



EMORY
LIBRARIES &
INFORMATION
TECHNOLOGY

OpenEmory

Identification of 22q13 genes most likely to contribute to Phelan McDermid syndrome

Andrew R. Mitz, *National Institute of Mental Health*

Travis J. Philyaw, *University of Utah*

Luigi Boccutto, *Greenwood Genetic Center*

Aleksandr Shcheglovitov, *University of Utah*

Sara M. Sarasua, *Care Coordination Institute*

[Walter E. Kaufmann](#), *Emory University*

Audrey Thurm, *National Institute of Mental Health*

Journal Title: European Journal of Human Genetics

Volume: Volume 26, Number 3

Publisher: Springer Nature [academic journals on nature.com]: Hybrid Journals | 2018-01-22, Pages 293-302

Type of Work: Article | Final Publisher PDF

Publisher DOI: 10.1038/s41431-017-0042-x

Permanent URL: <https://pid.emory.edu/ark:/25593/t0jp7>

Final published version: <http://dx.doi.org/10.1038/s41431-017-0042-x>

Copyright information:

© 2018 The author.

This is an Open Access work distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Accessed December 11, 2019 1:02 AM EST



Identification of 22q13 genes most likely to contribute to Phelan McDermid syndrome

Andrew R. Mitz¹ · Travis J. Philyaw² · Luigi Boccutto^{3,4} · Aleksandr Shcheglovitov² · Sara M. Sarasua⁵ · Walter E. Kaufmann^{3,4,6,7} · Audrey Thurm⁸

Received: 1 June 2017 / Revised: 4 September 2017 / Accepted: 31 October 2017 / Published online: 22 January 2018
© The Author(s) 2018. This article is published with open access

Abstract

Chromosome 22q13.3 deletion (Phelan McDermid) syndrome (PMS) is a rare genetic neurodevelopmental disorder resulting from deletions or other genetic variants on distal 22q. Pathological variants of the *SHANK3* gene have been identified, but terminal chromosomal deletions including *SHANK3* are most common. Terminal deletions disrupt up to 108 protein-coding genes. The impact of these losses is highly variable and includes both significantly impairing neurodevelopmental and somatic manifestations. The current review combines two metrics, prevalence of gene loss and predicted loss pathogenicity, to identify likely contributors to phenotypic expression. These genes are grouped according to function as follows: molecular signaling at glutamate synapses, phenotypes involving neuropsychiatric disorders, involvement in multicellular organization, cerebellar development and functioning, and mitochondrial. The likely most impactful genes are reviewed to provide information for future clinical and translational investigations.

Introduction

22q13.3 deletion syndrome, also called Phelan–McDermid syndrome (PMS or PHMDS, OMIM 606232) is a

neurodevelopmental disorder-associated syndrome caused by heterozygous contiguous gene deletions at chromosome 22q13 or by pathological variants of *SHANK3* [1]. The clinical PMS phenotype can vary widely and lacks a distinguishing feature. PMS is characterized by intellectual disability (global developmental delays), loss or absence of speech, neonatal hypotonia, autism spectrum symptoms, mild to severe motor impairments, and minor dysmorphic traits [2], resulting in significant, life-long care needs. The overwhelming majority of cases identified thus far involve *SHANK3*, a distal gene of 22q13.33 [2–9]. Further, *SHANK3* variants have been identified in individuals with a full range of phenotypical severity [7, 10–13]. In general, *SHANK3* variants have been associated with cases of intellectual disability (ID) at a rate of about 2% [14], autism spectrum disorder (ASD) at a rate of about 0.7% [14] and schizophrenia at a rate of 0.6–2.16% [15, 16], making it an important target for possible pharmacological intervention in these major neuropsychiatric and neurodevelopmental disorders.

The genetic landscape of PMS is large. Terminal deletions can vary in size from roughly 0.2 to 9.2 Mbp, impacting up to 108 separate protein-coding genes (PMS-associated genes, “PMS genes”). The 108 PMS genes extend from 22q13.2 to 22q13.33 [5, 7, 10, 11, 17–19].

Electronic supplementary material The online version of this article (<https://doi.org/10.1038/s41431-017-0042-x>) contains supplementary material, which is available to authorized users.

✉ Andrew R. Mitz
arm@nih.gov

¹ Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA

² Department of Neurobiology and Anatomy, University of Utah, Salt Lake City, UT, USA

³ Greenwood Genetic Center, Greenwood, SC, USA

⁴ Clemson University School of Health Research, Clemson, SC, USA

⁵ Care Coordination Institute, Greenwood, SC, USA

⁶ Department of Neurology, Boston Children’s Hospital, Boston, MA, USA

⁷ Department of Human Genetics, Emory University School of Medicine, Atlanta, GA, USA

⁸ Neurodevelopmental and Behavioral Phenotyping Service, Office of the Clinical Director, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA

Studying the influences of genes other than *SHANK3* has been difficult because terminal deletions that include *SHANK3* comprise the bulk of reported cases. Some individuals with interstitial deletions, who retain both copies (presumed without deleterious variants) of *SHANK3*, can have phenotypes indistinguishable from individuals with *SHANK3* involvement [5, 10, 17, 18], raising the possibility that much of what is seen in the broader PMS population may be under the influence of many genes of 22q13.3. Further, cohort studies have consistently demonstrated deletion size effects on both core features and other manifestations of PMS [9, 11, 20–22].

The large and, as yet, poorly documented range of genetic contributions to PMS impacts the search for therapeutic interventions. Individuals with PMS provide the only available cohorts for *SHANK3*-targeted clinical studies of ASD [23–25]. Cohorts are assembled from fewer than 1600 cases worldwide [26]. The majority of this population have terminal deletions with no evidence for a preferred breakpoint [7, 10]. A possible recurrent breakpoint of ~150 bp in *SHANK3* was postulated based on three cases with similar breakpoints [27], however, these authors later reported no evidence of preferred deletion size for terminal deletions in 30 subjects [11]. Calculating the mean and median terminal deletion sizes from the graphical data in Sarasua et al. [10] (5.20 and 4.77 Mbp, respectively) and tabular data in Soorya et al. [7] (4.27 and 4.28 Mbp, respectively), it was determined that the mean and median deletion sizes are similar both within studies and between studies as would be expected for a uniform distribution. If the average terminal deletion size is taken as 4.5 Mbp, that value represents a loss of just under half of the 108 PMS genes (Supplementary Figure S1). Therefore, PMS commonly entails the loss of many genes. The frequency of loss for individual genes is quantified in Methods.

In the present study, we describe multiple genes near *SHANK3* that are highly expressed in the brain and have a very high probability of “loss of function intolerance” as described in the bioinformatics study by Lek et al. [28, 29]. Loss of function intolerance is based on genetic variations in the general population that reflect locations on the genome where variation is likely to impact normal gene function in a way that negatively influences the individual’s wellbeing. Based on these probabilities, we identify both well-characterized and poorly-characterized PMS genes that may contribute to clinically relevant manifestations of the syndrome. We expect these results to guide validating functional studies, aid interpretation of *SHANK3*-targeted clinical studies and provide data for identification of new therapeutic targets for improving functioning of individuals with PMS.

Materials and methods

Methods for gene function discovery and likelihood estimation of gene contributors to PMS

Identifying potential contributors to features associated with PMS involves two processes that are not mutually exclusive: (1) matching up gene function with PMS phenotype (function discovery) and (2) estimating the probability that a gene or a group of related genes contribute to the disorder (likelihood estimation). Figure 1 shows the decision pipeline.

Function discovery

We applied several methods for function discovery. The Online Mendelian Inheritance in Man (OMIM) catalog of genetic disorders [30] was searched to identify which PMS genes might be associated with reported phenotype characteristics of PMS. Genes associated with recessive inheritance were included under the assumption that they may contribute to the PMS phenotype with only single copy loss when these genes have a high likelihood of contribution [29].

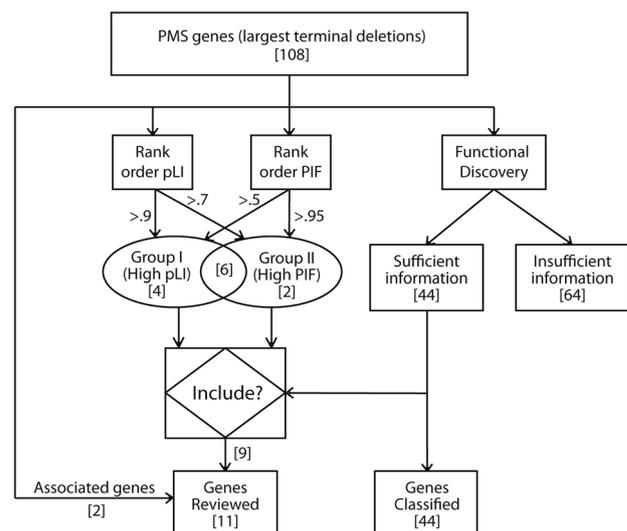


Fig. 1 Gene selection and classification pipeline. PMS genes were selected for classification through a “functional discovery” literature review, as described in Materials and methods. PMS genes were independently rank ordered based on “probability of loss of function intolerance” (pLI) scores from the ExAC database (<http://exac.broadinstitute.org/>) and the population impact factor (PIF), as described in Materials and methods. Two overlapping groups were created: Group I (High pLI, pLI > 0.9 and PIF > 0.5) and Group II (High PIF, pLI > 0.7 and PIF > 0.95). Genes in either group with sufficient information were included in the gene reviews along with incidentally associated genes (see Table 1). Numbers in brackets are number of genes. Groups I and II share six genes

Our second approach for function discovery was to survey existing literature on individual genes, including cellular expression, regional expression, metabolic pathways, disease association, animal models and in vitro models. This survey began with four widely used online resources: PubMed [31], Ensembl [32], Gene Ontology Consortium (GO) [33] and GeneCards [34]. The evidence for functional association is discussed separately for each PMS gene that is reviewed here. We considered any manifestation that has been reported with a frequency $\geq 20\%$ in at least one of the following studies of 10 or more individuals [6–11, 35–37]. For the purpose of function discovery, we assume that loss of one or more PMS genes drives that trait. Poorly characterized genes (unknown function) are treated as their own “phenotype” class.

We used curated databases and genome-wide association studies (GWAS) for both functional discovery and likelihood of contribution of neuropsychiatric-neurodevelopmental genes by locating PMS genes found in the following sources: Class I–III Genes from ASD quartets [38], TADA-65 genes [39], Category S and Category 1–3 SFARI Gene as of March 2017 [40]. We did not explicitly include genes from the Autism Speaks MSSING project [41], because our likelihood estimation criteria are a superset of that project's initial screening criteria.

Likelihood estimation

Likelihood estimation is based on the probability of loss of function (LoF) intolerance (pLI) score from the aggregation of over 60,000 genomes provided by the ExAC project [28, 29]. pLI was computationally determined from high-quality exome sequencing of adults with no reported major childhood/developmental disorders. The full database was used to create a reference sequence for each gene, then variants from this reference were counted for each gene in each individual. Conceptually, impactful genes that are sensitive to variants will be selected out of a population, leaving fewer variants than expected by chance [42]. Lek et al. [29] developed a model that probabilistically assigns to each gene the likelihood that it should be classified into these categories: not sensitive to variants, haplosufficient and haploinsufficient, based on the difference between expected number of variants and observed number of variants. The degree to which each gene might belong in one, two, or all three of these categories, was fit with empirically derived frequencies of haplosufficient and haploinsufficient genes using Bayesian maximum-likelihood estimation. The pLI is calculated by normalizing the probability that a gene is haploinsufficient relative to the three possible categories. Thus, pLI is intolerance for either heterozygous or homozygous LoF. A pLI value of 1 is a high likelihood of intolerance (predicted fully haploinsufficient gene). We

used a pLI > 0.9 for LoF variants as the threshold for likely candidates except as noted. In some cases, likelihood is discussed based on other evidence, which will be described in detail.

Plotting of gene positions

To capture the probable impact of deletion size on phenotype, we plot gene candidates as a function of chromosome position. Gene loci are plotted based upon the most distal exon of the most common isoform. Two-dimensional plots use the same positional criteria for the abscissa and the likelihood estimation (pLI) for the ordinate.

Population impact factor and gene selection

Genes other than *SHANK3* were selected as candidates for detailed review of their possible contribution to PMS. Initial gene selection was based on pLI and the estimated population impact factor (PIF). The PIF is the proportion of the population in which the gene has been disrupted. The PIF is estimated for a given gene by dividing the number of cases of gene disruption by the size of the population. We use terminal deletions from Sarasua et al. [10] for this estimation because it is the largest published sample currently available, although the results are unlikely to vary substantially if we used data from other published studies because of the uniformity in breakpoint locations [7, 11]. Disruption is defined as any deletion that removes part, or all, of the canonical transcript of a gene. To calculate the PIF value, the breakpoints of all 105 terminal deletion cases from Sarasua et al. were located on genome build hg38 based upon the reported deletion sizes. This list was compared to the locus of the last exon of the conical transcript for each gene. The PIF is calculated as the number of case breakpoints that are greater than the locus of the last exon divided by the total number of cases (105). That is, the PIF is the number of individuals who experienced loss or disruption of the gene divided by the total number of individuals in the population.

Two overlapping groups of genes were selected for review based on a combination of pLI and PIF. The first group (Group I) has pLI > 0.9 and PIF > 0.5 (very probably intolerant of LoF variants and over half of the population potentially affected by LoF). The PIF criterion was relaxed for *SULT4A1* because it has been specifically implicated in a study of interstitial deletions [18]. In that study, Disciglio et al. identified a minimal deleted region for nine cases of interstitial deletions roughly 3–7 Mbp in size. *SULT4A1* and *PARVB* were identified as functionally the most likely candidates among 12 genes in the minimal deleted region [18] (see Supplementary Figure S1). Relaxing the PIF criterion for *PARVB* would have no effect given its very low pLI.

The second group of genes (Group II) has $pLI > 0.7$ and $PIF > 0.95$ (likely intolerant of LoF variants and 95% of the population potentially affected by LoF), to not overlook somewhat less likely pLI genes that could impact a high percentage of the population.

Genes were not reviewed if we could not find sufficient evidence to identify a possible association between a phenotypic manifestation of PMS and the known clinical or molecular role of the gene, either because of insufficient literature or an unclear association. Genes were not reviewed if no pLI values were reported for those genes [29]. PMS genes closely associated with a reviewed gene are included in the review for completeness.

Results

Function discovery

Manual curation provided gene candidates that might contribute to common PMS characteristics reported in both retrospective and prospective studies [7–11, 19–22, 43]. We were unable to find specific studies on some genes (e.g., *TTC38*, *PRR34*), whereas other genes have a rich literature (e.g., *SHANK3*). We found a mean of about 2 references/gene ($227/108 = 2.10$). Using the production of a transgenic mouse model as a measure of research investment into gene function, we found that mouse models have been created and validated for nearly one third of PMS genes ($34/108 = 31\%$). The results are summarized using the format recommended by Cody and Hale [44] in Supplementary Table S1. Curated databases and GWAS identified a short list of ASD-associated neuropsychiatric genes (*SHANK3*, *CHKB*, *SBF1*, and *TCF20*).

Classification

Function discovery provided sufficient information to arrange 44 genes into 14 phenotype-related classes (Supplementary Table S2). Classes are not mutually exclusive to allow for pleiotropy. Twelve of these classes are summarized in Fig. 2. All classes, genes assigned to each class, evidence for the assignment and all associated pLI and PIF scores are listed in Supplementary Table S2.

Functional discovery identified 43 of the 108 PMS genes with possible phenotype association. Eighteen of the forty-three genes have either a direct association with a neurodevelopmental disorder or regulate a key feature of neurodevelopment (e.g., neuronal migration, differentiation, cell polarity). Eleven of these genes reside on the distal 1 Mbp of the chromosome (Fig. 2 expanded graph and Supplementary Table S2). However, the remaining neurodevelopmental-associated genes are distributed

throughout the adjacent ~8 Mbp of chromosome 22. Notably, genes associated with known autosomal recessive degenerative diseases (*ARSA*, *CHKB*, *ODF3B*, *SCO2*, *SBF1*, *TUBGCP6*, *MLC1*, *ALG12*, *TBC1D22A*, *ATXN10*) are all localized to distal ~5 Mbp of the chromosome.

Likelihood estimation

Of the 108 PMS genes, 100 meet criteria for assigning a pLI . The genes that do not meet criteria are: *SYCE3*, *LMF2*, *TTL8*, *C22orf34*, *PRR34*, *ARHGAP8*, *SHISA8*, and *C22orf46* [29]. The pLI values for PMS genes are tabulated in Supplementary Table S3 and plotted as a function of position on chromosome 22 in Supplementary Figure S2. Figure 3 shows the distribution of the 18 PMS genes with a $pLI > 0.9$, 7 of which can be found within the distal 1 Mbp of 22q13.33.

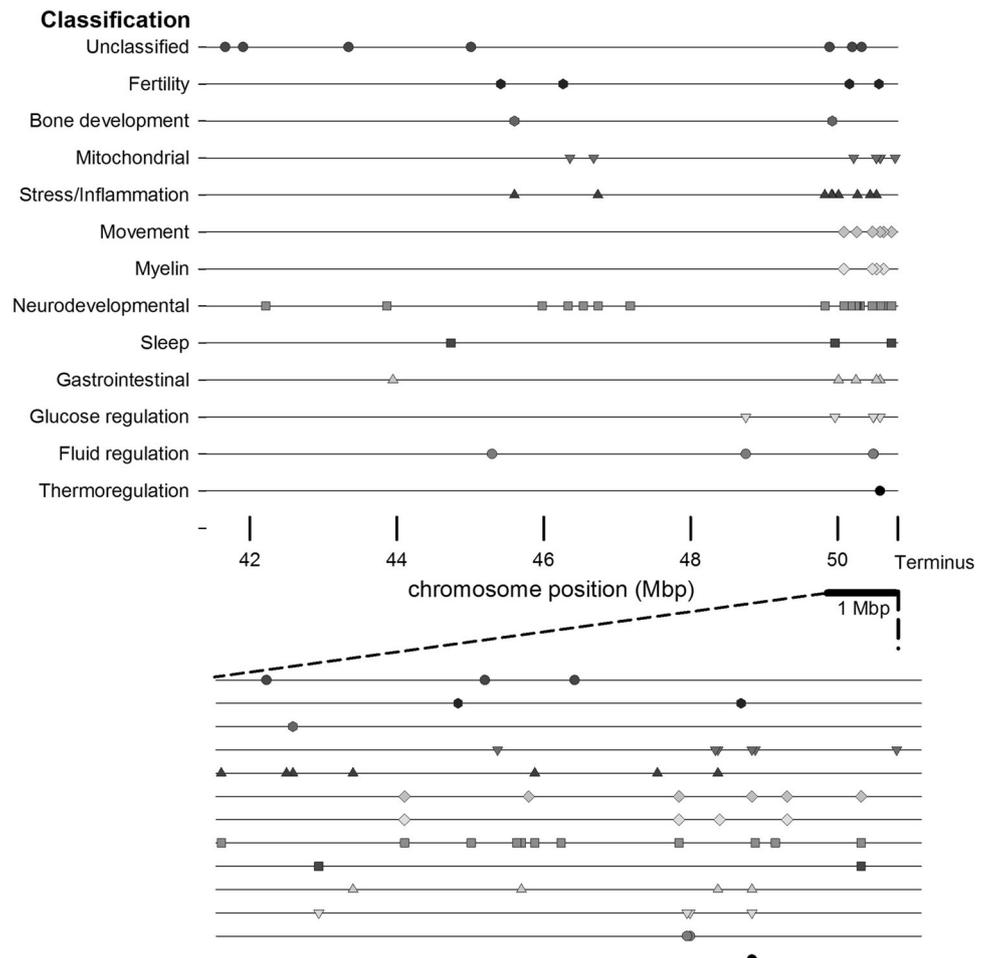
Reviews

Genes other than *SHANK3* that meet the Group I and/or Group II criteria are listed in Table 1. Well-documented genes from this table are reviewed in detail along with genes that are incidentally associated with these genes (Supplementary Document S1). General features emerging from these reviews are summarized here.

The picture that emerges from reviewing high pLI and high PIF genes of PMS is that groups of genes may contribute to common features of the disorder, not unlike the groupings for all PMS genes (Fig. 2). Some of these groups can be tied to synaptic function and thus potentially contribute directly to the phenotypic manifestations associated with *SHANK3* haploinsufficiency. Other groups are associated with more fundamental processes, such as brain development and neuronal organization, which may impact brain function separately from synaptic function per se. Such groups have the potential for pleiotropic impact. Five groups are described here and in detail in Supplementary Document S1.

The first group of PMS genes is the group of MAP kinase genes. There are three members of this group deleted in PMS patients, *MAPK8IP2* ($pLI = 1.00$), *MAPK11* ($pLI = 0.47$), and *MAPK12* ($pLI = 0.01$). The MAPK pathway is a potent regulator of plasticity-related signaling cascades at excitatory synapses [45, 46]. MAP kinases are expressed in multiple brain regions (prefrontal cortex, hippocampus, cerebellum) and have been implicated in regulatory functions for different glutamate receptors (AMPA, NMDA, mGluR) as well as actin dynamics [47–53]. *MAPK8IP2* has preclinical evidence as a potential contributor to the PMS phenotype, including behavioral, motor and motor learning deficits observed in homozygous (null) mice [47]. However, there is substantial potential for additional impact of

Fig. 2 Gene function discovery for 12 phenotype characteristics. Top part of figure shows the 12 characteristics + unclassified plotted across 9.2 Mbp of 22q13.3. The terminus is at the right edge of the plot. Numbers below the top plot show chromosome locations in millions of bp (Mbp) referenced to human genome build hg38. The bottom part is an expanded plot of the distal 1 Mbp of the chromosome. Symbols are the same within a characteristic. Pleiotropic genes are represented more than once in some cases. pLI estimation is evaluated separately. There are 64 unclassified genes, only genes with $pLI > 9$ are plotted



MAPK11 and *MAPK12* loss that would modify losses associated with *SHANK3* and *MAPK8IP2* because of their common pathways, association with neurotransmitters and association with the *MEF2C* transcription factor (Supplementary Document S1).

The second group of PMS genes is a pair of adjacent genes, each of which has been independently associated with neuropsychiatric disorders: *BRD1* ($pLI = 0.99$) and *ZBED4* ($pLI = 0.99$). Both genes are highly expressed in the brain prenatally and both are DNA-binding regulatory genes. Haploinsufficiency of *BRD1* has a wide-ranging impact on neurodevelopment, generally and schizophrenia, specifically [54–56].

The third group includes three PMS genes associated with multicellular organization. Complementary to the *SHANK3* protein's role as a scaffolding for internal cell structure [57], *CELSR1* ($pLI = 1.00$), Plexin-B2 (*PLXNB2*, $pLI = 0.99$), and Fibulin-1 (*FBLN1*, $pLI = 0.85$) proteins provide critical external structure for neural assemblies. *CELSR1* is involved in tissue patterning in the cochlea, lungs, brain and elsewhere (see Supplementary Document S1). Plexin-B2 regulates the timing of neuronal

differentiation and migration during development [58]. Fibulin-1, an extracellular matrix protein, is associated with syndactyly and thinning of the myocardium in mice [59] and homozygous variants have been associated with syndactyly in humans [60]. Its impact on the nervous system is limited to cranial nerves IX and X, but it may also affect other organs [59].

The fourth group of PMS genes impact cerebellar development and function. That group entails six PMS genes (counting *SHANK3*). These genes are highly LoF intolerant and are highly and/or differentially expressed in the cerebellum: *SHANK3* (granule cells) [61], *ZBED4* (cerebellum) [62], *MAPK8IP2* (Purkinje cells) [47], *SULT4A1* (granule cell layer) [63], *PIM3* (upregulated in cerebellum after seizures) [64], and *PLXNB2* (granule cell differentiation and migration) [65]. In one study, structural abnormalities of the cerebellum were observed in 8 out of 10 subjects with PMS [37]. Testing of cerebellar motor function may be an especially useful biomarker to investigate the impact of polygenic loss on central nervous system (CNS) function (e.g., refs. [66, 67]).

Fig. 3 Probability of LoF intolerance for PMS genes. Genes with high pLI scores ($p > 0.9$) are plotted as a function of chromosome position. Inset graph expands the distal chromosome. Each gene is assigned to only one class

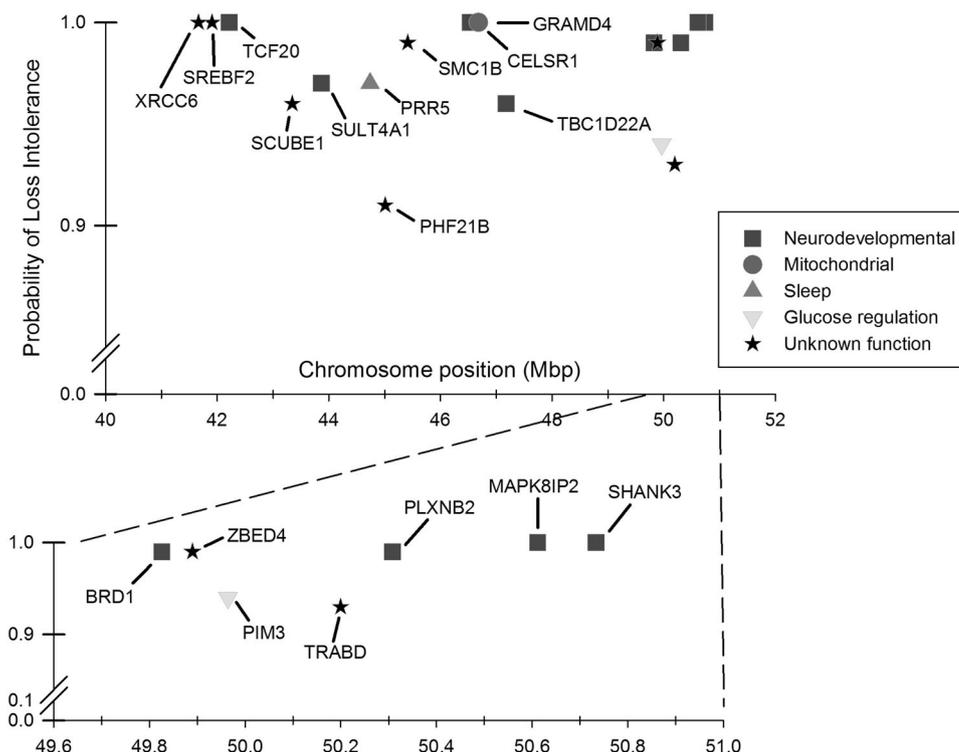


Table 1 Genes that met one or both group criteria

Gene	Group I (pLI > 0.9, PIF > 0.5)	Group II (pLI > 0.7, PIF > 0.95)	Well documented
<i>MAPK8IP2</i>	X	X	X
<i>PLXNB2</i>	X	X	X
<i>ZBED4</i>	X	X	X
<i>BRD1</i>	X	X	X
<i>PIM3</i>	X	X	X
<i>TRABD</i>	X	X	
<i>TBC1D22A</i>	X		
<i>GRAMD4</i>	X		
<i>CELSR1</i>	X		X
<i>SULT4A1</i> ^a	X		X
<i>RABL2B</i>		X	X
<i>PPP6R2</i>		X	X

Only well-documented genes were reviewed

pLI probability of loss of function intolerance, PIF population impact factor

^aPIF for *SULT4A1* is below 0.5; *SULTA1* is included by exception as described in Materials and methods

The last group of PMS genes is mitochondria-associated genes. Four of these genes have previously been identified as a functional group associated with PMS, *GRAMD4* (pLI = 1.00), *SCO2* (pLI = 0), *TYMP* (pLI = 0), and *CPT1B* (pLI = 0) [43]. The role of mitochondrial genes in PMS is

unclear, since there was no consistent phenotype or correlation with deletion size [43]. The review of *RABL2B* (pLI = 0.77) provides a possible alternative association with mitochondrial function. *Rabl2*, the mouse ortholog of human *RABL2A/B*, is important for mitochondria transport as demonstrated in the mouse liver [68]. It is notable that mitochondria are normally moved to post-synaptic boutons in an activity-dependent manner [69, 70]. We speculate that poor activity-dependent transport of mitochondria might exacerbate activity-dependent deficits of *SHANK3* at glutamate post-synaptic terminals in the face of *SHANK3* haploinsufficiency in a pattern that follows the differential patterning of *RABL2A* versus *RABL2B* in the brain [71]. Unfortunately, *RABL2B* specific function cannot be studied using typical methods in model species because of the unique evolutionary pattern of *RABL2B* in humans. Human-induced pluripotent stem cells (iPSCs) and iPSC-derived neurons from patients might provide an opportunity for studying the loss of *RABL2B* in PMS.

Discussion

To the best of our knowledge, this is the first systematic overview that identifies protein-coding genes commonly deleted in the 22q13 region that are likely contributors to manifestations of PMS. We focused on genes most frequently lost [7, 10] and those most likely to influence

survival based on computational modeling of variant frequency (LoF intolerance) in the general population [29], using pre-determined selection criteria (Fig. 1). In parallel, we performed a literature-based functional discovery to classify all 108 PMS genes. We show that for $\text{PIF} > 0.95$ (95% of terminal deletions), six genes are very highly LoF intolerant and 5/6 are potential contributors to CNS dysfunction over and above *SHANK3* haploinsufficiency.

Classification

Of the 108 PMS genes, we found sufficient information to classify 44 into functional categories (Supplementary Table S2). Genes that met the pLI and PIF criteria fell into five groups of potential impact on CNS function, including MAP kinases and genes associated with DNA regulation, multicellular organization, cerebellar function, and mitochondria. Identifying functional groups has two implications. First, multiple gene losses within a group may have cumulative effects. Second, if there are cumulative effects, therapeutic target selection may benefit from that knowledge.

Interstitial deletions

By design, the PIF metric favors distal genes to estimate population impact on the most common cases of PMS, terminal deletions. However, PMS does not exclude interstitial deletions that spare *SHANK3* [2] and there is no distinguishing phenotype difference between terminal deletions and interstitial deletions that do not include *SHANK3* [5, 10, 17, 18]. This similarity suggests that interstitial genes may contribute significantly to the common characteristics of PMS. Although the impact of deleted non-coding regions of chromosome 22 on *SHANK3* regulation cannot be ruled-out, it seems likely that high pLI genes, notably *SULT4A1* [18], are potential causal genes for cases of interstitial deletions. The possibility of undetected pathological *SHANK3* variants must also be ruled out.

Pleiotropic effects

Some of the genes we reviewed have evidence for pleiotropic effects. *PIM3* and *MAPK8IP2*, for example, both participate in the regulation of glucose, and *PIM3* has been associated with regulation of the circadian rhythm. *ZBED4*, *FBLN1*, and *CELSR1* all influence kidney development (Supplementary Document S1). One third (4/12) of the genes most likely to contribute to the PMS phenotype are poorly characterized, but these genes may also be pleiotropic and they may not be any less important than the well characterized genes [72]. Some pleiotropic effects may be more penetrant than others for a given gene, and thus may

contribute to the variability of features reported in individuals with PMS, such as sleep disturbances or kidney disorders.

Limitations of the study

Our results are computationally-based and, as such, can only identify candidate genes. The actual contributions of each candidate to the PMS phenotype will depend on clinical and functional studies, with supporting evidence from animal models and cellular/molecular studies. Crucial genes may be missed using the pLI as a predictor [29]. The pLI is also not aimed at major polygenic losses as in PMS, whereas other approaches do address polygenic loss [73]. We note that our function discovery is based on previously identified roles for each gene, which may not capture critical additional (e.g., pleiotropic) action in the central nervous system. Separating the impact of *SHANK3* disruption from the disruption of other genes is problematic, perhaps best addressed through the study of individuals with restricted interstitial deletions that spare *SHANK3*. However, this is challenging work considering the relatively small number of individuals with such genetic abnormalities.

Future directions

PMS cases with a deleted or an otherwise disrupted copy of *SHANK3* provide valuable information on phenotype–genotype investigations of this important ASD-associated gene. Interpreting these studies, however, must consider the broader genetic landscape. The present work begins to fill in the picture, although considerably more needs to be done to understand the genetic origins of the PMS phenotype. Protein-coding genes are only part of the DNA lost in terminal or interstitial deletions. More detailed work is required to account for other genes, regulatory elements [74], and even the possible contributions of small open reading frames [75]. The unknown functions of poorly characterized protein-coding genes with very high pLI need to be addressed. Most importantly, more clinical studies are needed to validate the contributions of individual genes to the PMS phenotype, especially studies that include interstitial deletions. These steps will bring an opportunity to identify new therapeutic targets and bring aid to those awaiting relief from the difficult realities of PMS.

Acknowledgements We thank Per Qvist for help with the BRD1 and ZBED4 gene reviews. We thank Drs J Giza, C Sala, C Verpelli, and Y-H Jiang for their comments on an early draft of the manuscript. This research was supported (in part) by the Intramural Research Program of the National Institute of Mental Health, National Institutes of Health (Annual Report numbers: ZIA MH002868 and ZIAMH002886), by a Seed Grant from the Simons Center for the Social Brain, at the

Massachusetts Institute of Technology, and support from the Simons Foundation to WEK, and by the Slifka/Ritvo Innovation in Autism Award and Whitehall (2015-08-86), Brain Research Foundation (BRFSG-2016-08) and Brain and Behaviour Foundation (25456) Grants to AS.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, and provide a link to the Creative Commons license. You do not have permission under this license to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Phelan MC, Rogers RC, Saul RA, et al. 22q13 deletion syndrome. *Am J Med Genet.* 2001;101:91–99.
- Phelan K, McDermid HE. The 22q13.3 deletion syndrome (Phelan-McDermid Syndrome). *Mol Syndromol.* 2012;2:186–201.
- Nesslinger NJ, Gorski JL, Kurczynski TW, et al. Clinical, cytogenetic, and molecular characterization of seven patients with deletions of chromosome 22q13.3. *Am J Hum Genet.* 1994;54:464–72.
- Doheny KF, McDermid HE, Harum K, Thomas GH, Raymond GV. Cryptic terminal rearrangement of chromosome 22q13.32 detected by FISH in two unrelated patients. *J Med Genet.* 1997;34:640–4.
- Wilson HL, Crolla JA, Walker D, et al. Interstitial 22q13 deletions: genes other than SHANK3 have major effects on cognitive and language development. *Eur J Hum Genet.* 2008;16:1301–10.
- Dhar SU, del Gaudio D, German JR, et al. 22q13.3 deletion syndrome: clinical and molecular analysis using array CGH. *Am J Med Genet A.* 2010;152A:573–81.
- Soorya L, Kolevzon A, Zweifach J, et al. Prospective investigation of autism and genotype-phenotype correlations in 22q13 deletion syndrome and SHANK3 deficiency. *Mol Autism.* 2013;4:18.
- Zwanenburg RJ, Ruiters SA, van den Heuvel ER, Flapper BC, Van Ravenswaaij-Arts CM. Developmental phenotype in Phelan-McDermid (22q13.3 deletion) syndrome: a systematic and prospective study in 34 children. *J Neurodev Disord.* 2016;8:16.
- Jeffries AR, Curran S, Elmslie F, et al. Molecular and phenotypic characterization of ring chromosome 22. *Am J Med Genet A.* 2005;137:139–47.
- Sarasua SM, Boccutto L, Sharp JL, et al. Clinical and genomic evaluation of 201 patients with Phelan-McDermid syndrome. *Hum Genet.* 2014;133:847–59.
- Bonaglia MC, Giorda R, Beri S, et al. Molecular mechanisms generating and stabilizing terminal 22q13 deletions in 44 subjects with Phelan/McDermid syndrome. *PLoS Genet.* 2011;7:e1002173.
- Macedoni-Luksic M, Krgovic D, Zagradisnik B, Kokalj-Vokac N. Deletion of the last exon of SHANK3 gene produces the full Phelan-McDermid phenotype: a case report. *Gene.* 2013;524:386–9.
- Serret S, Thummler S, Dor E, Vesperini S, Santos A, Askenazy F. Lithium as a rescue therapy for regression and catatonia features in two SHANK3 patients with autism spectrum disorder: case reports. *BMC Psychiatry.* 2015;15:107.
- Leblond CS, Nava C, Polge A, et al. Meta-analysis of SHANK mutations in autism spectrum disorders: a gradient of severity in cognitive impairments. *PLoS Genet.* 2014;10:e1004580.
- de Sena Cortabitarte A, Degenhardt F, Strohmaier J, et al. Investigation of SHANK3 in schizophrenia. *Am J Med Genet B Neuropsychiatr Genet.* 2017;174:390–8.
- Gauthier J, Champagne N, Lafreniere RG, et al. De novo mutations in the gene encoding the synaptic scaffolding protein SHANK3 in patients ascertained for schizophrenia. *Proc Natl Acad Sci USA.* 2010;107:7863–8.
- Simenson K, Oiglane-Shlik E, Teek R, Kuuse K, Ounap K. A patient with the classic features of Phelan-McDermid syndrome and a high immunoglobulin E level caused by a cryptic interstitial 0.72-Mb deletion in the 22q13.2 region. *Am J Med Genet A.* 2014;164A:806–9.
- Disciglio V, Lo Rizzo C, Mencarelli MA, et al. Interstitial 22q13 deletions not involving SHANK3 gene: a new contiguous gene syndrome. *Am J Med Genet A.* 2014;164A:1666–76.
- Denayer A, Van Esch H, de Ravel T, et al. Neuropsychopathology in 7 patients with the 22q13 deletion syndrome: presence of bipolar disorder and progressive loss of skills. *Mol Syndromol.* 2012;3:14–20.
- Sarasua SM, Dwivedi A, Boccutto L, et al. Association between deletion size and important phenotypes expands the genomic region of interest in Phelan-McDermid syndrome (22q13 deletion syndrome). *J Med Genet.* 2011;48:761–6.
- Luciani JJ, de Mas P, Depetris D, et al. Telomeric 22q13 deletions resulting from rings, simple deletions, and translocations: cytogenetic, molecular, and clinical analyses of 32 new observations. *J Med Genet.* 2003;40:690–6.
- Hannachi H, Mougou S, Benabdallah I, et al. Molecular and phenotypic characterization of ring chromosome 22 in two unrelated patients. *Cytogenet Genome Res.* 2013;140:1–11.
- Kolevzon A, Bush L, Wang AT, et al. A pilot controlled trial of insulin-like growth factor-1 in children with Phelan-McDermid syndrome. *Mol Autism.* 2014;5:54.
- Zwanenburg RJ, Bocca G, Ruiters SA, et al. Is there an effect of intranasal insulin on development and behaviour in Phelan-McDermid syndrome? A randomized, double-blind, placebo-controlled trial. *Eur J Hum Genet.* 2016;24:1696–701.
- Harony-Nicolas H, De Rubeis S, Kolevzon A, Buxbaum JD. Phelan McDermid syndrome: from genetic discoveries to animal models and treatment. *J Child Neurol.* 2015;30:1861–70.
- Phelan McDermid Syndrome Foundation: About Us. 2016. http://22q13.org/j15/index.php?option=com_content&view=article&id=115&Itemid=111. Accessed 29 November 2016.
- Bonaglia MC, Giorda R, Mani E, et al. Identification of a recurrent breakpoint within the SHANK3 gene in the 22q13.3 deletion syndrome. *J Med Genet.* 2006;43:822–8.
- Exome Aggregation Consortium: ExAC Browser: Broad Institute, 2016. <http://exac.broadinstitute.org/>. Accessed 14 October 2016.
- Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536:285–91.
- McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University. Online Mendelian Inheritance in

- Man (OMIM). 2017. <http://omim.org>. Accessed 16 March 2017.
31. NCBI Resource Coordinators. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 2017; **45**: D12–17.
 32. Flicek P, Amode MR, Barrell D, et al. Ensembl 2014. *Nucleic Acids Res.* 2014;42:D749–755.
 33. Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucleic Acids Res.* 2015;43:D1049–56.
 34. LifeMap Sciences: GeneCards Human Gene Database, 2016. <https://www.genecards.org/>. Accessed 29 November 2016.
 35. Shaw SR, Rahman A, Sharma A. Behavioral profiles in Phelan-McDermid syndrome: focus on mental health. *J Ment Health Res Intellect Disabil.* 2011;4:1–18.
 36. Oberman LM, Boccuto L, Cascio L, Sarasua S, Kaufmann WE. Autism spectrum disorder in Phelan-McDermid syndrome: initial characterization and genotype-phenotype correlations. *Orphanet J Rare Dis.* 2015;10:105.
 37. Aldinger KA, Kogan J, Kimonis V, et al. Cerebellar and posterior fossa malformations in patients with autism-associated chromosome 22q13 terminal deletion. *Am J Med Genet A.* 2013;161A:131–6.
 38. Yuen RK, Thiruvahindrapuram B, Merico D, et al. Whole-genome sequencing of quartet families with autism spectrum disorder. *Nat Med.* 2015;21:185–91.
 39. Sanders SJ, He X, Willsey AJ, et al. Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron.* 2015;87:1215–33.
 40. Abrahams BS, Arking DE, Campbell DB, et al. SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Mol Autism.* 2013;4:36.
 41. Yuen RKC, Merico D, Bookman M, et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. *Nat Neurosci.* 2017;20:602–11.
 42. Samocha KE, Robinson EB, Sanders SJ, et al. A framework for the interpretation of de novo mutation in human disease. *Nat Genet.* 2014;46:944–50.
 43. Frye RE, Cox D, Slattery J, et al. Mitochondrial dysfunction may explain symptom variation in Phelan-McDermid syndrome. *Sci Rep.* 2016;6:19544.
 44. Cody JD, Hale DE. Linking chromosome abnormality and copy number variation. *Am J Med Genet A.* 2011;155A:469–75.
 45. Dhanasekaran DN, Kashef K, Lee CM, Xu H, Reddy EP. Scaffold proteins of MAP-kinase modules. *Oncogene.* 2007;26:3185–202.
 46. Yasuda J, Whitmarsh AJ, Cavanagh J, Sharma M, Davis RJ. The JIP group of mitogen-activated protein kinase scaffold proteins. *Mol Cell Biol.* 1999;19:7245–54.
 47. Giza J, Urbanski MJ, Prestori F, et al. Behavioral and cerebellar transmission deficits in mice lacking the autism-linked gene *islet brain-2*. *J Neurosci.* 2010;30:14805–16.
 48. Kennedy NJ, Martin G, Ehrhardt AG, et al. Requirement of JIP scaffold proteins for NMDA-mediated signal transduction. *Genes Dev.* 2007;21:2336–46.
 49. Bolshakov VY, Carboni L, Cobb MH, Siegelbaum SA, Belardetti F. Dual MAP kinase pathways mediate opposing forms of long-term plasticity at CA3-CA1 synapses. *Nat Neurosci.* 2000;3:1107–12.
 50. Rush AM, Wu J, Rowan MJ, Anwyl R. Group I metabotropic glutamate receptor (mGluR)-dependent long-term depression mediated via p38 mitogen-activated protein kinase is inhibited by previous high-frequency stimulation and activation of mGluRs and protein kinase C in the rat dentate gyrus in vitro. *J Neurosci.* 2002;22:6121–8.
 51. Zhong P, Liu W, Gu Z, Yan Z. Serotonin facilitates long-term depression induction in prefrontal cortex via p38 MAPK/Rab5-mediated enhancement of AMPA receptor internalization. *J Physiol.* 2008;586:4465–79.
 52. Eales KL, Palygin O, O'Loughlin T, et al. The MK2/3 cascade regulates AMPAR trafficking and cognitive flexibility. *Nat Commun.* 2014;5:4701.
 53. Sanderson TM, Hogg EL, Collingridge GL, Correa SA. Hippocampal metabotropic glutamate receptor long-term depression in health and disease: focus on mitogen-activated protein kinase pathways. *J Neurochem.* 2016;139(Suppl 2):200–14.
 54. Severinsen JE, Bjarkam CR, Kiaer-Larsen S, et al. Evidence implicating BRD1 with brain development and susceptibility to both schizophrenia and bipolar affective disorder. *Mol Psychiatry.* 2006;11:1126–38.
 55. Fryland T, Christensen JH, Pallesen J, et al. Identification of the BRD1 interaction network and its impact on mental disorder risk. *Genome Med.* 2016;8:53.
 56. Qvist P, Christensen JH, Vardya I, et al. The schizophrenia associated *brd1* gene regulates behavior, neurotransmission, and expression of schizophrenia risk enriched gene sets in mice. *Biol Psychiatry.* 2016;82:62–76.
 57. Boeckers TM, Bockmann J, Kreutz MR, Gundelfinger ED. ProSAP/Shank proteins - a family of higher order organizing molecules of the postsynaptic density with an emerging role in human neurological disease. *J Neurochem.* 2002;81:903–10.
 58. Deng S, Hirschberg A, Worzfeld T, et al. Plexin-B2, but not Plexin-B1, critically modulates neuronal migration and patterning of the developing nervous system in vivo. *J Neurosci.* 2007;27:6333–47.
 59. Cooley MA, Kern CB, Fresco VM, et al. Fibulin-1 is required for morphogenesis of neural crest-derived structures. *Dev Biol.* 2008;319:336–45.
 60. Debeer P, Schoenmakers EF, Twal WO, et al. The fibulin-1 gene (*FBLN1*) is disrupted in a t(12;22) associated with a complex type of synpolydactyly. *J Med Genet.* 2002;39:98–104.
 61. Bockers TM, Segger-Junius M, Iglauer P, et al. Differential expression and dendritic transcript localization of Shank family members: identification of a dendritic targeting element in the 3' untranslated region of *Shank1* mRNA. *Mol Cell Neurosci.* 2004;26:182–90.
 62. Carithers LJ, Moore HM. The Genotype-Tissue Expression (GTEx) Project. *Biopreserv Biobank.* 2015;13:307–8.
 63. Liyou NE, Buller KM, Tresillian MJ, et al. Localization of a brain sulfotransferase, *SULT4A1*, in the human and rat brain: an immunohistochemical study. *J Histochem Cytochem.* 2003;51:1655–64.
 64. Feldman JD, Vician L, Crispino M, et al. *KID-1*, a protein kinase induced by depolarization in brain. *J Biol Chem.* 1998;273:16535–43.
 65. Friedel RH, Kerjan G, Rayburn H, et al. Plexin-B2 controls the development of cerebellar granule cells. *J Neurosci.* 2007;27:3921–32.
 66. Wang Z, Magnon GC, White SP, Greene RK, Vaillancourt DE, Mosconi MW. Individuals with autism spectrum disorder show abnormalities during initial and subsequent phases of precision gripping. *J Neurophysiol.* 2015;113:1989–2001.
 67. Schmitt LM, Cook EH, Sweeney JA, Mosconi MW. Saccadic eye movement abnormalities in autism spectrum disorder indicate dysfunctions in cerebellum and brainstem. *Mol Autism.* 2014;5:47.
 68. Yi LoJC, O'Connor AE, Andrews ZB, et al. *RABL2* is required for hepatic fatty acid homeostasis and its dysfunction leads to steatosis and a diabetes-like state. *Endocrinology.* 2016;157:4732–43.

69. Zhou B, Yu P, Lin MY, Sun T, Chen Y, Sheng ZH. Facilitation of axon regeneration by enhancing mitochondrial transport and rescuing energy deficits. *J Cell Biol.* 2016;214:103–19.
70. Todorova V, Blokland A. Mitochondria and synaptic plasticity in the mature and aging nervous system. *Curr Neuropharmacol.* 2017;15:166–73.
71. Kramer M, Backhaus O, Rosenstiel P, et al. Analysis of relative gene dosage and expression differences of the paralogs RABL2A and RABL2B by pyrosequencing. *Gene.* 2010;455:1–7.
72. Pandey AK, Lu L, Wang X, Homayouni R, Williams RW. Functionally enigmatic genes: a case study of the brain ignormome. *PLoS One.* 2014;9:e88889.
73. Andrews T, Honti F, Pfundt R, et al. The clustering of functionally related genes contributes to CNV-mediated disease. *Genome Res.* 2015;25:802–13.
74. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489:57–74.
75. Couso JP, Patraquim P. Classification and function of small open reading frames. *Nat Rev Mol Cell Biol.* 2017;18:575–89.