ARTICLE Open Access

Risk gene-set and pathways in 22q11.2 deletion-related schizophrenia: a genealogical molecular approach

Elena Michaelovsky[1,2], Miri Carmel[1,2], Amos Frisch[1,2], Mali Salmon-Divon [3], Metsada Pasmanik-Chor[4], Abraham Weizman[1,2,5,6] and Doron Gothelf[1,6,7]

Abstract

The 22q11.2 deletion is a strong, but insufficient, "first hit" genetic risk factor for schizophrenia (SZ). We attempted to identify "second hits" from the entire genome in a unique multiplex 22q11.2 deletion syndrome (DS) family. Bioinformatic analysis of whole-exome sequencing and comparative-genomic hybridization array identified de novo and inherited, rare and damaging variants, including copy number variations, outside the 22q11.2 region. A specific 22q11.2-haplotype was associated with psychosis. The interaction of the identified "second hits" with the 22q11.2 haploinsufficiency may affect neurodevelopmental processes, including neuron projection, cytoskeleton activity, and histone modification in 22q11.2DS-rrlated psychosis. A larger load of variants, involved in neurodevelopment, in combination with additional molecular events that affect sensory perception, olfactory transduction and G-protein-coupled receptor signaling may account for the development of 22q11.2DS-related SZ. Comprehensive analysis of multiplex families is a promising approach to the elucidation of the molecular pathophysiology of 22q11.2DS-related SZ with potential relevance to treatment.

Introduction

Schizophrenia (SZ) is a disabling mental disorder affecting 1% of the general population with a strong genetic component. Case–control studies of thousands participants have identified many molecular variants carrying genetic risk for SZ[1–3]. The highest known genetic risk for SZ is the 22q11.2 microdeletion[1]. Individuals carrying a hemizygous 22q11.2 deletion, "first hit", have ~40% chance of developing SZ-spectrum disorders in adults[4].

Since not all individuals who carry the 22q11.2 deletion develop SZ, it is plausible that genetic variants outside the 22q11.2 region ("second hits") contribute to the risk to develop SZ[5]. In an attempt to identify genetic risk factors contributing to SZ development, we focused on a large multiplex 22q11.2DS family. Studying multiplex families is a promising approach since the risk variants in the SZ affected individuals are expected to be enriched compared to the non-affected family members. A genealogical analysis, in contrast to case–control study, eliminates the need for population stratification. Moreover, it is a straightforward and less expensive strategy to analyze the association between clinical and genetic data in a particular family under a long-lasting close clinical follow-up.

We took the advantage of a unique multiplex 22q11.2DS family with high load of psychosis expression (Fig. 1) and attempted to identify "second hits" that may contribute to psychosis and SZ developing, using bioinformatic analysis of whole-exome sequencing (WES) and comparative-genomic hybridization (CGH) array.

Correspondence: Elena Michaelovsky (michaelo@tauex.tau.ac.il) or (elenami2@clalit.org.il)
[1]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel
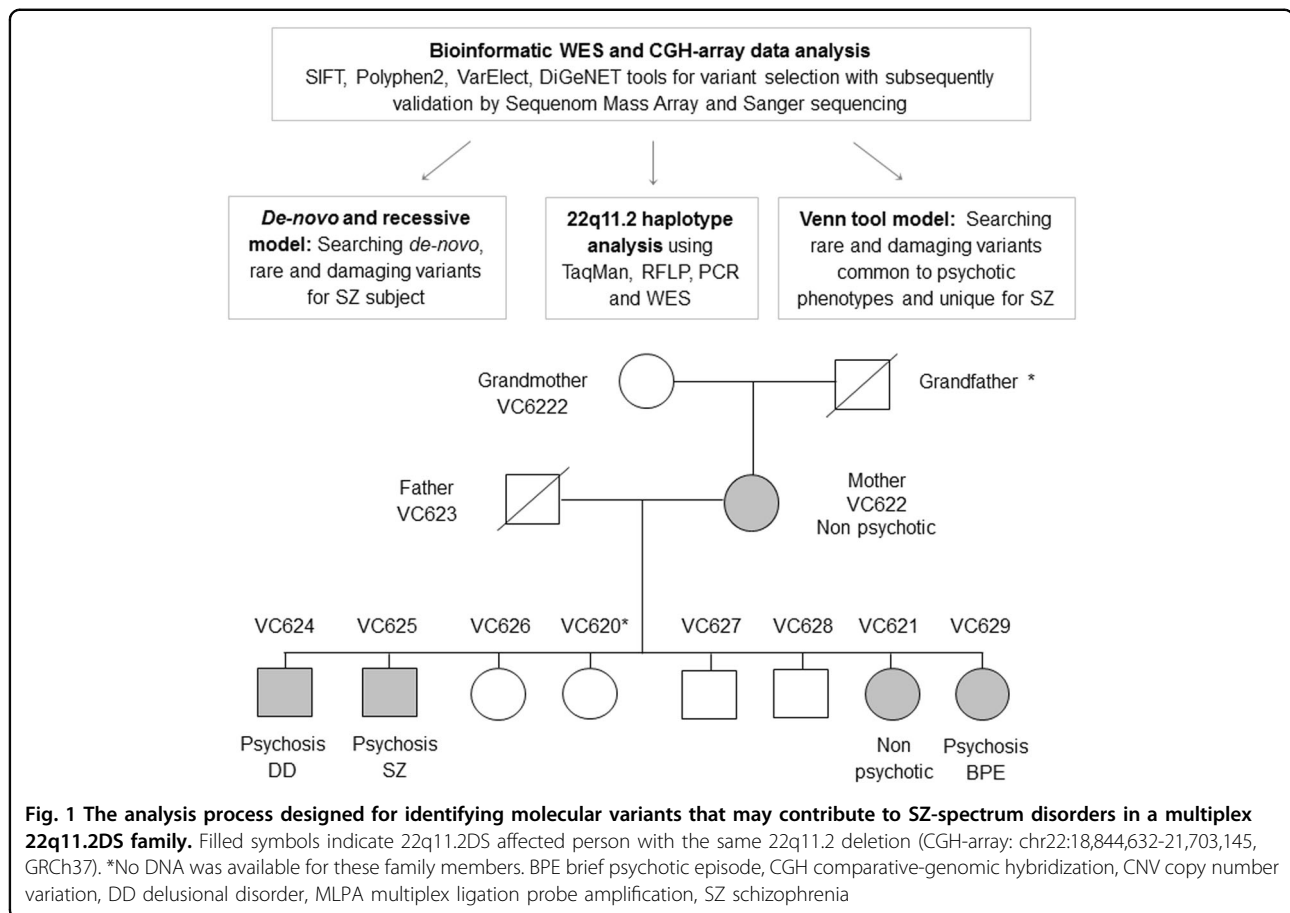[2]Felsenstein Medical Research Center, Petah Tikva, Israel
Full list of author information is available at the end of the article.
These authors contributed equally: Elena Michaelovsky, Miri Carmel

Fig. 1 The analysis process designed for identifying molecular variants that may contribute to SZ-spectrum disorders in a multiplex 22q11.2DS family. Filled symbols indicate 22q11.2DS affected person with the same 22q11.2 deletion (CGH-array: chr22:18,844,632-21,703,145, GRCh37). *No DNA was available for these family members. BPE brief psychotic episode, CGH comparative-genomic hybridization, CNV copy number variation, DD delusional disorder, MLPA multiplex ligation probe amplification, SZ schizophrenia

Materials and methods

Subjects

The multigeneration Jewish family of Moroccan origin, consisted of a grandmother (mother side), parents and 8 children, referred to the Behavioral Neurogenetics Center with suspected 22q11.2DS in one of the children (VC621). Genetic assessment has demonstrated that VC621, her mother and three of her seven siblings are affected with 22q11.2DS (Fig. 1). All family members with 22q11.2DS had psychiatric manifestations with varying degrees of severity, including specific and social anxiety disorder, generalized anxiety disorder (GAD), attention deficit/hyperactivity disorder (ADHD), and SZ-spectrum disorders, including SZ, delusional disorder (DD), and history of brief psychotic episode (BPE) (Supplementary Table 1).

The SZ-VC625 patient was hospitalized at the age of 25 years due to an acute psychotic episode that was triggered by use of 3,4-methylenedioxymethamphetamine (MDMA) and heroin. He was diagnosed with paranoid SZ and substance use disorder. Since then he suffers from chronic psychotic symptoms and marked negative symptoms with poor response to antipsychotics and he is currently living in a sheltered home. VC624 was diagnosed as suffering from a DD at the age 43. VC629 was diagnosed at the age 12 with ADHD, GAD, specific and social anxiety disorders and experienced a BPE at age 14. Unfortunately, no DNA was available for the sibling VC620 (Fig. 1) and therefore she was not included in the molecular analyses.

Psychiatric and cognitive assessments

The participants were interviewed by experienced clinicians using the Structured Clinical Interview for Axis I DSM-IV Disorders (SCID)[6]. The ADHD module from the Schedule for Affective Disorders and Schizophrenia for School-Aged Children, Present and Lifetime[7] was added to the SCID to evaluate the presence of ADHD. Cognitive evaluation was conducted with the age appropriate versions of the Wechsler Intelligence Scale[8,9]. The study protocol was approved by the Sheba Medical Center Review Board. Written informed consent was obtained from all participants after the nature of this study was explained to the subjects.

Molecular testing for 22q11.2 microdeletion

To confirm 22q11.2 microdeletion all family members with suspicious 22q11.2DS phenotype were tested by fluorescence in situ hybridization (FISH), using the LSI TUPLE1 (*HIRA*) or N25 commercial probes (Vysis Inc., Downers Grove, IL, USA), and by multiplex ligation probe amplification (MLPA) using SALSA MLPA P250-A1 DiGeorge kit (MRC-Holland, Amsterdam, Netherlands) analyses. In addition comparative-genomic hybridization (CGH) analysis, using Illumina HumanOmni Express-12 v 1.1 BeadChip (Illumina Inc, San Diego, CA) was conducted for all 10 family members (grandmother, parents and 7 children).

Genotyping in 22q11.2 region

COMT Val158Met polymorphism (rs4680) was genotyped by the C25746809-50 TaqMan kit (Applied Biosystems Incorporated, Foster City, CA) using the ABI 7000 instrument. The genotyping of the seven additional SNPs from *COMT-ARVCF* region (rs2097603, rs4633, rs4818, rs3838146, rs165599, rs2073748, rs2240717) is described in our previous study[10] and the three short tandem repeats (STRs) markers (D22S1648, D22S941, and D22S944) were genotyped by Morrow et al.[11] and Perez et al.[12]. The genotyping of the two *PRODH* functional SNPs rs2008720 and rs4819756 were described in our previous studies[13,14]. The additional nine *PRODH* SNPs from exon 12 (rs16983466, rs2238731, rs2904552, rs2904551, rs3970559, rs2238730, rs2904550, rs2870984, and rs2870983) were performed by PCR amplification of a 452 bp fragment (primers: F-5′gggacagaggttggaggc-3′; R-5′ggacacatgtggctgacaag-3′) followed by sequencing with BigDye Terminators using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, USA).

Whole-exome sequencing (WES) analysis

WES analysis of genomic DNA from 7 family members (parents and five of their children, which four of them are 22q11.2DS affected), was conducted by Edge-Bio Company (EgdeBio, Gaithersburg, MD). Exon capture was performed with the Nimblegen SeqCap EZ Human Exome v3 (Roche NimbleGen, Madison, Wisconsin, USA). Paired end library was constructed using the Illumina TruSeq Sample Prep. Exon libraries were sequenced with the Illumina Hiseq 2000 platform (Illumina, San Diego, California, USA), providing about 100 bp read length for each sample with theoretical average coverage—100×. Variant tools[15] "bwa_-gatk28_hg19" pipeline which followed the "Best GATK (v2.8-1-g932cd3a) practice" was used for read alignment and variant calling. Briefly, sequencing data were aligned to the human reference genome build 37 using the Burrows Wheeler Alignment tool (BWA 0.7.7) with default parameters. Duplicated reads were marked with

Picard tool (http://broadinstitute.github.io/picard/), and local realignment of reads around indels was performed in order to avoid misalignments due to indels. Next, base quality scores were recalibrated in order to get more accurate base qualities, which in turn improve the accuracy of the variant calls. Raw variants were called with the GATK's UnifiedGenotyper following SNP and indels recalibration process which assigned a well-calibrated probability to each variant call. Filtering was done based on this score, and only variant calls which assigned the "PASS" filter (see GATK documentation for more details) were considered in downstream analysis. Variants were annotated against the single nucleotide polymorphism database (dbSNPs v.138), the 1000 Genome database and the RefSeq database.

Bioinformatic databases and tools

The analysis of the WES data started from 1,215,052 variants, including exonic, 5′/3′ UTR, intergenic, intragenic and splice-sites. Filtering for only exonic variations revealed 27,988 variants out of which 13,854 were nonsynonymous.

The validation of WES results for selected candidate variants were performed using SNP genotyping using the Sequenom Mass Array platform[16] and Sanger sequencing. Seven SNPs within the 22q11.2 were validated using TaqMan and RFLP analysis methods.

Predicted level of variant penetrance was obtained using SIFT (http://sift.jcvi.org/)[17] and Polyphen (http://genetics.bwh.harvard.edu/pph2/)[18]. VarElect (http://varelect.genecards.org/)[19], DisGeNET (http://www.disgenet.org/web/DisGeNET/menu)[20], and Schizophrenia Exome Sequencing Genebook[2] were also used to characterize variations. All genomic data for molecular variants in this study were compatible with Genome build GRCh37. NCBI gene database (https://www.ncbi.nlm.nih.gov/gene), Ensembl[21], Exome aggregation consortium (ExAC)[22], Vertebrate Genome Annotation database (Vega; http://vega.sanger.ac.uk/index.html)[23], and UCSC (https://genome.ucsc.edu/)[24] were used for genome builds and general gene and SNP (rs ID) annotations.

Database of genomic variants (DGV; http://dgv.tcag.ca/dgv/app/home)[25] and Integrative Genomics Viewer (IGV; http://software.broadinstitute.org/software/igv/)[26] were used for SNP analysis. Functional and pathway of specific genes resulting from WES and CGH-array were analyzed using DAVID (The Database for Annotation, Visualization and Integrated Discovery) Functional Annotation Bioinformatics Microarray Analysis (https://david.ncifcrf.gov/)[27]. KEGG database was used to investigate biological pathways[28].

Protein–protein interaction analysis of selected genes was performed using STRING database (http://string-db.org/)[29] and GeneMania (http://genemania.org/)[30].

Results

De novo and inherited rare CNVs contribution

CGH-array analysis was performed for searching of de novo and rare CNVs and determination of the deletion borders in 10 family members (Fig. 1) using Illumina HumanOmni Express-12 v 1.1 BeadChip. Fifteen CNVs were detected in this family (Supplementary Table 2). De novo CNVs (17q21.31-gain and 11q14.1-loss) were identified only in the SZ-VC625 individual. Out of the 13 parental CNVs, rare paternal 3p26.3-loss (DGV[25] frequency 0.06%) was transmitted to all 22q11.2DS offspring and another one, rare maternal 14q11.2-gain (DGV frequency 0.09%), to the SZ sibling. Relatively high rates of parental CNVs were found in the SZ-VC625 (5/13) and DD-VC624 (7/13) siblings in comparison to their non-psychotic sister (1/13) (Supplementary Table 2).

22q11.2-haplotype association with psychosis

To examine the association between 22q11.2 intact region and psychosis, we performed haplotype analysis using 73 polymorphisms from the deleted region (chr22:18,900,750-21,350,369; GRCh37): 51 variants, covered by WES and filtered by GATK "PASS" filter (see Materials and methods), and 22 additional variants by TaqMan, PCR, RFLP and Sanger sequencing. The haplotype analysis revealed transmission of a "risk" haplotype from the father to the three psychotic offspring, while the putative "protective" haplotype was transmitted to their non-psychotic daughter. These two haplotypes differed in 42 polymorphic sites out of the 73 (Supplementary Table 3). The "risk" haplotype encompassed two "damaging" variants, *PRODH*-rs2904552/T and *CLTCL1*-rs1061325/ C, that are located on histone binding site, including H3K27me3 and H3K36me3 modifications (http://www. mutationtaster.org/)[31]. These two variants were not presented in the non-psychotic mother (VC622) and daughter (VC621) that carried different haplotypes (Supplementary Table 3).

SZ gene-set and putative pathways suggested by WES and CGH-array bioinformatic analysis

To search for de novo, rare/damaging variants, common to the three psychotic siblings as well as those unique for SZ-VC625 bioinformatic selection analysis (PolyPhen2[18] and SIFT[32] tools) of 1,215,052 WES variants were conducted using recessive, de novo, and Venn tool models. The results of the bioinformatic selection are presented in Supplementary Tables 4 and 5. Following the bioinformatics selection, a further screening was conducted by examining the overlap with gene-sets established by eleven WES[2,33–37] studies in SZ, including 22q11.2DS-related SZ[36], and autism spectrum disorder (ASD)[38–41] (Supplementary Table 6). The putative genetic contributors that were common to the three psychotic

siblings and unique for SZ individual in the studied family, based on WES-bioinformatic selection and CGH-array are presented in a Venn diagram with emphasis on the rare (≤0.05) and de novo variants (Fig. 2).

Candidates common to the psychosis phenotype

CGH-array identified a 3p26.3 rare deletion (DGV)[25] transmitted from the father to the three psychotic siblings. This CNV includes *CNTN4* gene that has a role in formation of axon connections, neuronal network formation/plasticity and synaptogenesis in the developing nervous system (GeneCards: https://www.genecards. org/). In addition, 14 variants with damaging prediction (SIFT and Polyphen2) were selected as common to psychosis phenotype by WES-bioinformatic analysis (Supplementary Table 4). Five of them (36%), located in *ACSF2*, *MPRIP*, *PER3*, *PHC1*, and *RERE* genes, were rare variants (Fig. 2, Supplementary Table 4). *RERE* and *PER3* are located in the same 1p36.23 locus. *RERE* and *PHC1* are transcriptional repressors involved in chromatin organization and histone modifications during embryonic development. *PHC1* is associated with microcephaly and intellectual disability (DisGeNET). *PER3* is also a transcription factor that involved in circadian entrainment. Three keratin intermediate filament candidate genes (*KRT13*, *KRT32*, and *KRT40*, Fig. 2), clustered on 17q21, also related to circadian entrainment superpathway. Interestingly, *ACSF2* that involved in fatty acyl-CoA biosynthesis is located on the same 17q21 locus. Another cytoskeleton-related candidate, *MPRIP*, regulates disassembly of stress fibers in neuronal cells (GeneCards).

Candidates for SZ phenotype and pathway enrichments

In addition to the candidates common for psychosis phenotype, we attempted to identify genetic factors unique to the SZ patient. Using de novo, recessive and Venn tool models additional variants with damaging SIFT/Polyphen2 predictions were selected (Supplementary Table 5). De novo CNVs, as described in the above section, were detected only in the SZ-VC625 sibling. Interestingly, 17q21.31-gain is located on 17q21 region that covered candidate genes common to psychotic siblings, namely, *ACSF2* and keratin genes. In addition, two de novo missense mutations, located in the *DNAH7* and *KRT6C* genes, were detected in the SZ sibling (Supplementary Table 5). All bioinformatically selected genetic contributors, including de novo and heritable, common to psychosis phenotypes and unique to SZ, are presented in the Fig. 2. As presented in this figure, the rate of the heritable rare variants (excluding CNV variants) that were unique for the SZ sibling was 52% (14/27), whereas the rate of the rare variants common to the three psychotic siblings was 36% (5/14).

The 120 candidates selected for the SZ sibling (Supplementary Table 6), including 28 genes from the 22q11.2 deleted region and 82 additional from the entire genome, were used for running the DAVID[27] functional enrichment analysis. The statistically significant DAVID functional enrichments are presented in Fig. 3 (details in Supplementary Table 7).

Three signaling pathways, including transmembrane signaling receptor activity, detection of stimulus involved in sensory perception and olfactory transduction remained significant ($p < 0.04$) following Benjamini correction for multiple testing. The G-protein-coupled receptor (GPCR) signaling and activity also revealed significant results ($p < 0.009$) but did not survive after post hoc correction. Additional significant DAVID clustering identified neurodevelopmental pathways, including cranial nerve development/morphogenesis and neuron projection extension/morphogenesis. It is of note that the majority (~80%) of the genes enriched for sensory perception, olfactory transduction, and GPCR signaling were consisted of genes unique to the SZ patient, while the majority (75–88%) of the genes contributing to neurodevelopmental processes, including neuron projection extension/morphogenesis, were common to the psychosis phenotype (Supplementary Table 7).
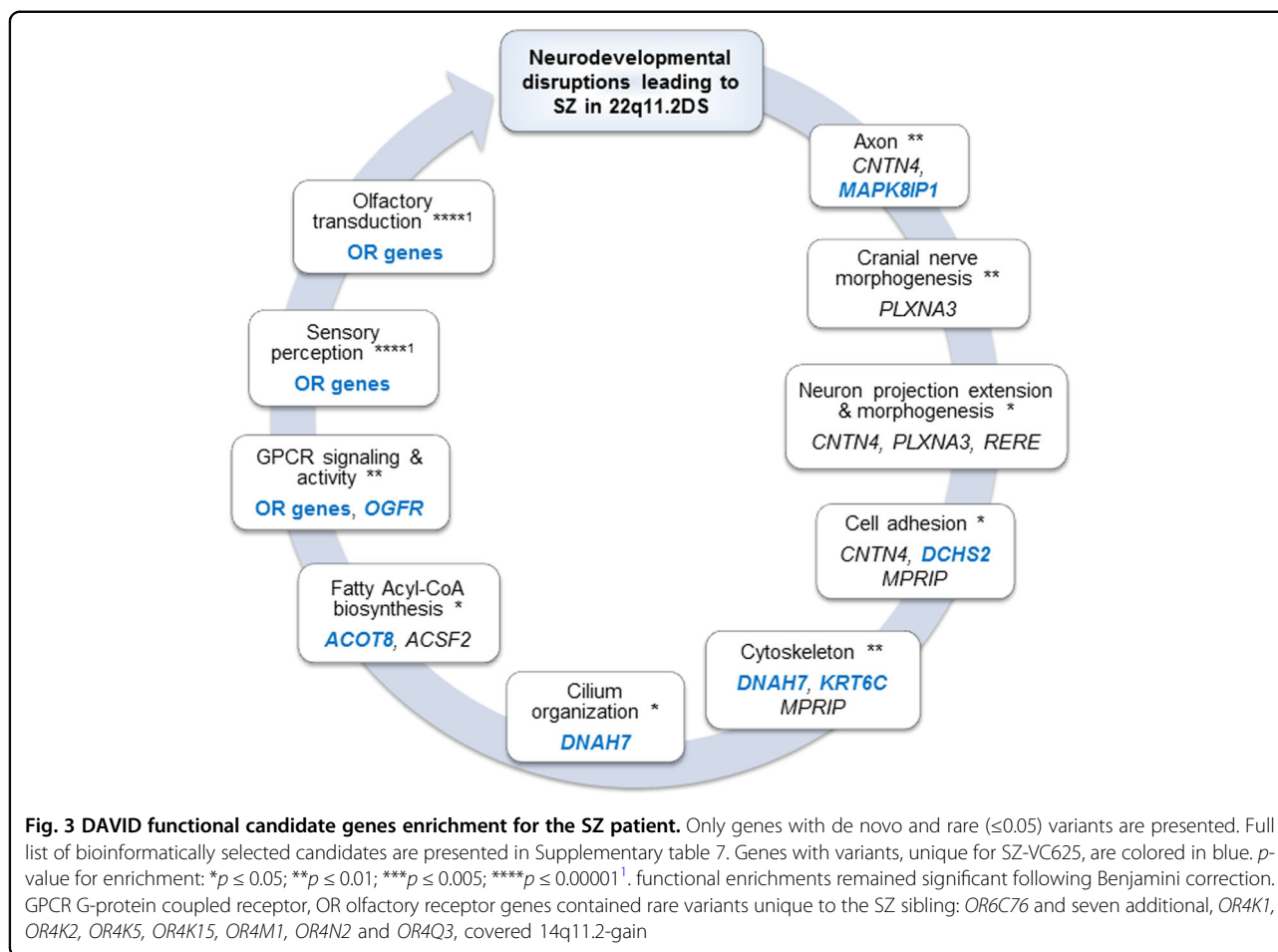
The histone modification processes did not appear to be significantly enriched by DAVID analysis. However, genes involved in these process included de novo (*KANSL1*) and inheritable (*CNTN4*, *RERE*, *NCOA3*, and *PHC1*) rare mutations (Supplementary Table 7). Specifically, *RERE*, *PHC1* and *CNTN4* variants were common to the three psychotic siblings and *KANSL1* and *NCOA3* variants were unique for SZ patient. In addition, fatty acyl-CoA biosynthesis and charge relay system processes, which involved in histone acylation and deacethylation, revealed DAVID statistical significance ($p = 0.02$ and $p = 0.03$, respectively) (Supplementary Table 7).

Discussion

The goal of this study was to search for "second hits" that may contribute to the development of psychosis and SZ in a unique multiplex family with high load of
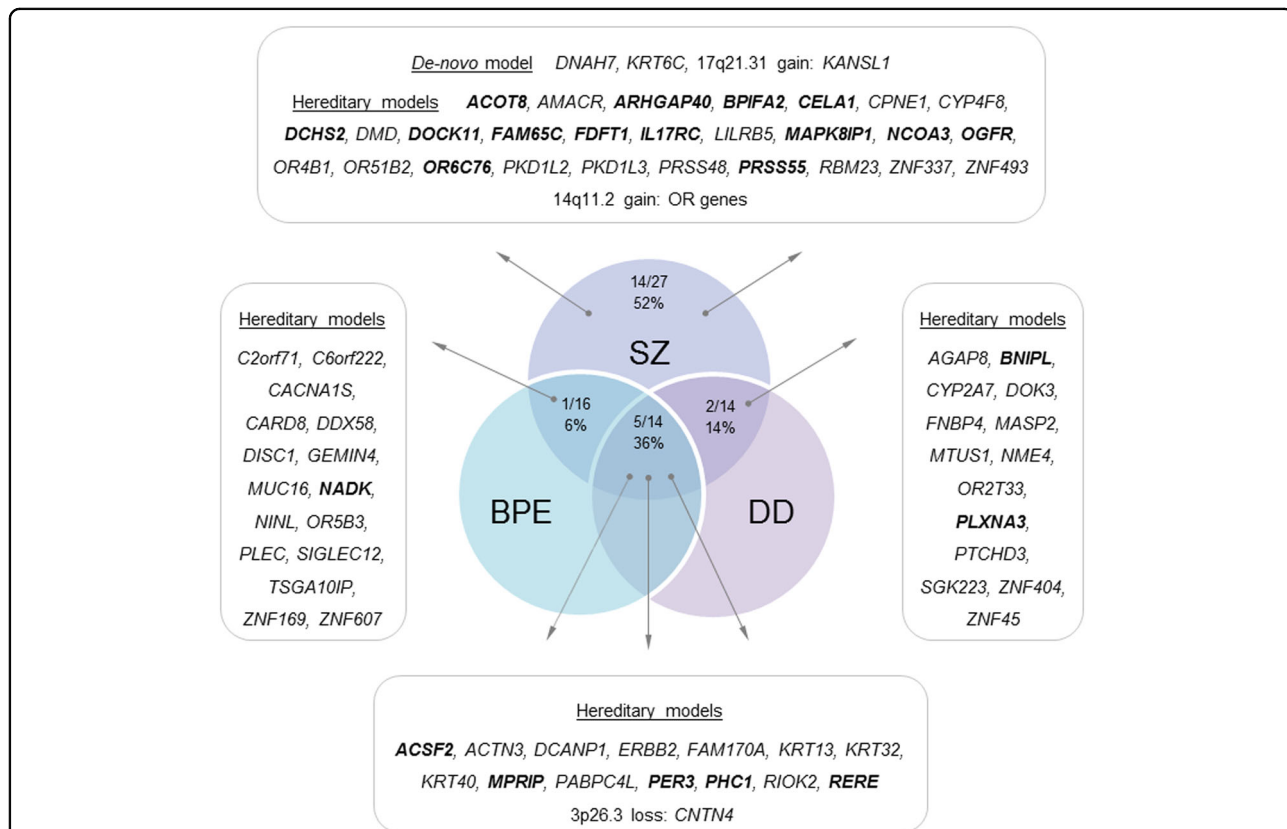
Fig. 2 Genetic contributors to SZ-spectrum disorders in the 22q11.2DS family. Genes with damaging variants that are common to the three psychotic siblings as well as those unique to the SZ sibling are presented. The damaging variants were selected using SIFT and Polyphen2 prediction (Supplementary Table 4 and 5) and overlapping with gene-set of 11 whole-exome sequencing studies in SZ and autism spectrum disorder populations (Supplementary Table 6). Genes with rare (≤0.05) variants, excluding de novo and CNVs, are bolded and presented in rates. BPE brief psychotic episode, DD delusional disorder, OR olfactory receptor genes, *OR4K1, OR4K2, OR4K5, OR4K15, OR4M1, OR4N2,* and *OR4Q3*, covered 14q11.2-gain, SZ schizophrenia

Fig. 3 DAVID functional candidate genes enrichment for the SZ patient. Only genes with de novo and rare (≤0.05) variants are presented. Full list of bioinformatically selected candidates are presented in Supplementary table 7. Genes with variants, unique for SZ-VC625, are colored in blue. *p*-value for enrichment: *$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.005$; ****$p \leq 0.00001$[1]. functional enrichments remained significant following Benjamini correction. GPCR G-protein coupled receptor, OR olfactory receptor genes contained rare variants unique to the SZ sibling: *OR6C76* and seven additional, *OR4K1, OR4K2, OR4K5, OR4K15, OR4M1, OR4N2* and *OR4Q3*, covered 14q11.2-gain

psychotic expression affected by the inherited 22q11.2 deletion.

Candidates common to psychosis phenotype

Firstly, the putative contribution of the 22q11.2 intact chromosome to psychosis phenotype was examined. In the analyzed pedigree, one of the interesting finding is that the same haplotype in the 22q11.2 region is transmitted by the unaffected father to each of the three psychotic offspring. The corresponding 22q11.2-haplotype analysis revealed an association between a specific "risk" haplotype, contained two "damaging" variants in *PRODH* and *CLTCL1* genes, and psychosis. The identified rare and damaging *PRODH*-rs2904552 variant as well as *CLTCL1* de novo mutation were previously reported to be associated with SZ[42,43]. The *CLTCL1* has a role in synaptic plasticity (KEGG:hsa04721), neural crest development[44], neuronal depolarization[45], and histone modification[46]. Two previous case–control studies reported that variants on the 22q11.2 intact chromosome are not major contributors to SZ[36] and psychosis[47]. The haplotype analysis of large multiplex families, in contrast to case–control studies, enables to examine contribution of the 22q11.2

gene variants, inherited together as a haplotype, to psychosis using internal control for comparison that eliminates the need for population stratification.

When a CNV is not de novo, the pedigree offers an opportunity to identify other genetic factors than the 22q11.2-loss, "second hits", contributing to the transmission of psychosis. Bioinformatic analysis of WES and CGH-array data selected 15 common for psychosis phenotype damaging variants and 6 of them were rare mutations, including paternal 3p26.3-loss and 5 additional missense mutations located in *ACSF2, MPRIP, PHC1, PER3*, and *RERE* genes. It is of note that the identified candidates were previously reported to be involved in neurodevelopmental processes and to be associated with psychosis and SZ. The 3p26.3-loss, that covered *CNTN4*, was previously reported to be associated with psychosis in 22q11.2DS[48]. *CNTN4* gene is involved in synaptic plasticity[49], prefrontal neuron-specific histone methylation[50], response to antipsychotics[51], and was implicated in both SZ[52] and autism[53]. *RERE* and *PER3* are located on the 1p36.23, one of the reported 108 SZ-associated loci[3]. *RERE* and *PHC1* are transcriptional repressors involved in chromatin organization and histone modifications during

embryonic development. PER3 is a transcriptional factor that together with the three keratin genes, established as candidates to psychosis in the current study, is related to circadian entrainment superpathway, recently reported as an overlapped superpathway for SZ, ASD, and bipolar disorder (BD)[54]. Interestingly, four candidates for psychosis (three keratin genes and *ACSF2*), identified in the current study, are located on 17q21 locus that was suggested to be associated with SZ in a linkage study that analyzed 175 families with at least two siblings suffering from SZ/SZ-affective disorders[55].

Candidates unique to SZ phenotype and functional enrichment

The next step of the study focused on search for risk gene-set and putative pathways that may contribute to the development of SZ. The detection of de novo mutations in the SZ sibling including CNVs (17q21.31-gain and 11q14.1-loss) and missense mutations (*DNAH7* and *KRT6C*) supports a de novo mutational paradigm for SZ[37,56,57]. The 17q21.31-gain may be a relevant SZ risk factor since it contains the neuronal chromatin modifier *KANSL1* gene[58] and causes microduplication syndrome expressed in psychomotor retardation, poor social interaction and intellectual disability[59]. In addition, 17q21.31-gain was established in childhood onset SZ[60]. Notably, the 17q21 region was demonstrated in the present study to be associated as a risk region common to the three psychotic siblings and as discussed above was previously reported to be associated with SZ[55]. The 11q14.1-loss did not contain coding genes, however in 1.6 Mb upstream to this loss a rare deletion was previously reported in a patient with sporadic SZ[57], and the same region was found to be highly associated (LOD = 3.44) with BD in a linkage study[61].

DAVID[27] functional enrichment analysis that was performed for bioinformatically selected SZ gene-set, identified several statistically significant biological pathways and processes. Sensory perception and neuron projection extension/morphogenesis established in the current study are consistent with previously reported findings in 22q11.2DS-related SZ[36]. The GPCR signaling and olfactory transduction pathways, revealed in the current study, were previously reported to be involved in idiopathic[37,62] and early-onset[63] SZ, respectively.

Histone modification process, implicated in SZ in a large (60,000 participants) GWAS study[64] did not appear to be significantly enriched by DAVID analysis, however, the candidate genes involved in this process were affected by rare de novo and inherited variants. These variants include CNVs covered two histone modification genes, *KANSL1*[58,65] (17q21.31-gain) and *CNTN4*[50] (3p26.3-loss), and missense mutations in *RERE*, *NCOA3* and *PHC1* genes. Moreover, "fatty acyl-CoA biosynthesis" and "charge relay system" processes, which were previously reported to be involved in histone acylation and deacethylation[66,67], exhibited DAVID significant results.

The major weakness of this study is the small sample size of the pedigree, which limits the potential of reaching statistical significance. However, the studied pedigree is unique multiplex 22q11.2DS family with high load of psychosis expression comprised of eight offspring which half of them inherited the maternal deletion. In addition, the fact that one of the three siblings was diagnosed with SZ enabled us to attempt to identify both, variants common to the psychosis phenotype as well as variants that are unique to the SZ individual.

Thus, it appears that the strength of our pedigree design is that all cases share the same genetic causes identical by descent (IBD) that enabled the identification of molecular events relevant to the psychosis phenotype. Furthermore, the value of the pedigree is the exclusion of events not shared by all the affected offspring.

We suggest that specific molecular events, CNVs and missense mutations, outside the 22q11.2, account for disruptions in neurodevelopmental processes, including neuron projection extension and morphogenesis, cytoskeleton signaling and histone modification, that may contribute to the psychosis phenotype in 22q11.2DS. Apparently, the combination of a larger load of damaging molecular events with greater harmful effects on neurodevelopmental processes, together with additional effects on sensory perception, olfactory transduction and GPCR signaling may explain development of SZ in 22q11.2DS. A future study including large 22q11.2DS families with SZ and non-SZ psychosis, as well as a comparison sample of large families with non-22q11.2DS SZ may help in establishing biological pathways relevant to the pathophysiology of SZ and may lead to the development of novel treatments.

Author details
[1]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. [2]Felsenstein Medical Research Center, Petah Tikva, Israel. [3]Department of Molecular Biology, Ariel University, Ariel, Israel. [4]Bioinformatics Unit, G.S. Wise Faculty of Life Science, Tel Aviv University, Tel Aviv, Israel. [5]Geha Mental Health Center, Petah Tikva, Israel. [6]Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel. [7]The Behavioral Neurogenetics Center, Sheba Medical Center, Tel Hashomer, Israel

Authors' contributions
E.M., M.C. and A.W. designed the project, carried out the molecular analyses, and prepared the first draft of the manuscript; D.G. performed the clinical characterization of the participants and collected the blood samples; M.P.-C. and M.S.-D. performed the bioinformatic analyses; E.M., M.C., A.F., A.W. and D.G. conceived the research approach and managed the preparation of the final draft.

Conflict of interest

The authors declare that they have no conflict of interest.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Marshall, C. R. et al. Contribution of copy number variants to schizophrenia from a genome-wide study of 41,321 subjects. *Nat. Genet.* **49**, 27–35 (2017).
2. Purcell, S. M. et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **506**, 185–190 (2014).
3. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
4. Schneider, M. et al. Psychiatric disorders from childhood to adulthood in 22q11.2 deletion syndrome: results from the International Consortium on Brain and Behavior in 22q11.2 Deletion Syndrome. *Am. J. Psychiatry* **171**, 627–639 (2014).
5. Bassett, A. S. et al. Rare genome-wide copy number variation and expression of schizophrenia in 22q11.2 deletion syndrome. *Am. J. Psychiatry* **174**, 1054–1063 (2017).
6. Kaufman, J. et al. Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *J. Am. Acad. Child Adolesc. Psychiatry* **36**, 980–988 (1997).
7. First, M. B., Spitzer, R. L., Gibbon, M. & Williams, J. B. W. *User's Guide for the Structured Clinical Interview for DSM-IV Axis I Disorders SCID-I: Clinician Version* (American Psychiatric Press, Washington, D.C., 1997).
8. Wechsler, D. & Corporation, P. *WISC-III: Wechsler Intelligence Scale for Children: Manual*, (Psychological Corporation, San Antonio, TX, 1991).
9. Wechsler, D. *WAIS-III, Wechsler Adult Intelligence Scale: Administration and Scoring Manual*, (Psychological Corporation, San Antonio, TX, 1997).
10. Michaelovsky, E. et al. Association between a common haplotype in the COMT gene region and psychiatric disorders in individuals with 22q11.2DS. *Int. J. Neuropsychopharmacol.* **11**, 351–363 (2008).
11. Morrow, B. et al. Molecular definition of the 22q11 deletions in velo-cardio-facial syndrome. *Am. J. Hum. Genet.* **56**, 1391–1403 (1995).
12. Perez, E. & Sullivan, K. E. Chromosome 22q11.2 deletion syndrome (DiGeorge and velocardiofacial syndromes). *Curr. Opin. Pediatr.* **14**, 678–683 (2002).
13. Carmel, M. et al. Association of COMT and PRODH gene variants with intelligence quotient (IQ) and executive functions in 22q11.2DS subjects. *J. Psychiatr. Res.* **56**, 28–35 (2014).
14. Zarchi, O. et al. Schizophrenia-like neurophysiological abnormalities in 22q11.2 deletion syndrome and their association to COMT and PRODH genotypes. *J. Psychiatr. Res.* **47**, 1623–1629 (2013).
15. Wang, G. T., Peng, B. & Leal, S. M. Variant association tools for quality control and analysis of large-scale sequence and genotyping array data. *Am. J. Hum. Genet.* **94**, 770–783 (2014).
16. Gabriel, S., Ziaugra, L. & Tabbaa, D. SNP genotyping using the Sequenom MassARRAY iPLEXplatform. *Curr. Protoc. Hum. Genet.* Chapter 2, Unit 2 (2009).
17. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding nonsynonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
18. Adzhubei, I. A. et al. A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
19. Stelzer, G. et al. VarElect: the phenotype-based variation prioritizer of the GeneCards Suite. *BMC Genom.* **17**(Suppl 2), 444 (2016).
20. Pinero, J. et al. DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database* **2015**, bav028 (2015).
21. Hubbard, T. et al. The Ensembl genome database project. *Nucleic Acids Res.* **30**, 38–41 (2002).
22. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).
23. Ashurst, J. L. et al. The vertebrate genome annotation (Vega) database. *Nucleic Acids Res.* **33**, D459–D465 (2005).
24. Karolchik, D., Hinrichs, A. S. & Kent, W. J. The UCSC genome browser. *Curr. Protoc. Bioinformatics* Chapter 1, Unit 14 (2009).
25. MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L. & Scherer, S. W. The database of genomic variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res.* **42**, D986–D992 (2014).
26. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
27. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
28. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).
29. Szklarczyk, D. et al. STRINGv10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447–D452 (2015).
30. Warde-Farley, D. et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* **38**, W214–W220 (2010).
31. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* **11**, 361–362 (2014).
32. Sim, N. L. et al. SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* **40**, W452–W457 (2012).
33. Fromer, M. et al. De novo mutations in schizophrenia implicate synaptic networks. *Nature* **506**, 179–184 (2014).
34. Guipponi, M. et al. Correction: exome sequencing in 53 sporadic cases of schizophrenia identifies 18 putative candidate genes. *PLoS ONE* **10**, e0141630 (2015).
35. McCarthy, S. E. et al. De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. *Mol. Psychiatry* **19**, 652–658 (2014).
36. Merico, D. et al. Whole-genome sequencing suggests schizophrenia risk mechanisms in humans with 22q11.2 deletion syndrome. *G3* **5**, 2453–2461 (2015).
37. Xu, B. et al. Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nat. Genet.* **43**, 864–868 (2011).
38. Iossifov, I. et al. De novo gene disruptions in children on the autistic spectrum. *Neuron* **74**, 285–299 (2012).
39. Neale, B. M. et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **485**, 242–245 (2012).
40. O'Roak, B. J. et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485**, 246–250 (2012).
41. Sanders, S. J. et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237–241 (2012).
42. Gulsuner, S. et al. Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* **154**, 518–529 (2013).
43. Jacquet, H. et al. PRODH mutations and hyperprolinemia in a subset of schizophrenic patients. *Hum. Mol. Genet.* **11**, 2243–2249 (2002).
44. Nahorski, M. S. et al. A novel disorder reveals clathrin heavy chain-22 is essential for human pain and touch development. *Brain* **138**, 2147–2160 (2015).
45. Chahrour, M. H. et al. Whole-exome sequencing and homozygosity analysis implicate depolarization-regulated neuronal genes in autism. *PLoS Genet.* **8**, e1002635 (2012).
46. Weinstein, J. S. et al. Global transcriptome analysis and enhancer landscape of human primary T follicular helper and T effector lymphocytes. *Blood* **124**, 3719–3729 (2014).
47. Guipponi, M. et al. No evidence for the presence of genetic variants predisposing to psychotic disorders on the non-deleted 22q11.2 allele of VCFS patients. *Transl. Psychiatry* **7**, e1039 (2017).
48. Williams, H. J. et al. Schizophrenia two-hit hypothesis in velo-cardio facial syndrome. *Am. J. Med. Genet. B* **162B**, 177–182 (2013).
49. Landek-Salgado, M. A., Faust, T. E. & Sawa, A. Molecular substrates of schizophrenia: homeostatic signaling to connectivity. *Mol. Psychiatry* **21**, 10–28 (2016).

50. Shulha, H. P. et al. Human-specific histone methylation signatures at transcription start sites in prefrontal neurons. *PLoS Biol.* **10**, e1001427 (2012).
51. Liu, Q., Jamba, M., Patrick, C. 3rd, Padmanabhan, S. & Brennan, M. D. Targeted pharmacogenetic analysis of antipsychotic response in the CATIE study. *Pharmacogenomics* **13**, 1227–1237 (2012).
52. Zhao, Z. et al. Association study of 167 candidate genes for schizophrenia selected by a multi-domain evidence-based prioritization algorithm and neurodevelopmental hypothesis. *PLoS ONE* **8**, e67776 (2013).
53. Glessner, J. T. et al. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **459**, 569–573 (2009).
54. Khanzada, N. S., Butler, M. G. & Manzardo, A. M. Gene analytics pathway analysis and genetic overlap among autism spectrum disorder, bipolar disorder and schizophrenia. *Int. J. Mol. Sci.* **18**, E527 (2017).
55. Escamilla, M. et al. A schizophrenia gene locus on chromosome 17q21 in a new set of families of Mexican and central american ancestry: evidence from the NIMH Genetics of schizophrenia in latino populations study. *Am. J. Psychiatry* **166**, 442–449 (2009).
56. Girard, S. L. et al. Increased exonic de novo mutation rate in individuals with schizophrenia. *Nat. Genet.* **43**, 860–863 (2011).
57. Xu, B. et al. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat. Genet.* **40**, 880–885 (2008).
58. Koolen, D. A. et al. Mutations in the chromatin modifier gene KANSL1 cause the 17q21.31 microdeletion syndrome. *Nat. Genet.* **44**, 639–641 (2012).
59. Natacci, F. et al. Chromosome 17q21.31 duplication syndrome: description of a new familiar case and further delineation of the clinical spectrum. *Eur. J. Paediatr. Neurol.* **20**, 183–187 (2016).
60. Ahn, K. et al. High rate of disease-related copy number variations in childhood onset schizophrenia. *Mol. Psychiatry* **19**, 568–572 (2014).
61. Fanous, A. H. et al. Genetic overlap of schizophrenia and bipolar disorder in a high-density linkage survey in the Portuguese Island population. *Am. J. Med. Genet. B* **159B**, 383–391 (2012).
62. Kim, M. J. et al. Functional analysis of rare variants found in schizophrenia implicates a critical role for GIT1-PAK3 signaling in neuroplasticity. *Mol. Psychiatry* **22**, 417–429 (2017).
63. Xu, Y. et al. Altered expression of mRNA profiles in blood of early-onset schizophrenia. *Sci. Rep.* **6**, 16767 (2016).
64. Network & Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat. Neurosci.* **18**, 199–209 (2015).
65. Meunier, S. et al. An epigenetic regulator emerges as microtubule minus-end binding and stabilizing factor in mitosis. *Nat. Commun.* **6**, 7889 (2015).
66. Finnin, M. S. et al. Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. *Nature* **401**, 188–193 (1999).
67. Sabari, B. R., Zhang, D., Allis, C. D. & Zhao, Y. Metabolic regulation of gene expression through histone acylations. *Nat. Rev. Mol. Cell Biol.* **18**, 90–101 (2017).