo, autosomal recessive and X-linked recessive/dominant). In this analysis, a heterozygous *de novo* frameshift variant in *KMT2E* was detected, which we classified as pathogenic in accordance with the ACMG variant classification guidelines.

Patient #8

Clinical Exome Analysis was performed (Parkville, VC, Australia) with a targeted mean coverage of 100x and a minimum of 90% of bases sequenced to at least 15%. Data was processed using Cpipe (Sadedin, S et al. (2015) Genome Medicine 7:68). Annotated variant calls were generated within the target region (coding exons +/- 2bp) using reference genome (GRCh37). Variants were annotated against all gene transcripts, with the reporting of variants against the HGNC recommended transcript. Curation of variants was phenotype-driven with gene lists used for variant prioritization. Regarding the patient's phenotype, two variants were identified as possibly causative of disease: c.554T>C; p.(Leu185Pro) in *KAT6A* and c.2164_2167del; p.(Lys722Valfs*17) in *KMT2E*. The *KAT6A* heterozygous missense variant was predicted to result in a moderate amino acid change from leucine to proline at position 185 of the protein, is not located within a well-established functional domain and has been seen clinically in unaffected individuals. This variant was classified as a "Variant of Uncertain Significance (VUS)." As the patient's clinicians felt this *KAT6A* variant was likely irrelevant, no parental testing was pursued. The *KMT2E* deletion was predicted to create a frameshift starting at amino acid position 722 introducing a stop codon resulting in loss of protein function through nonsense-mediated decay (NMD). Although this variant was not seen before in clinical cases, other variants predicted to cause NMD have been described in *KMT2E*-related neurodevelopmental disorders (ClinVar, [2]). This variant was Sanger sequenced in both parents and confirmed de novo. A new report was issued classifying this variant as "Pathogenic."

Patient #10

Next-generation sequencing (massively parallel sequencing, exome trio sequencing, AgilentSureSelectXT Human All Exon V7. lllumlna^{R} sequencing technology) was carried out on an Illumina NextSeq 500 system (Illumina, San Diego, CA) as 150 bp paired-end runs using v2.0 SBS chemistry. Sequencing reads were aligned to the human reference genome (GRCh37/hg19) using BWA (v0.7. 13-r1126) with standard parameters. Statistics on coverage and sequencing depth on the clinical targeted regions (i.e. RefSeq coding exons and +/-5 intronic region) were calculated with a custom script. SNV and INDEL calling on the nuclear genes was conducted using SAMtools (v1.3.1) with subsequent coverage and quality dependent filter steps. Variant annotation was performed with snpEff (v4.2) and Alamut-Batch (v1.4.4). Only variants (SNVs/small INDELs) in the coding and flanking intronic regions (±50 bp) were evaluated.

The *de novo* variant c.4829dup p.(Leu1610Phefs*259) was classified as "Likely pathogenic" with the ACMG criteria PS2 (confirmed *de novo* occurrence), PVS1 at moderate strength[3] (null variant in established gene with known loss-of-function pathomechanism) and PM2 (absent from controls).

NGS-based copy number variation analysis of the gene panel dataset did not reveal any deleterious CNVs.

Patient #14 and #15

In both siblings, aCGH analyses using CGX™ v1.1 8-plex platform with 60K probes returned normal results. Singleton exome sequencing was performed for patient 14.1 by a commercial laboratory (Genome Diagnostics Nijmegen, NL) in 2017. Exome enrichment (Agilent SureSelectXT Human All Exon 50Mb) and sequencing (Illumina HiSeq) was performed. Variant calling and annotation were performed using the laboratory's in-house developed strategy. A homozygous *GJB2*:NM_004004.5:c.109G>A, p.(Val37Ile)

variant was identified. This is a known pathogenic variant associated with non-syndromic hearing loss[4], explaining the bilateral sensorineural hearing impairment in patient 14.1. No other relevant variant in relation to his clinical presentation was identified. In the meantime, samples of both siblings were also sent to the Broad Institute and included in an exome research project led by the Autism Sequencing Consortium (ASC). Exome enrichment was performed using the Illumina Nextera exome capture kit, with sequencing performed on an Illumina HiSeq sequencer. Raw exome data was retrieved from ASC in 2019 and processed using HKU in-house bioinformatics pipeline based on GATK v3.4 and ANNOVAR. We filtered for variants with gnomAD allele frequency <1% and prioritized reported pathogenic/likely pathogenic variants in the ClinVar database, loss-of-function variants, as well as missense variants with deleterious bioinformatics prediction. Our analysis identified a paternally inherited splice variant *KMT2E*:NM_018682.3:c.768+1G>A in both siblings and was confirmed by Sanger sequencing. Based on the ACMG/AMP 2015 guidelines, it was classified as likely pathogenic with the criteria PVS1 (null variant in a gene in which loss-of-function is a known disease mechanism) and PM2 (variant not found in population database gnomAD). For the other variants in the analysis, none had a clear association with neurodevelopmental disorders.

Patient #16

Singleton whole exome sequencing was performed using an Agilent Clinical Research Exome v2 enrichment kit. Sequencing was done on an Illumina HiSeq4000 platform. A heterozygous canonical splice variant at c.2848-2A>C was identified and confirmed by Sanger sequencing. The patient's mother did not carry this variant, but the father was unavailable for testing.

Patient #17

Conventional karyotyping returned normal results (46,XY). We then performed an Oligo array CGH (44K BlueGnome; Cambridge, United Kingdom) according to manufacturer procedures. Data were analyzed using BlueFuse MULTI v4.1 software (Illumina) and Cartagenia software (Agilent). Assembly hg19 (GRCh37) of the genome was used as a reference. Array CGH analyses revealed a heterozygous deletion on the long arm of chromosome 7 (7q22.3 deletion) spanning 710 kb and including 10 consecutive oligonucleotide probes encompassing the last 3' exons of *KMT2E* (from exon 4 to exon 15; arr[hg19] 7q22.3(104,696,686-105,407,628)x1). Conventional FISH with Bacterial Artificial Chromosome (BAC) probe RP11-195N21 mapping on chromosome 7q22.3 was used to confirm the chromosomal rearrangement detected by array CGH and to perform parental segregation analysis. The microdeletion was confirmed *de novo*.

1. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine : official journal of the American College of Medical Genetics 2015;**17**(5):405-24 doi: 10.1038/gim.2015.30[published Online First: Epub Date]|.
2. O'Donnell-Luria AH, Pais LS, Faundes V, et al. Heterozygous Variants in KMT2E Cause a Spectrum of Neurodevelopmental Disorders and Epilepsy. American journal of human genetics 2019;**104**(6):1210-22 doi: 10.1016/j.ajhg.2019.03.021[published Online First: Epub Date]|.
3. Abou Tayoun AN, Pesaran T, DiStefano MT, et al. Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. Human mutation 2018;**39**(11):1517-24 doi: 10.1002/humu.23626[published Online First: Epub Date]|.

4. Oza AM, DiStefano MT, Hemphill SE, et al. Expert specification of the ACMG/AMP variant interpretation guidelines for genetic hearing loss. Human mutation 2018;**39**(11):1593-613 doi: 10.1002/humu.23630[published Online First: Epub Date]|.