Narrowing the Genetic Causes of Language Dysfunction in the 1q21.1 Microduplication Syndrome

Antonio Benítez-Burraco [1]*, Montserrat Barcos-Martínez [2,3], Isabel Espejo-Portero [2,3], Maite Fernández-Urquiza [4], Raúl Torres-Ruiz [5], Sandra Rodríguez-Perales [5] and Mª Salud Jiménez-Romero [6]

[1] Department of Spanish, Linguistics, and Theory of Literature, University of Seville, Seville, Spain, [2] Laboratory of Molecular Genetics, University Hospital "Reina Sofía", Córdoba, Spain, [3] Maimónides Institute of Biomedical Research, Córdoba, Spain, [4] Department of Spanish Philology, University of Oviedo, Oviedo, Spain, [5] Molecular Cytogenetics Group, Centro Nacional Investigaciones Oncológicas, Madrid, Spain, [6] Department of Psychology, University of Córdoba, Córdoba, Spain

The chromosome 1q21.1 duplication syndrome (OMIM# 612475) is characterized by head anomalies, mild facial dysmorphisms, and cognitive problems, including autistic features, mental retardation, developmental delay, and learning disabilities. Speech and language development are sometimes impaired, but no detailed characterization of language problems in this condition has been provided to date. We report in detail on the cognitive and language phenotype of a child who presents with a duplication in 1q21.1 (arr[hg19] 1q21.1q21.2(145,764,455-147,824,207)×3), and who exhibits cognitive delay and behavioral disturbances. Language is significantly perturbed, being the expressive domain the most impaired area (with significant dysphemic features in absence of pure motor speech deficits), although language comprehension and use (pragmatics) are also affected. Among the genes found duplicated in the child, *CDH1L* is upregulated in the blood of the proband. *ROBO1*, a candidate for dyslexia, is also highly upregulated, whereas, TLE3, a target of FOXP2, is significantly downregulated. These changes might explain language, and particularly speech dysfunction in the proband.

Keywords: chromosome 1q21.1 duplication syndrome, cognitive delay, language deficits, speech problems, CDH1L, ROBO1

INTRODUCTION

Copy number variant (CNV) syndromes entailing language impairment provide important evidence of how changes in gene dosage impact on the wiring and function of brain areas involved in language processing. The 1q21.1 region contains several segmental-duplication blocks which favor the occurrence of microdeletions and microduplications. The chromosome 1q21.1 duplication syndrome (OMIM# 612475) results from microduplications of the BP3-BP4 region, with a minimum duplicated region of ~1.2 Mb of unique DNA sequence, which includes at least seven genes, although a less common variant, resulting from microduplications of the BP2-BP4 region, with a duplicated region of ~2 Mb and encompassing the region deleted in the TAR syndrome (OMIM# 274000), has been also reported (1). These two variants are commonly referred as Class I 1q21.1 duplication syndrome and Class II 1q21.1 duplication syndrome, respectively.

Many of the subjects bearing the 1q21.1 microduplication exhibit autistic features, attention deficit-hyperactivity disorder, mild or moderate mental retardation, relative macrocephaly, mild dysmorphic features, heart disease, and hypotonia (1–3). Some of these features are shared with carriers of the reciprocal deletion, including cognitive dysfunction, motor problems, articulation deficits, and hypotonia (4). Interestingly, microcephaly is prevalent in deletion carriers, whereas macrocephaly is common in duplication carriers; likewise, deletions are usually associated with schizophrenia, while carriers of the duplication usually exhibit autistic features [(4) and references herein]. Interestingly too, the adult phenotype seems to be milder than the child presentation of the syndrome and includes macrocephaly and abnormalities of possible connective tissue origin, and occasionally, schizophrenia (3). Overall, because the microduplication (as well as the reciprocal deletion) is associated with a highly variable phenotype, and because it exhibits an incomplete penetrance and variable expressivity, it has been suggested that this is a susceptibility locus, rather than a clinically distinct syndrome, which predisposes to neurodevelopmental problems (5).

Little is known about the genetic causes of the neuropsychiatric phenotype of 1q21.1 duplications. The distinctive genomic architecture of the 1q21.1 region, which accounts for its genomic instability, seems to be evolutionarily advantageous, and includes multiple human-specific changes (compared to extant primates) important for brain evolution, such as the expansion of genes coding for proteins with DUF1220 protein motifs, and the duplicative transpositions of non-1q21.1 genes, like *SRGAP2* and *HYDIN* (6). According to Dolcetti et al. (3), gene expression data suggest that candidate genes for the distinctive cognitive profile of people with the 1q21.1 microduplication may include *BCL9*, *GJA8*, *PDZK1*, and *PRKAB2*. Nonetheless, no conclusive genotype-phenotype correlations have been found. This is particularly true of the language problems that are commonly observed in patients, among other reasons, because we still lack a comprehensive characterization of the language phenotype associated to this condition.

In this paper, we report on a boy with a microduplication of the 1q21.1 region and provide a detailed description of his language (dis)abilities. Relying on our current knowledge of the genetic aspects of language development and language evolution, but also on *ad-hoc* analyses of the expression pattern of selected genes in the blood of our proband, we propose *BCL9* and *CDH1L* as principal candidates for language dysfunction in this syndrome.

MATERIAL AND METHODS
Linguistic, Cognitive, and behavioral Assessment
The global developmental profile of the proband was evaluated with the Spanish versions of the Battelle Developmental Inventories (7), the Luria-Nebraska Neuropsychological Battery

for Children (LNNB-C) (8), the Kaufman Assessment Battery for Children (K-ABC) (9), and the Inventory for Client and Agency Planning (ICAP) (10). The Spanish version of the Autism Diagnostic Observation Schedule- Revised (ADOS-R) was used, specifically, for screening autistic features in the child.

Battelle Developmental Inventories
Battelle Developmental Inventories consist of 341 items and are aimed to evaluate personal/social development, adaptive capabilities, motor abilities (gross and fine), communication skills (in the receptive and expressive domains), and cognitive development.

Luria-Nebraska Neuropsychological Battery for Children (LNNB-C)
The LNNB-C is aimed to evaluate (and anticipate) the child's learning, experience, and cognitive skills, as well as his neuropsychological deficiencies. The test comprises 10 scales that are aimed to measure executive functioning in several areas of interest (manual mobility, left-right orientation, gestures and praxias, verbal regulation, spatial orientation) as well as language functioning in different domains (object and drawing naming, phonological awareness, vocabulary in pictures, similarities and dissimilarities, and numerical operations).

Kaufman Assessment Battery for Children (K-ABC)
The K-ABC is aimed to evaluate diverse processing and cognitive abilities as they are put into use, either simultaneously or sequentially, to resolve specific problems. The test comprises 16 subtests arranged into three scales: Simultaneous Processing, Sequential Processing, and Knowledge.

Inventory for Client and Agency Planning (ICAP)
The ICAP is aimed to evaluate the subject's functional abilities and maladaptive behaviors in the following general areas: motor skills, social and communication skills, personal living skills, and community living skills. The ICAP measures the frequency and severity of eight types of behavioral disturbances, which are organized in three subscales: asocial maladaptive behavior (uncooperative behavior and socially offensive behavior), internalized maladaptive behavior (withdrawn or inattentive behavior, unusual or repetitive habits, and self-harm), and externalized maladaptive behavior (disruptive behavior, destructive to property, and hurtful to others). Behavior is rated as normal or abnormal, whereas behavioral problems are subsequently rated as marginally serious, moderately serious, serious, or very serious.

Autism Diagnostic Observation Schedule-Revised (ADOS-R)
This test is designed for toddlers between 18 and 60 months and comprises 23 questions. A positive score is indicative of the possibility of suffering from the disease.

Regarding the language profile of the child, his abilities were evaluated in detail with the Prueba de Lenguaje Oral de Navarra-Revisada [Navarra Oral Language Test, Revised] (PLON-R) (11),

and the Spanish version of the Peabody Picture Vocabulary Test, Third Edition [PPVT-3; (12)].

Prueba de Lenguaje Oral de Navarra-Revisada (PLON-R)

This test was used to assess in depth the child's expressive abilities. The PLON-R is aimed for children between 3 and 6 years, and characterizes language production both structurally and functionally. The composite score "form" evaluates speech sound production at the level of single word, phonological awareness, as well as morphological and syntactic features of the child's imitated and elicited verbal production. The composite score "contents" assesses word and sentence meanings. Finally, the test also evaluates different aspects of language use (planning, autoregulation, and modification under environmental cues) in semi-spontaneous conditions.

Peabody Picture Vocabulary Test (PPVT-3)

This test was used to evaluate in more detail the receptive abilities of the proband. The PPVT-3 assesses the correct acquisition of words by the child via several tasks in which the child is asked to point to one of four color pictures on a page after hearing a word.

Additionally, language in use and overall communication skills in spontaneous conditions were assessed through the systematic analysis of a 17-min sample of the child's talk in interaction with his mother. The conversation was video-recorded and then transcribed and coded using CHAT (*Codes for the Human Analysis of Transcripts*) software, a tool of the CHILDES Project (13). CHAT allows to code for speech production phenomena (including information about articulation, prosody, and fluency) in the main lines of the transcript, and for phonological, morphological, syntactic and pragmatic phenomena in dependent lines.

It has to be pointed out that, when evaluating language in use, every linguistic phenomenon is to be assessed from a pragmatic perspective (14). Thus, we first tagged the linguistic errors made by the child indicating the linguistic components affected (phonology, morphology, syntax, lexical semantics). Then we evaluated the communicative effects of each error and labelled it accordingly on the pragmatic level, using the Pragmatic Evaluation Protocol for the analysis of oral Corpora (PREP-CORP), which has been satisfactorily used to provide the linguistic profile of other neurogenic disorders (15–17). The number of occurrences of each labeled phenomenon was automatically computed by means of CLAN (*Computerized Language Analysis*) (18), allowing us to build up a global description of the pragmatic linguistic profile of the proband, that is to say, of the most salient phenomena of his language in use.

Finally, the linguistic profile of the proband's parents was assessed with the verbal scale of the Spanish version of the WAIS-III (19), the Test de Aptitud Verbal "Buenos Aires" [Buenos Aires Verbal Aptitude Test] (BAIRES) (20), the Batería para la Evaluación de los Procesos Lectores en Secundaria y Bachillerato [Battery for the evaluation of reading abilities in high school students] (PROLEC-SE) (21), and one specific task aimed to evaluate the comprehension of passives and correferences.

WAIS-III

The verbal scale of the WAIS-III comprises six tasks aimed to assess the subject's abilities in two different domains of language processing: verbal comprehension [Similarities (S), Vocabulary (V), Information (I), and Comprehension (CO)] and working memory [Digits (D) and Arithmetic (A)].

BAIRES

The BAIRES is aimed to evaluate language comprehension and production in subjects above 16 years old. It comprises two tasks which assesses the subject's ability to understand and provide with synonymic words and verbal definitions, respectively.

PROLEC-SE

The PROLEC-SE is aimed to assess the reading abilities by adolescents between 12 and 18 years old. It comprises six tasks which evaluate three different processes involved in reading: lexical (word reading, pseudoword reading), semantic (text comprehension, text structure), and syntactic (picture-sentence mapping, punctuation marks).

Ad-hoc task

The *ad-hoc* task was a sentence-picture-matching task aimed to evaluate the comprehension of bound anaphora in complex sentences (that is, the ability to properly bind a determiner phrase in the subordinated clause to a referential element in the main clause) and of canonical long passive structures (that is, with a *by*-phrase). While the former task enquires about highly-demanding computational abilities in the domain of language processing, the latter is usually difficult for people with poor reading abilities, because in Spanish passives are much more frequent in the literary register.

Molecular and cytogenetic Analysis

DNA from the patient and his parents was extracted from 100 μl of EDTA-anticoagulated whole blood using MagNA Pure (Roche Diagnostics, West Sussex, UK) and used for subsequent analyses.

Karyotype Analysis

Peripheral venous blood lymphocytes were grown following standard protocols and collected after 72 h. A moderate resolution G-banding (550 bands) karyotyping by trypsin (Gibco 1x trypsin® and Leishmann stain) was subsequently performed. Microscopic analysis was performed with a Nikon® eclipse 50i optical microscope and the IKAROS Karyotyping System (MetaSystem® software).

Fragile X Syndrome Analysis

CGG expansions affecting the gene *FMR1* (the main determinant for Fragile X syndrome) were analyzed in the patient according to standard protocols. Polymerase chain reaction (PCR) of the fragile site was performed with specific primers for the fragile region of the *FMR1 locus* and the trinucleotide repeat size of the resulting fragments was evaluated by electrophoresis in agarose gel.

Multiplex Ligation-Dependent Probe Amplification (MLPA)

MLPA was performed to detect abnormal CNV of the subtelomeric regions of the probands' chromosomes. MLPA is based on the amplification (by use of a single PCR primer pair) of many different probes, each of which detecting a specific subtelomeric DNA sequence. Three different kits from MRC-Holland were used: SALSA® MLPA® probemix P036-E3 SUBTELOMERES MIX 1, SALSA® MLPA® probemix P070-B3 SUBTELOMERES MIX 2B, and SALSA® MLPA® probemix P245-B1 Microdeletion Syndromes-1A. The PCR products were analyzed by capillary electrophoresis in an automatic sequencer Hitachi 3500 and further analyzed with the Coffalyser V 1.0 software from MRC-Holland.

Microrrays for CNVs Search and Chromosome Aberrations Analysis

The DNA from the patient and his parents was hybridized on a CGH platform (Agilent Technologies). The derivative log ratio spread (DLRS) value was 0.172204. The platform included 60.000 probes. Data were analyzed with Agilent CytoGenomics 3.0.5.1 and qGenViewer, and the ADM-2 algorithm (threshold = 6.0; abs (log2ratio) = 0.25; aberrant regions had more than 4 consecutive probes).

qPCR

The expression pattern of genes of interest was assessed in blood by qRT-PCR in order to determine possible differences between the proband and his parents. Total RNA extraction and qRT-PCR were done as previously described (22). Briefly, total RNA was extracted from cells using Trizol (Sigma Aldrich) according to the manufacturer's procedure, followed by a treatment with RNase-free DNase (Roche). cDNA was synthesized from 1 μg of total RNA using a Superscript III First Strand cDNA synthesis kit (LifeTechnologies). Amounts of specific mRNAs in samples were quantified by qRT-PCR using an ABIPrism 7900 HT Detection System (Applied Biosystems) and SYBR green detection. PCR was performed in 96-well microtest plates (Applied Biosystems) with 0.5 units of Taq Polymerase (Applied Biosystems) per well and 35–40 cycles. In all experiments, mRNA amounts were normalized to the total amount of cDNA by using amplification signals for hGUSB. Each sample was determined in triplicate, and at least three independent samples of each patient were analyzed. Primer sequences and PCR conditions are listed in **Supplementary Table 1**.

Ethics approval for this research was granted by the Comité Ético del Hospital "Reina Sofía." Written informed consent was obtained from the proband's parents for publication of this case report and of any accompanying tables and images.

RESULTS

Clinical History

The proband (**Figure 1A**) is a boy born by normal delivery after 37 weeks of risk gestation because of placenta previa with bleeding and threatened abortion. The mother was a 28-year-old woman, who suffered f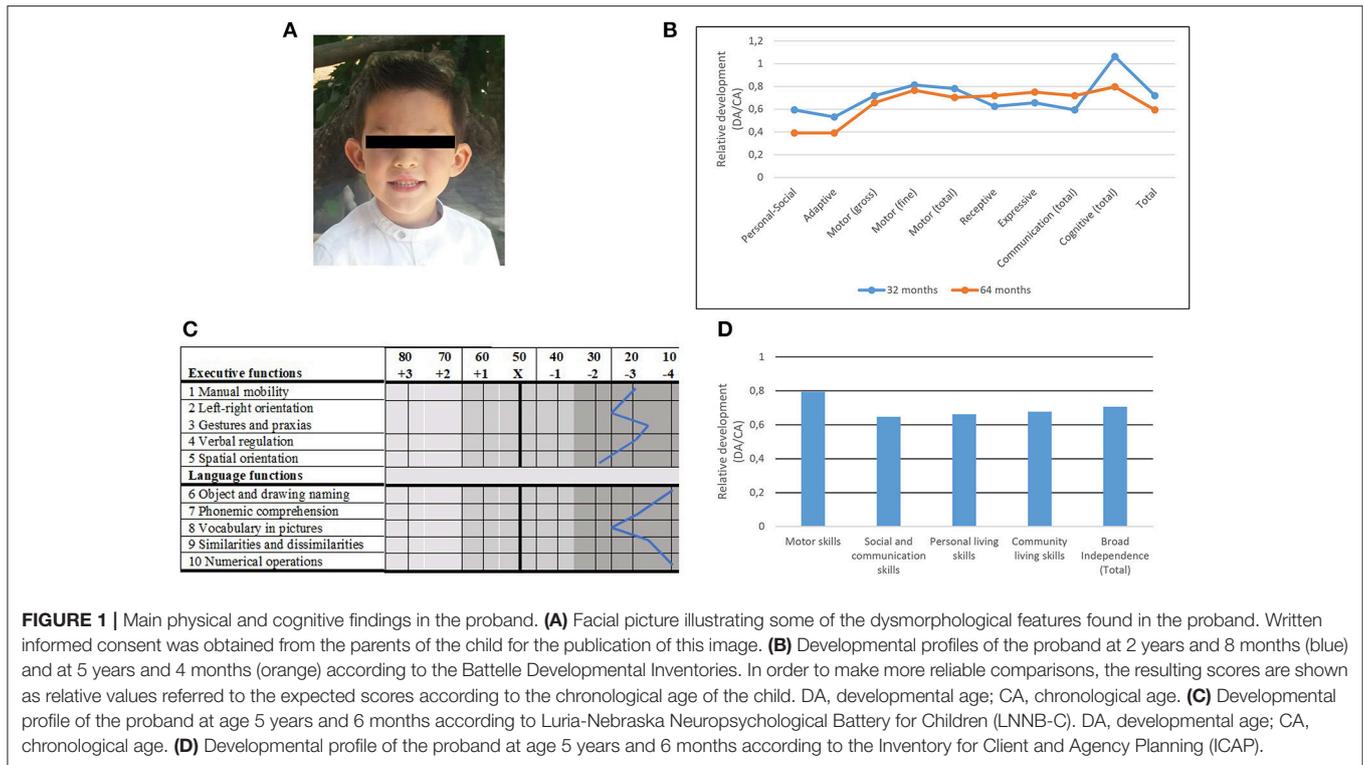rom high blood pressure and diabetes mellitus during pregnancy and who consumed anxiolytic drugs during the first trimester. At the delivery, no signs of disease were observed in the newborn. The child suffered from recurrent infantile colic during the first months of life, as well as from irritability that seriously disturbed his sleeping. At age 2 years and 3 months, his weight was 9.490 kg (percentile 1), his height 86.5 cm (percentile 18) and the occipitofrontal circumference 41.2 cm (below percentile 1). At present, when he is 6 years old, his weight is 19 kg (percentile 21), his height 117 cm (percentile 62) and the occipitofrontal circumference 50 cm (around percentile 30). Cranial magnetic resonance imaging (MRI), performed at 5 years and 7 months of age, yielded normal results. The child was set twice a tympanic drainage in conjunction with adenoidectomy (at 3 years and 4 years). At that time, the Eustachian tube was reported to be underdeveloped, although no auditory impairment was observed. One paternal cousin of the proband is dyslexic, whereas a second paternal cousin (from a different brother of the father) suffers from language developmental delay. Finally, the child exhibits facial dysmorphisms, including mild hypertelorism, small nose, and protruding ears.

Language and Cognitive Development

Developmental anomalies were first reported by the teachers of the nursery school when the proband was 2 years old. The child was unable to stand up without aid and fine and gross motor abilities seemed severely delayed. Bladder and bowel control had not been achieved either. The child only babbled and no single word was observed. Stereotypic behavior was reported. The child received early childhood intervention from age 2 years and 4 months. At age 2 years and 8 months his global development was assessed in detail with the Spanish version of the Battelle Developmental Inventories. The resulting scores were suggestive of a mild developmental delay, mostly impacting on his language, social and adaptive abilities, whereas his cognitive skills were slightly above the mean (**Figure 1B**). When he was 3 years old, the boy started attending a normal nursery school.

At age 5, the parents reported a generalized cognitive and behavioral regression. Stereotypic behaviors increased and the child exhibited lack of cognitive flexibility, occasional spatial disorientation, obsessive behavior, perseverations, unmotivated fears to darkness or load sounds, and lack of interaction with his peers at school. Although he seemed to have mastered a good vocabulary, he failed to put it into use for normal interaction. He also had problems with the conceptualization of time. Moreover, although he received speech and cognitive aid and therapy on a daily basis, at school he failed in achieving the learning objectives planned for that age.

At age 5 years and 4 months, the Spanish version of the Battelle Developmental Inventories was administered again. The obtained scores were not suggestive of a generalized regression, although the child's cognitive skills seemed more impaired than at younger ages (**Figure 1B**). The Spanish version of the Kaufman Assessment Battery for Children was also administered to evaluate his learning abilities. The most impaired areas were the general mental processing abilities (very significant deficit) and the sequential processing abilities

FIGURE 1 | Main physical and cognitive findings in the proband. **(A)** Facial picture illustrating some of the dysmorphological features found in the proband. Written informed consent was obtained from the parents of the child for the publication of this image. **(B)** Developmental profiles of the proband at 2 years and 8 months (blue) and at 5 years and 4 months (orange) according to the Battelle Developmental Inventories. In order to make more reliable comparisons, the resulting scores are shown as relative values referred to the expected scores according to the chronological age of the child. DA, developmental age; CA, chronological age. **(C)** Developmental profile of the proband at age 5 years and 6 months according to Luria-Nebrasca Neuropsychological Battery for Children (LNNB-C). DA, developmental age; CA, chronological age. **(D)** Developmental profile of the proband at age 5 years and 6 months according to the Inventory for Client and Agency Planning (ICAP).

(significant deficit), although the child scored lower than his peers in the remaining domains too. These results reinforce the view that the widespread regression reported by the parents was not for real, and that the child's poor performance in daily tasks, which becomes more evident as the boy grows older, might result from a moderate learning deficit caused by his linguistic and cognitive deficits. Additionally, in view of the behavioral problems reported by parents and teachers, but also of the scores obtained in the Batelle Developmental Inventories, the Spanish version of the Autism Diagnostic Observation Schedule- Revised (ADOS-R) was administered at the age of 5 years and 5 months. The child did not fulfill the requirements for Autism Spectrum Disorders (ASDs). He scored normally in all the tasks, including the use of non-verbal aspects of social interaction, interactive symbolic play, reciprocity during interaction and play, and understanding of, and response to, others' beliefs. The boy usually paid attention to others and was able to express his own feelings toward them and about himself. What is more, the qualitative analysis of the interactive aspects of the spontaneous conversation with his mother confirmed the absence of autism. In truth, the autistic-like behaviors initially reported by parents and teachers (avoidance behaviors, obsessive behaviors, lack of visual contact) were mostly observed when the child was placed in an unfamiliar environment and felt stressed, unsecure, or anxious. It is also possible that he suffered from some sort of compulsive-obsessive disorder.

When the child was 5 years and 6 months old, his global linguistic and cognitive profile, as well as his learning abilities

were also evaluated with the Luria-Nebraska Neuropsychological Battery for Children (LNNB-C). The obtained scores were suggestive of a deep impairment of language and executive functions (in all the assessed domains the scores corresponded to percentiles 1–2) (**Figure 1C**). Likewise, the Inventory for Client and Agency Planning (ICAP) was administered for assessing the child's adaptive behavior. The obtained scores (**Figure 1D**) were indicative of a delay of nearly 2 years in his global development, being social and communication skills the most impaired domain.

Because of the significant impairment and delay observed in the language domain, we also evaluated in detail the linguistic abilities of the proband at the age of 5 years and 6 months. The global score obtained in the Prueba de Lenguaje Oral de Navarra-Revisada (PLON-R), which evaluates speech sound production at the single word level as well as expressive abilities, was suggestive of a mild delay in the acquisition of the expressive component of language, being phonological awareness and morpho-syntax the most impaired domains. Likewise, a mild delay was suggested regarding the use of language. On the contrary, the child scored like their peers regarding the meaning of words and sentences. Concerning his receptive abilities in the domain of language, the global score obtained in the Spanish version of the Peabody Picture Vocabulary Test (PPVT-3) was suggestive of a delay of 12 months in the acquisition of the meaning of words, with an expected negative impact of his comprehension abilities.

Lastly, we examined in detail the child's language in use and his communication skills through the analysis of a

naturally-occurring interaction with his mother. The proband's speech production abilities in natural settings were severely impaired, hindering the child's intelligibility and communication success. Fluency problems included syllabified and effortful articulation, reduplication of syllables in any position of the word, and severe alterations of rhythm, use of pauses, volume, and prosody. Assimilation of specific sounds was frequent in the boy's speech, as well as simplifications of consonant clusters and deletions of weak syllables. Omissions of strong syllables (those bearing the primary stress) and of sounds in initial position were occasionally observed too. Concerning language structure and complexity, the child's mean length of utterance in words (MLUw) was 2.6, which is within the range observed in children aged between 2 years and 6 months, and 2 years and 11 months (23), and thus remarkably lower than the expected value by age. In longer utterances, the child produced morpho-syntactic errors, such as additions or substitutions of prepositions, or omissions of weak pronouns (usually, personal pronouns functioning as the object of the sentence [*me* "me" vs. *yo* "I"]), which may be due to the widespread weak-syllable deletions observed in his discourse. Finally, his interactive pragmatic skills seemed to be preserved: visual contact and joint attention mechanisms were normal, as well as patterns of turn taking and construction of collaborative turns during conversation. However, enunciative pragmatic skills (that is, principles of language use regulating how utterances are interpreted in their contexts of usage), such as the observation of maxims of quality ("be truthful and do not give information that is false or that is not supported by evidence") and relation ("be relevant, and say things that are pertinent to the discussion") (24), were slightly impaired, which might be related to his attention problems and/or intellectual disability.

Cytogenetic and Molecular Analyses

Routine cytogenetic and molecular analyses of the proband were performed when he was 2 years and 7 months old. PCR analysis of the FMR1 fragile site was normal. MLPA of subtelomeric regions was normal too. A comparative genomic hybridization array (array-CGH) identified a duplication of 2.1 Mb in the 1q21.1 region (arr[hg19] 1q21.1-q21.2(145,764,455-147,824,207)×3) (**Figure 2A**). The duplicated region corresponds to the region duplicated in Class II 1q21.1 duplication syndrome and includes 38 genes of which 13 are protein-coding (**Figure 2B**).

As noted, our proband exhibits most of the physical, behavioral, and cognitive features of the 1q21.1 microduplication syndrome (**Table 1**). The duplication was inherited from the proband's non-symptomatic mother (**Figure 3A**). The woman exhibited a verbal IQ slightly above the mean, although with peaks and valleys, which are suggestive of relative strengths in the domain of attention and auditory memory (task Digits) and of relative weaknesses in the domain of symbol processing (task Arithmetic) (**Figure 3B**). Additionally, she exhibited a very good command of verbal abilities as assessed by the BAIRES test, around percentile 90th, and normal-to-high reading abilities (**Figure 3C**). We have not assessed other features of the microduplication. The fact that she suffered from anxiety during

pregnancy might be indicative of some affectedness in other domains [see (4), Table 2], but as far as language and cognition are concerned, she performed normally.

In order to delve into the molecular causes of the speech and language problems exhibited by the child, we surveyed the literature looking for other clinical cases in which language deficits have been linked or associated to the mutation or the dysfunction of any of the genes duplicated in our proband. We found that *CHD1L* is a candidate for autism spectrum disorder (ASD) and attention deficit hyperactivity disorder (ADHD) (25, 26), whereas *PRKAB2* and *BCL9* have been associated to schizophrenia (27–30).

We also relied on DECIPHER (https://decipher.sanger.ac.uk/) to find cases of CNVs affecting chromosomal fragments smaller than the region duplicated in the child, that may help associate language dysfunctions to specific genes. We found that patients bearing microduplications not encompassing the genes *CHD1L* and *BCL9* do not usually exhibit language problems. Conversely, duplications involving *CHD1L* entail delayed speech and language development. No patient bearing a duplication affecting the *BCL9* only has been found (**Figure 4**).

Finally, we used String 10.5 (www.string-db.org) for examining potential functional links between the genes duplicated in our proband and genes important for language, in particular, known candidate genes for language disorders [dyslexia and SLI, as compiled by (31–33)] and for language evolution, as posited by Boeckx and Benítez-Burraco (34, 35) [importantly, many of these genes are also related to language dysfunction in broader cognitive disorders, particularly, autism and schizophrenia, as discussed in (36–38)] (**Supplementary Table 2**). String 10.5 predicts physical and functional associations between proteins relying on different sources (genomic context, high-throughput experiments, conserved coexpression, and text mining) (39). Several proteins were predicted to be direct interactors of some of the genes found in the duplicated fragment: FOXO1, TP53, TSC1, for PRKAB2; PARP1, for CHD1L, and CTNNB1, ARX, TL3, and TL2, for BCL9 (**Figure 5**).

In order to check the biological reliability of these potential links and the possibility that other robust candidates for language dysfunction were affected in our proband, we performed qRT-PCR analyses of selected genes using blood from the proband and his parents. We determined the expression level of (i) protein-coding genes located within the duplicated fragment (*BCL9, GJA8, GJA5, ACP6, PDZK1, PRKAB2, CHD1L, HYDIN2*), (ii) functional partners of the protein-coding genes located within the duplicated fragment with a known role in cognition/language development and/or impairment, as found in the literature and as predicted by String 10.5 (*PARP1, CTNNB1, ARX, FOXP2, TLE2, TLE3, FOXO1, TSC1*), and (iii) other strong candidates for language disorders and language evolution, as discussed in Boeckx and Benítez-Burraco (34, 35) (*AUTS2, BAZ1B, BMP2, CNTNAP2, DCDC2, DLX1, ELP4, FOXP1, GRIN2A, POU3F2, ROBO1, RUNX2, SLIT2*, and *SOX9*) (**Figure 4**). We found that 4 genes are significantly upregulated or downregulated in the proband compared to his healthy parents: *CHD1L, ACP6, TL3*, and *ROBO1* (**Figure 6** and **Supplementary Table 3**).
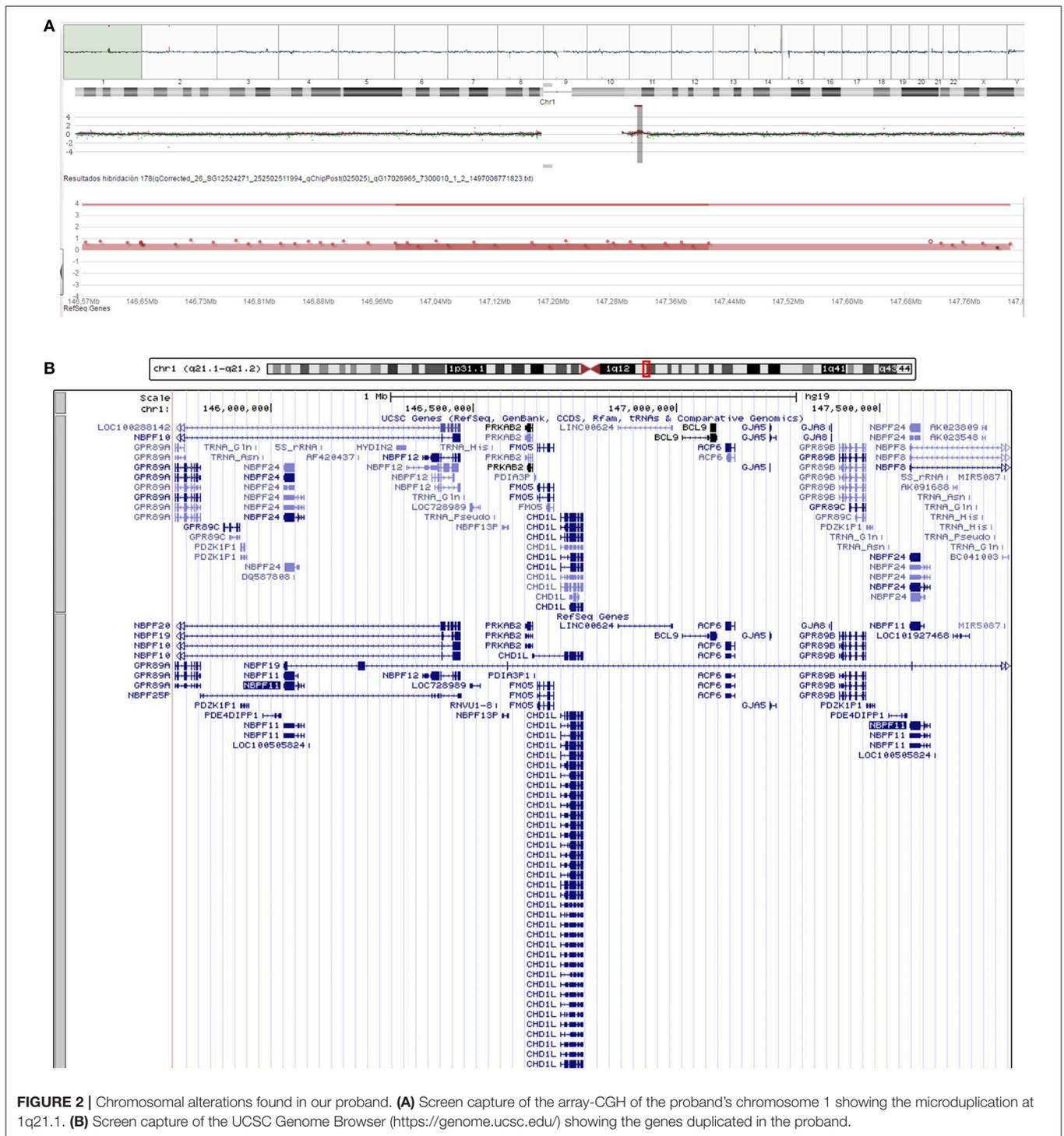
FIGURE 2 | Chromosomal alterations found in our proband. **(A)** Screen capture of the array-CGH of the proband's chromosome 1 showing the microduplication at 1q21.1. **(B)** Screen capture of the UCSC Genome Browser (https://genome.ucsc.edu/) showing the genes duplicated in the proband.

DISCUSSION

The widespread application of next generation sequencing facilities to the genetic diagnosis of people with cognitive and language disorders has resulted in a growing number of genes and chromosomal regions associated to these conditions. Nonetheless, in most cases, no robust genotype-to-phenotype correlations have been found. In this paper, we have characterized the cognitive profile of a boy bearing a microduplication in 1q21.1, with a focus on his language deficits. The boy exhibits many of the cognitive, behavioral, and physical symptoms of the 1q21.1 microduplication syndrome (**Table 1**). Regarding the language (dis)abilities of the affected people, the most detailed account is Bernier et al. (4), who found that ~10% of their

TABLE 1 | Summary table with the most relevant clinical features of our proband compared to other patients with the 1q21.1 microduplication syndrome [source: Unique (https://www.rarechromo.org/), (1, 2, 4)].

Clinical features	
COGNITIVE/BEHAVIORAL FEATURES	
Autism or autistic behaviors	
Anxiety and mood disorders	x
Attention deficit-hyperactivity disorder (ADHD)	
Mild-to-moderate mental retardation	x
Mild-to-moderate developmental delay	x
Learning disabilities	x
Language disorder	x
NEUROLOGICAL FINDINGS	
Fine motor impairment	x
Coordination disorder	x
Hypotonia	x
Articulatory problems/stuttering	x
Seizures	
Sleep disturbances	x
Agility abnormalities	x
PHYSICAL FEATURES	
Macrocephaly (occasional craniosynostosis)	
(Mild) dysmorphic facial features	
Frontal bossing	
Wider space between the eyebrows and above the nose	
Mild hypertelorism	x
Small nose	x
Prominent epicanthic folds	
Wide, flat nasal bridge	
Low-set ears	x
Stature below the average	
Scoliosis	
OTHER MEDICAL PROBLEMS	
Gastric ulcers	
Heart disease	
Minor genital anomalies (hypospadias, testicles undescended at birth)	

subjects suffered from language disorder, whereas phonological processing disorder was observed in ∼ 5% of cases. Likewise, around 10% of DECIPHER patients with a duplication within, overlapping, or encompassing the region duplicated in our proband (N = 224) are reported to suffer from language problems. Nonetheless, the linguistic evaluation of patients usually involves psychological tests only. To the best of our knowledge, this is the first time a detailed linguistic profile of a subject with the 1q21.1 microduplication has been provided.

Accordingly, we have conducted a fine-grained analysis of the language strengths and weaknesses of our subject with the 1q21.1 microduplication syndrome using a battery of specific diagnostic tools, as well as examining his language performance in spontaneous conditions, both at the structural and the functional levels. Overall, we found conclusive evidence of a global delay, which plausibly results from delayed and atypical
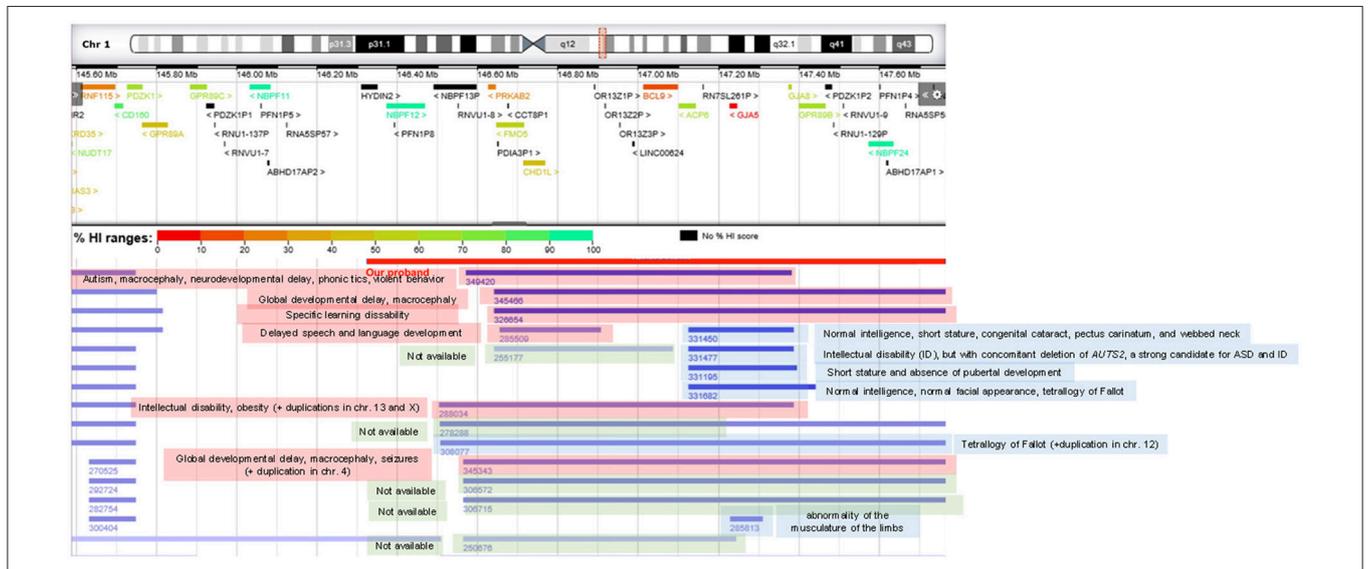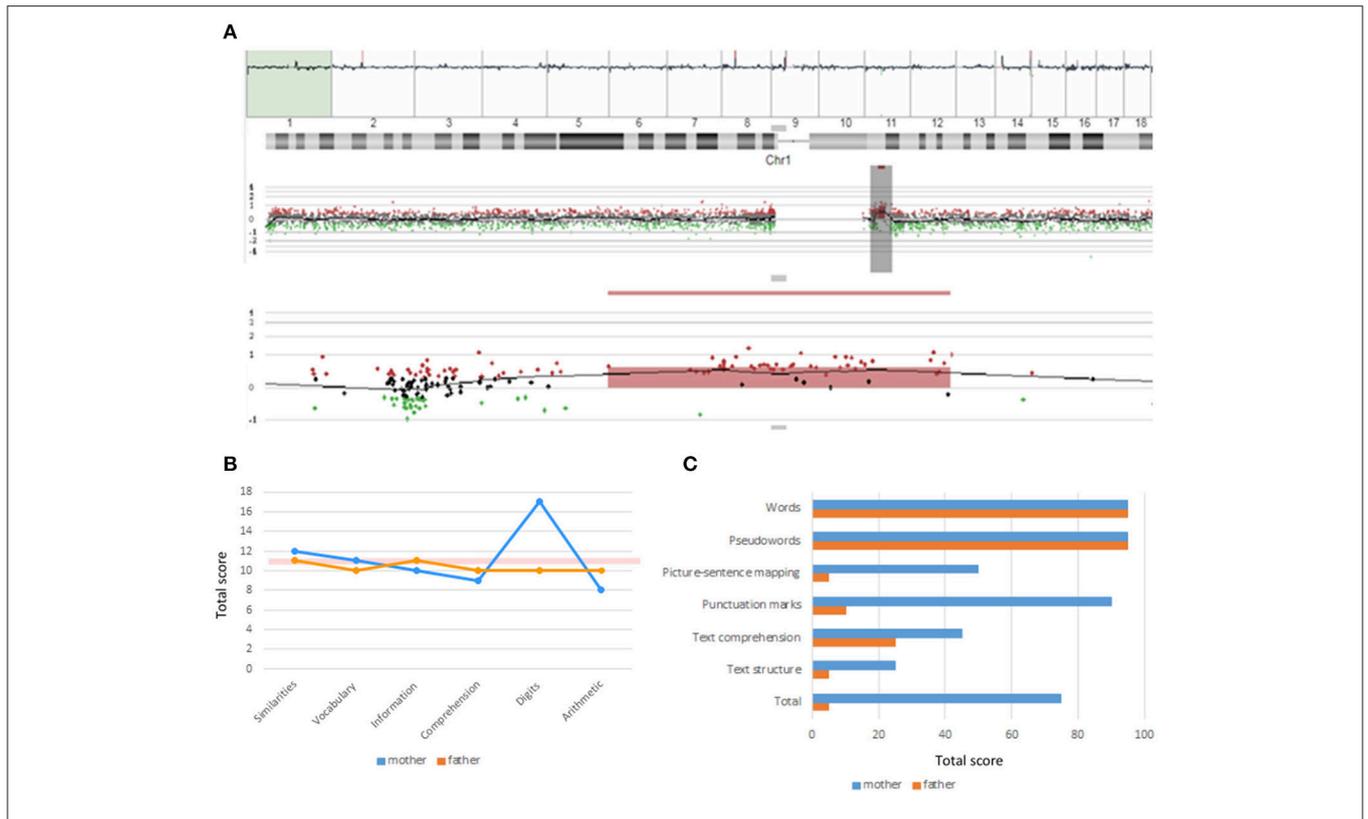
language development in the expressive and receptive domains. The proband's language impairment would be on the basis of a learning disability that precludes the normal acquisition of executive functions and adaptive behavior, hindering the child's interaction with his social environment.

The expressive problems exhibited by the child seem to be also due to altered speech processes, in the form of dysphemia (frequent repetitions or prolongations of sounds and syllables, as well as frequent pauses interrupting the flow of the discourse) and dysprosody (abnormal presentation of prosodic features, like stress, intonation, melody, or rhythmic patterns), as could be observed during the natural spontaneous interaction between the child and his mother. These problems might result in part from deficient executive functions, because we have found no evidence of pure motor speech disorders, like apraxia or dysarthria. Finally, our proband seemingly suffers from a mild pragmatic deficit, because of the also attested delay in the mastering of conversational maxims [see (40, 41) for the neurotypical population].

Regarding the molecular causes of the observed symptoms, Dolcetti et al. (3) point out that UCSC gene expression data are suggestive of the involvement of the genes *BCL9*, *GJA8*, *PDZK1*, and *PRKAB2* in the neuropsychiatric phenotype of the 1q21.1 duplication syndrome. O'Bleness et al. (6) have pointed out that the region found duplicated in our proband contains several candidates for human-lineage-specific neurodevelopmental changes, including most of the genes encoding DUF1220 protein domains (*NBPF10*, *NBPF11*, *NBPF12*, *NBPF24*), as well as *PDZK1* and *HYDIN2*. Among the genes located within the duplicated fragment, we have found that 2 genes have changed their expression level in the blood of the affected subject compared to his healthy father and his asymptomatic carrier mother. Hence, *ACP6* is significantly dowregulated, whereas *CHD1L* is significantly upregulated (**Figure 6**).

ACP6 encodes a histidine acid phosphatase involved in G protein-coupled receptor signaling and the balancing of lipid composition within the cell. Interestingly, differences in the methylation status of the gene have been correlated with differences in hearing abilities associated to age-related hearing impairment (42). Likewise, pathogenic variants in *ACP6* have been recently identified in children with cerebral visual impairment, which results from the impairment in projection and/or interpretation of the visual input in the brain, and which usually co-occurs with intellectual disability (43). The gene is expressed at low levels in the brain (**Figure 7**).

Regarding *CHD1L*, this gene encodes a DNA helicase protein involved in the relaxation following DNA damage and ultimately, in DNA repair, and its overexpression has been linked to several types of cancers (45). CHD1L is activated by PARP1 during the initial steps of cellular reprogramming (46, 47). Interestingly, *PARP1* is one of the genes that are differentially expressed among (mammalian) vocal learners (48), and its product regulates as well the dyslexia-susceptibility gene *DYX1C1*, important for neuronal migration in the developing cortex (49). We didn't find evidence of a differential expression of *PARP1* in the blood of our patient compared to his mother, who does not exhibit speech or language problems, although
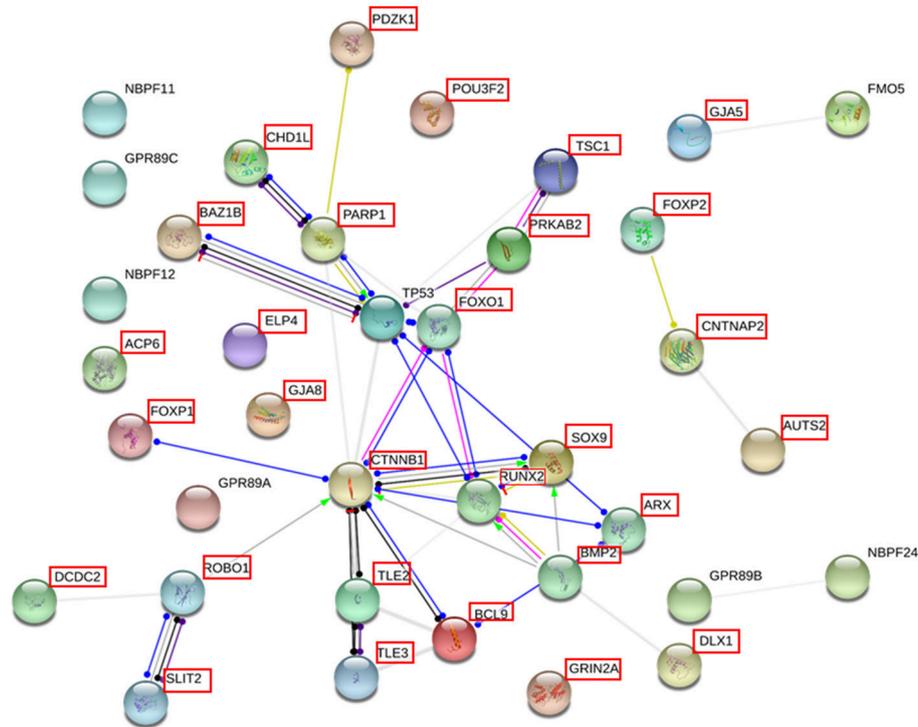
FIGURE 3 | Chromosomal alterations and linguistic profile of the proband's mother. **(A)** Screen capture of the array-CGH of the chromosome 1 of the proband's mother showing the microduplication at 1q21.1. **(B)** Developmental profile of the proband's mother according to the verbal component of the WAIS-III (for comparison, the figure includes the profile of the healthy, non-carrier father). The pink horizontal bar shows the normal values. **(C)** Reading abilities of the proband's mother according to the PROLEC-SE test (for comparison, the figure includes the profile of the healthy, non-carrier father: the low scores obtained by the father are seemingly explained by his low educational level).

FIGURE 4 | DECIPHER patients of interest bearing chromosomal duplications at 1q21.1 with similar or smaller sizes than the one found in the proband. Patients with cognitive deficits are highlighted in red, cases with no cognitive problems are highlighted in blue, whereas patients for whom no phenotypic profile is available are highlighted in green.

FIGURE 5 | Interaction network of the proteins coded for the most relevant genes duplicated in the proband. The network was drawn with String [version 10.5; (39)] license-free software (http://string-db.org/), using the molecular action visualization. It includes the products of the protein-coding genes duplicated in the subject, their potential interactors according to String, and the products of 15 strong candidates for language development/evolution. The proteins for which the expression level of the gene was measured in the blood of the patient by qRT-PCR are framed in red (the product of the pseudogene *HYDIN2* is not included in the network). Colored nodes symbolize gene/proteins included in the query (small nodes are for proteins with unknown 3D structure, while large nodes are for those with known structures). The color of the edges represent different kind of known protein-protein associations. Green: activation, red: inhibition, dark blue: binding, light blue: phenotype, dark purple: catalysis, light purple: post-translational modification, black: reaction, yellow: transcriptional regulation. Edges ending in an arrow symbolize positive effects, edges ending in a bar symbolize negative effects, whereas edges ending in a circle symbolize unspecified effects. Gray edges symbolize predicted links based on literature search (co-mentioned in PubMed abstracts). Stronger associations between proteins are represented by thicker lines. The medium confidence value was 0.0400 (a 40% probability that a predicted link exists between two enzymes in the same metabolic map in the KEGG database: http://www.genome.jp/kegg/pathway.html). The diagram only represents the potential connectivity between the involved proteins, which has to be mapped onto particular biochemical networks, signaling pathways, cellular properties, aspects of neuronal function, or cell-types of interest.

the gene is slightly upregulated in both carriers compared to the healthy father (**Supplementary Table 3**). *CHD1L* is also a candidate for ASD and ADHD (25, 26). Contrary to what is observed in patients with microdeletions of the 1q21.1 region, as the size of duplications encompassing *CHD1L* increases, ASD severity also increases, whereas nonverbal IQ remains unaffected (25). Interestingly, DECIPHER patient 285509, who bears a microduplication affecting only to part of the gene *FMO5* and the whole sequence of *CHD1L*, exhibits delayed speech and language development (**Figure 3**). *CHD1L* is highly expressed in different brain areas, particularly in the cerebellum, where it is significantly upregulated around birth (**Figure 7**). The cerebellum is involved in almost every aspect of speech and language processing, from syntax processing to phonological and semantic verbal fluency, verbal working memory and speech perception and planning (50). Moreover, it is involved as well in most of the processes that are impaired in ASD, from language and communication to social interaction and behavior, and cerebellar damage has been found to contribute to the autistic phenotype (51).

We have found in the literature evidence of the potential involvement of *BCL9* in the language deficits observed in patients with the 1q21.1 duplication syndrome. Similarly to *CHD1L*, *BCL9* is also highly expressed in several parts of the brain, and particularly, in the cerebellum (**Figure 7**). *BCL9* plays a significant role in embryonic body axis development (52), as part of the Wnt/β-catenin pathway (53). Specifically, BCL9 promotes transcriptional activity and nuclear retention of β-catenin (54). β-catenin is encoded by *CTNNB1*, a gene related to cognitive disorders entailing language deficits, particularly to schizophrenia (55, 56), but also to genes that have been hypothesized to contribute to language evolution in our species [see (34) for details]. *BCL9* itself is a candidate for schizophrenia (27) and interacts as well with genes important for language development and evolution. To begin with, in mice, *Bcl9* is a target of Foxp2, the renowned 'language gene' (57). Additionally, BCL9 interacts with ARX to regulate canonical Wnt signaling (58). *ARX* is also a target of FOXP2 (57, 59) and regulates the migration of GABAergic interneurons (60), as well as
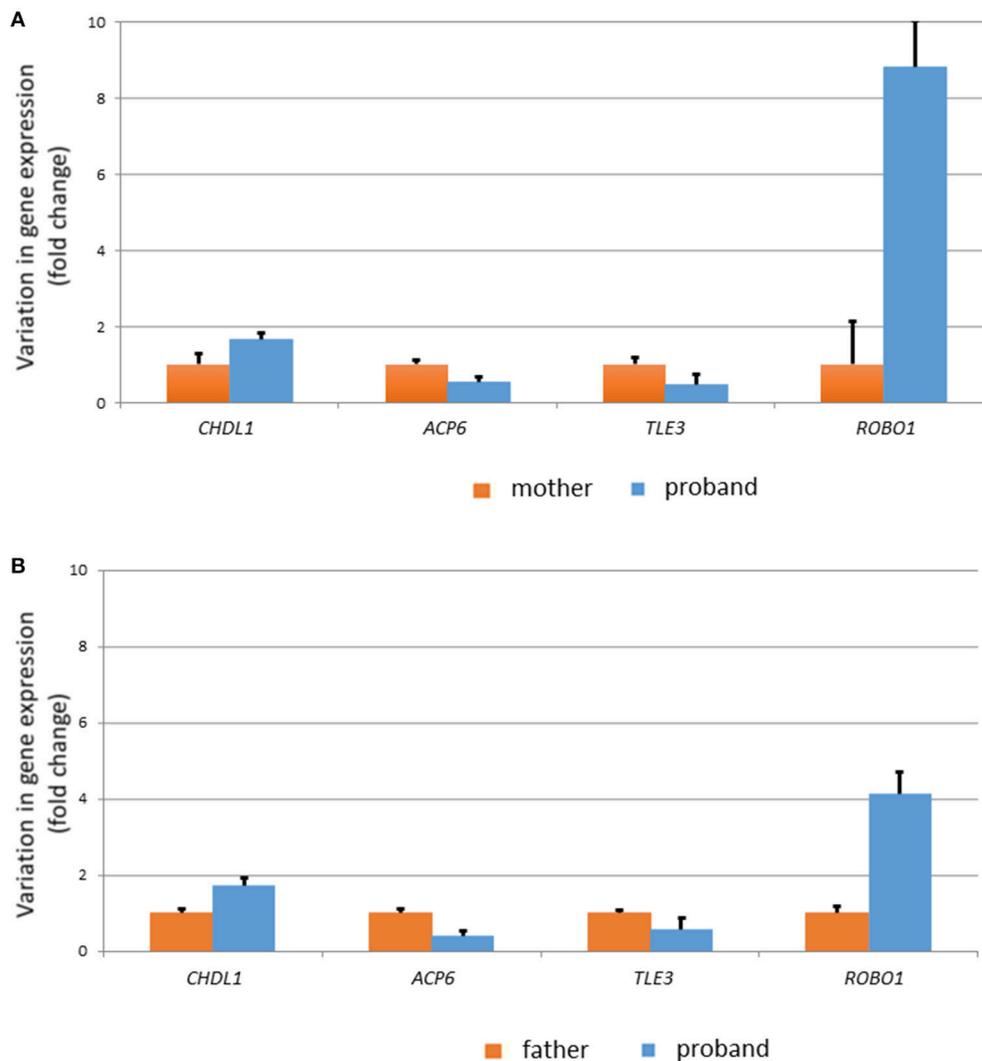
FIGURE 6 | Expression profile of genes of interest in the blood of the proband. Only the genes showing significant ($p < 0.05$) differences with his healthy carrier mother **(A)** and his healthy non-carrier father **(B)** are displayed. Variation in gene expression levels is expressed as fold changes.

dopaminergic neuron development (61). In humans, mutations of *ARX* result in mental retardation and interneuronopathies (62), as well as in agenesis of the corpus callosum (63), a condition entailing deficits in problem solving and social skills that often fall within the autistic spectrum (64, 65). Finally, according to the Reactome Pathway Database (https://reactome.org/), BCL9 interacts with both TLE2 and TLE3 as part of the pathway involved in the activation and deactivation of the β-catenin transactivating complex. Both genes encode corepressors of gene expression and play crucial functions in brain development (66, 67). *TLE3* is a target of FOXP2 in the inferior prefrontal cortex (68). Interestingly, we have found that *TL3* is significantly downregulated in the blood of our proband compared to their asymptomatic parents. The gene is upregulated in the embryonic brain, particularly, in the neocortex (**Figure 7**).

Finally, among the set of robust candidates for language disorders and language evolution, we have found that *ROBO1* is strongly upregulated in the blood of our subject compared to his healthy father and particularly, to his asymptomatic carrier mother. *ROBO1* encodes an axon guidance receptor involved in interneuron migration and axon tract development in the forebrain, in particular, of thalamocortical axons implicated in diverse cognitive functions, consciousness, and alertness (69, 70). Interestingly, *ROBO1* regulates interaural interaction in auditory pathways (71), and was co-opted for vocalization in songbirds as a component of the motor song output nucleus (72). Overall, as part of the SLIT/ROBO pathway, *ROBO1* has been hypothesized to be a crucial candidate for speech development and evolution [see (34) for details]. The gene is a candidate for clinical conditions entailing speech problems, like dyslexia and speech-sound disorder (73, 74). According to its known role in
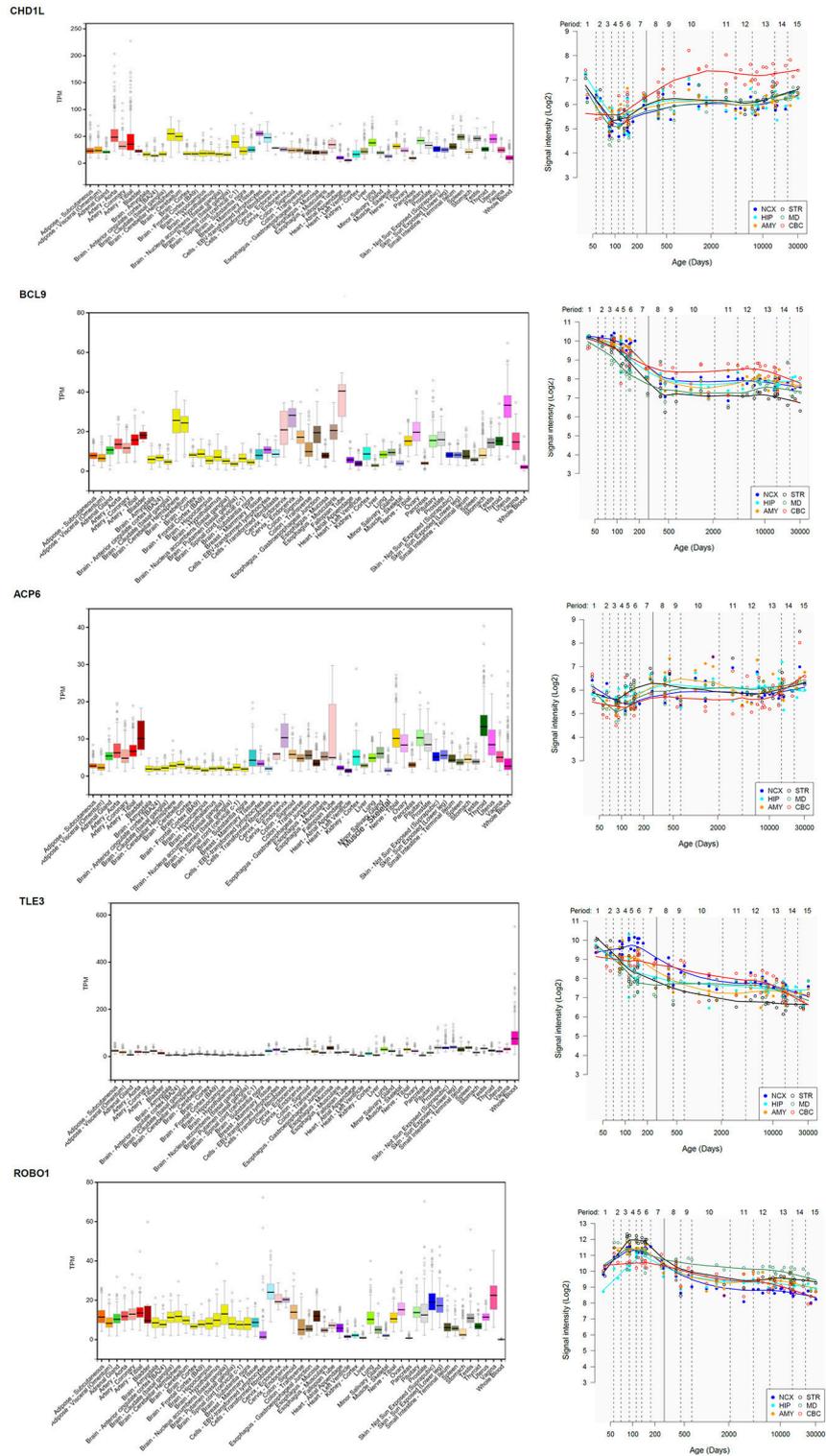
FIGURE 7 | Expression pattern in the body and the brain of genes of interest in the proband. **(Left)** Expression levels of the gene in different tissues according to the Genotype-Tissue Expression (GTEx) project (44). Statistical analysis and data interpretation was performed by The GTEx Consortium Analysis Working Group. Data was provided by the GTEx LDACC at The Broad Institute of MIT and Harvard. TPM, transcripts per million. **(Right)** Brain expression profile of the gene across development. The expression data are from the Human Brain Transcriptome Database (http://hbatlas.org/). Six different brain regions are considered: the cerebellar cortex (CBC), the mediodorsal nucleus of the thalamus (MD), the striatum (STR), the amygdala (AMY), the hippocampus (HIP), and 11 areas of neocortex (NCX).

vocalization, the gene is highly expressed in several subcortical regions (**Figure 7**). No DECIPHER patient bears a duplication of *ROBO1* only, so it is difficult to ascertain the effects in humans of an excessive dosage of the gene. Nonetheless, it is interesting that increased expression of *Robo1* in rats reactivates astrocytes after transient forebrain ischemia (75), whereas the increase of Robo1 function in mice restores normal axon guidance in several neuron populations (76).

In conclusion, although the exact molecular causes of the cognitive and linguistic profile of our proband remain to be fully elucidated, we hypothesize that his distinctive features might result, to a large extent, from the overexpression of *CHD1L*. We further hypothesize that his problems with speech might result from subtle changes in the expression level of some of the genes involved in the externalization of language as part of the FOXP2/ROBO/SLIT regulatory networks, in particular, of *ROBO1*, but also some of the FOXP2's targets, like *TLE2*. We wish to acknowledge two potential caveats of our study. First, we have determined changes in the expression level of genes of interest in the blood, whereas the phenotype under scrutiny has a brain-based locus of impairment. Nonetheless, as shown in **Figure 7**, our candidates are expressed in the blood and the brain at similar levels. Moreover, several studies point to a significant overlapping in gene expression patterns between both tissues, ranging from 20% (77, 78) to 55% (79). Accordingly, we regard that our findings in the blood can be confidently extrapolated to the brain. Second, because the child was born following a multi-factorial high-risk pregnancy that involved maternal diabetes and exposure to anxiolytic drugs during embryonic and early fetal development, we cannot rule out the possibility that these environmental influences contributed to his language problems. Such an effect is really difficult to estimate. Although recent research has acknowledged diabetes as a robust predictor of language delay in twins (80), the available literature is not fully conclusive [see (81) for a meta-review], and in fact, most children born from mothers with diabetes are not significantly affected if diabetes is properly controlled during pregnancy [see (82) for discussion]. The effect of anxiolytics is perhaps more elusive. Considering the phenotype of our proband, it is interesting that prenatal administration of diazepam delays the maturation of temporal acuity in the rat auditory system (83). This effect might result in altered speech perception in humans. However, no association has been found between anxiolytic exposure during pregnancy and long-term language competence (84). Accordingly, considering as well all the evidence discussed in the paper, we think that the language problems observed in our proband are mostly due to the microduplication in chromosome 1. We hope our findings contribute to a better understanding of the genetic underpinnings of human language.

AUTHOR CONTRIBUTIONS

AB-B conceived and planned the paper, analyzed the data, and wrote the manuscript. MB-M and IE-P performed the cytogenetic analyses and the microarrays, and analyzed the data. MF-U analyzed in detail the proband's language in use speech and revised the manuscript. RT-R and SR-P performed the qRT-PCRs and analyzed the data. MJ-R assessed the cognitive and linguistic problems of the child, analyzed the data, and revised the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fped.2018.00163/full#supplementary-material

Supplementary Table 1 | List of primers used for qRT-PCR analyses.

Supplementary Table 2 | The whole set of genes important for language used for the *in silico* analysis with String 10.5.

Supplementary Table 3 | Gene expression levels in the proband and his parents in selected genes of interest.

REFERENCES

1. Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, Sahoo T, et al. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat Genet*. (2008) **40**:1466–71. doi: 10.1038/ng.279
2. Mefford HC, Sharp AJ, Baker C, Itsara A, Jiang Z, Buysse K, et al. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N Engl J Med*. (2008) **359**:1685–99. doi: 10.1056/NEJMoa0805384
3. Dolcetti A, Silversides CK, Marshall CR, Lionel AC, Stavropoulos DJ, Scherer SW, et al. 1q21.1 Microduplication expression in adults. *Genet Med*. (2013) **15**:282–9. doi: 10.1038/gim.2012.129
4. Bernier R, Steinman KJ, Reilly B, Wallace AS, Sherr EH, Pojman N, et al. Clinical phenotype of the recurrent 1q21.1 copy-number variant. *Genet Med*. (2016) **18**:341–9. doi: 10.1038/gim.2015.78
5. Haldeman-Englert CR, Jewett T. 1q21.1 recurrent microdeletion. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Mefford HC, Stephens K, Amemiya A, Ledbetter N, editors. *GeneReviews®*. Seattle, WA: University of Washington (2015).
6. O'Bleness M, Searles VB, Varki A, Gagneux P, Sikela JM. Evolution of genetic and genomic features unique to the human lineage. *Nat Rev Genet*. (2012) 13:853–66. doi: 10.1038/nrg3336
7. De La Cruz López MV, González Criado M. *Adaptación Española del Inventario de Desarrollo Battelle*. Madrid: TEA (2011).
8. Manga D, Ramos F. *Luria-Inicial*. Evaluación *Neuropsicológica* en la *Edad Preescolar*. Madrid: TEA Ediciones (2006).
9. Conde E, Seisdedos N. *K-ABC: Batería de Evaluación de Kaufman Para Niños*. Manual de interpretación. Madrid: TEA Ediciones (2009).
10. Montero D. *Evaluación de la Conducta Adaptativa en Personas con Discapacidades*. Adaptación y validación del ICAP. Bilbao: Mensajero (1996).
11. Aguinaga G, Armentia M, Fraile A, Olangua P, Uriz N. *PLON.-R Prueba de Lenguaje Oral de Navarra*. Madrid: TEA (2005).
12. Dunn L, Dunn LM, Arribas D. *Peabody, Test de Vocabulario en Imágenes*. Madrid: TEA Ediciones (2006).

13. MacWhinney B. *The CHILDES Project: Tools for Analyzing Talk. Volume I: Transcription Format and Programs*. Mahwah, NJ: Lawrence Erlbaum (2000).

14. Verschueren J. *Understanding Pragmatics*. London: Edward Arnold (1999).

15. Fernández-Urquiza M, Viejo A, Cortiñas S, Huelmo J, Medina B, García I, et al. *Perfiles pragmáticos Comparados de síndromes genéticos neuroevolutivos* (S. Williams S. Down y S. X Frágil). In Diéguez-Vide F, editor. Temas de Lingüística Clínica. Barcelona: Horsori Editorial (2015). p. 89–90.

16. Fernández-Urquiza M, Diez-Itza E, Cortiñas S. PREP-CORP: sistema de etiquetado pragmático de corpus clínicos de lengua oral. In: Fernández López MC, Martí Sánchez M, Ruiz Martínez AM, editor. *Aplicaciones de la Lingüística*. Alcalá de Henares: Universidad de Alcalá (2017). p. 167–83.

17. Díez-Itza E, Martínez V, Pérez V, Fernández-Urquiza M. (2018). Explicit oral narrative intervention for students with Williams syndrome. *Front Psychol.* (2015) 8:2337. doi: 10.3389/fpsyg.2017.02337

18. Conti-Ramsden. G CLAN (Computerized Language Analysis). *Child Lang Teach Ther.* (1996) 12:345–9. doi: 10.1177/026565909601200308

19. Wechsler D. *WAIS – III. Escala de Inteligencia Para Adultos de Wechsler-Tercera Edición*. Buenos Aires: Paidós (2002).

20. Cortada de Kohan N. *BAIRES. Test de Aptitud Verbal "Buenos Aires"*. Madrid: TEA Ediciones (2004).

21. Ramos JL, Cuetos F. *PROLEC-SE. Evaluación de los Procesos Lectores. Manual*. Madrid: TEA Ediciones (2011).

22. Torres-Ruiz R, Martinez-Lage M, Martín MC, García A, Bueno C, Castaño J, et al. Efficient recreation of t(11;22) EWSR1-FLI1(+) in human stem cells using CRISPR/Cas9. *Stem Cell Rep.* (2017) 8:1408–20. doi: 10.1016/j.stemcr.2017.04.014

23. Rice ML, Smolik F, Perpich D, Thompson T, Rytting N, Blossom M. Mean Length of Utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. *J Speech Lang Hear Res.* (2010) 53:333–49. doi: 10.1044/1092-4388(2009/08-0183)

24. Grice HP. Logic and conversation. In: Cole P, Morgan J, editors. *Syntax and Semantics Volume 3: Speech Acts*. New York, NY: Academic Press (1975). p. 41–58.

25. Girirajan S, Dennis MY, Baker C, Malig M, Coe BP, Campbell CD, et al. Refinement and discovery of new hotspots of copy-number variation associated with autism spectrum disorder. *Am J Hum Genet.* (2013) 92:221–37. doi: 10.1016/j.ajhg.2012.12.016

26. Kim DS, Burt AA, Ranchalis JE, Wilmot B, Smith JD, Patterson KE, et al. Sequencing of sporadic Attention-Deficit Hyperactivity Disorder. (ADHD) identifies novel and potentially pathogenic de novo variants and excludes overlap with genes associated with autism spectrum disorder. *Am J Med Genet B Neuropsychiatr Genet.* (2017) 174:381–9. doi: 10.1002/ajmg.b.32527

27. Li J, Zhou G, Ji W, Feng G, Zhao Q, Liu J, et al. Common variants in the BCL9 gene conferring risk of schizophrenia. *Arch Gen Psychiatry* (2011) 68:232–40. doi: 10.1001/archgenpsychiatry.2011.1

28. Hosak L. New findings in the genetics of schizophrenia. *World J Psychiatry* (2013) 3:57–61. doi: 10.5498/wjp.v3.i3.57

29. Xu C, Aragam N, Li X, Villla EC, Wang L, Briones D, et al. BCL9 and C9orf5 are associated with negative symptoms in schizophrenia: meta-analysis of two genome-wide association studies. *PLoS ONE* (2013) 8:e51674. doi: 10.1371/journal.pone.0051674

30. Kimura H, Tanaka S, Kushima I, Koide T, Banno M, Kikuchi T, et al. Association study of BCL9 gene polymorphism rs583583 with schizophrenia and negative symptoms in Japanese population. *Sci Rep.* (2015) 5:15705. doi: 10.1038/srep15705

31. Paracchini S, Diaz R, Stein J. Advances in dyslexia genetics—new insights into the role of brain asymmetries. In: Friedmann T, Dunlap JC, Goodwin SF, editors. *Advances in Genetics*, Vol 96. London: Academic Press (2016). p. 53–97.

32. Pettigrew KA, Frinton E, Nudel R, Chan MTM, Thompson P, Hayiou-Thomas ME, et al. Further evidence for a parent-of-origin effect at the NOP9 locus on language-related phenotypes. *J Neurodev Disord.* (2016) 8:24. doi: 10.1186/s11689-016-9157-6

33. Chen XS, Reader RH, Hoischen A, Veltman JA, Simpson NH, Francks C, et al. Next-generation DNA sequencing identifies novel gene variants and pathways involved in specific language impairment. *Sci Rep.* (2017) 7:46105. doi: 10.1038/srep46105

34. Boeckx C, Benítez-Burraco A. Globularity and language-readiness: generating new predictions by expanding the set of genes of interest. *Front Psychol.* (2014) 5:1324. doi: 10.3389/fpsyg.2014.01324

35. Benítez-Burraco A, Boeckx C. Possible functional links among brain- and skull-related genes selected in modern humans. *Front Psychol.* (2015) 6:794. doi: 10.3389/fpsyg.2015.00794

36. Benítez-Burraco A, Murphy E. The oscillopathic nature of language deficits in autism: from genes to language evolution. *Front Hum Neurosci.* (2016) 10:120. doi: 10.3389/fnhum.2016.00120

37. Murphy E, Benítez-Burraco A. Language deficits in schizophrenia and autism as related oscillatory connectomopathies: an evolutionary account. *Neurosci Biobehav Rev.* (2016a) 83:742–64. doi: 10.1016/j.neubiorev.2016.07.029

38. Murphy E, Benítez-Burraco A. Bridging the gap between genes and language deficits in schizophrenia: an oscillopathic approach. *Front Hum Neurosci.* (2016b) 10:422. doi: 10.3389/fnhum.2016.00422

39. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* (2015) 43:D447–52. doi: 10.1093/nar/gku1003

40. Eskritt M, Whalen J, Lee K. Preschoolers can recognize violations of the Gricean maxims. *Br J Dev Psychol.* (2008) 26:435–43. doi: 10.1348/026151007X253260

41. Okanda M, Asada K, Moriguchi Y, Itakura S Understanding violations of Gricean maxims in preschoolers and adults. *Front Psychol.* (2015) 6:901. doi: 10.3389/fpsyg.2015.00901

42. Wolber LE, Steves CJ, Tsai PC, Deloukas P, Spector TD, Bell JT, et al. Epigenome-wide DNA methylation in hearing ability: new mechanisms for an old problem. *PLoS ONE* (2014) 9:e105729. doi: 10.1371/journal.pone.0105729

43. Bosch DG, Boonstra FN, de Leeuw N, Pfundt R, Nillesen WM, de Ligt J, et al. Novel genetic causes for cerebral visual impairment. *Eur J Hum Genet.* (2016) 24:660–5. doi: 10.1038/ejhg.2015.186

44. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet.* (2013) 45:580–5. doi: 10.1038/ng.2653

45. Cheng W, Su Y, Xu F. CHD1L: a novel oncogene. *Mol Cancer* (2013) 12:170. doi: 10.1186/1476-4598-12-170

46. Gottschalk AJ, Timinszky G, Kong SE, Jin J, Cai Y, Swanson SK, et al. Poly(ADP-ribosyl)ation directs recruitment and activation of an ATP-dependent chromatin remodeler. *Proc Natl Acad Sci USA.* (2009) 106:13770–4. doi: 10.1073/pnas.0906920106

47. Jiang BH, Chen WY, Li HY, Chien Y, Chang WC, Hsieh PC, et al. CHD1L regulated PARP1-driven pluripotency and chromatin remodeling during the early-stage cell reprogramming. *Stem Cells* (2015) 33:2961–72. doi: 10.1002/stem.2116

48. Wang R. *Dissecting the Genetic Basis of Convergent Complex Traits Based on Molecular Homoplasy*. Ph.D., Duke University (2011).

49. Tapia-Páez I, Tammimies K, Massinen S, Roy AL, Kere J. The complex of TFII-I, PARP1, and SFPQ proteins regulates the DYX1C1 gene implicated in neuronal migration and dyslexia. *FASEB J.* (2008) 22:3001–9. doi: 10.1096/fj.07-104455

50. Mariën P, Ackermann H, Adamaszek M, Barwood CH, Beaton A, Desmond J, et al. Consensus paper: language and the cerebellum: an ongoing enigma. *Cerebellum* (2014) 13:386–410. doi: 10.1007/s12311-013-0540-5

51. Hampson DR, Blatt GJ. Autism spectrum disorders and neuropathology of the cerebellum. *Front Neurosci.* (2015) 9:420. doi: 10.3389/fnins.2015.00420

52. Kennedy MW, Cha SW, Tadjuidje E, Andrews PG, Heasman J, Kao KR. A co-dependent requirement of xBcl9 and Pygopus for embryonic body axis development in Xenopus. *Dev Dyn.* (2010) 239:271–83. doi: 10.1002/dvdy.22133

53. van Tienen LM, Mieszczanek J, Fiedler M, Rutherford TJ, Bienz M. Constitutive scaffolding of multiple Wnt enhanceosome components by Legless/BCL9. *Elife* (2017) 6:e20882. doi: 10.7554/eLife.20882

54. Kramps T, Peter O, Brunner E, Nellen D, Froesch B, Chatterjee S, et al. Wnt/wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear beta-catenin-TCF complex. *Cell* (2002) 109:47–60. doi: 10.1016/S0092-8674(02)00679-7

55. Pedrosa E, Shah A, Tenore C, Capogna M, Villa C, Guo X, et al. β-catenin promoter ChIP-chip reveals potential schizophrenia and bipolar disorder gene network. *J Neurogenet.* (2010) 24:182–93. doi: 10.3109/01677063.2010.495182

56. Levchenko A, Davtian S, Freylichman O, Zagrivnaya M, Kostareva A, Malashichev Y. Beta-catenin in schizophrenia: possibly deleterious novel mutation. *Psychiatry Res.* (2015) **228**:843–8. doi: 10.1016/j.psychres.2015.05.014

57. Vernes SC, Oliver PL, Spiteri E, Lockstone HE, Puliyadi R, Taylor JM, et al. Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. *PLoS Genet.* (2011) **7**:e1002145. doi: 10.1371/journal.pgen.1002145

58. Cho IT, Lim Y, Golden JA, Cho G. Aristaless Related Homeobox. (ARX) interacts with β-Catenin, BCL9, and P300 to regulate canonical Wnt signaling. *PLoS ONE* (2017) **12**:e0170282. doi: 10.1371/journal.pone.0170282

59. Konopka G, Bomar JM, Winden K, Coppola G, Jonsson ZO, Gao F, et al. Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature* (2009) **462**:213–7. doi: 10.1038/nature08549

60. Colasante G, Collombat P, Raimondi V, Bonanomi D, Ferrai C, Maira M, et al. Arx is a direct target of Dlx2 and thereby contributes to the tangential migration of GABAergic interneurons. *J Neurosci.* (2008) **28**:10674–86. doi: 10.1523/JNEUROSCI.1283-08.2008

61. Filippi A, Jainok C, Driever W. Analysis of transcriptional codes for zebrafish dopaminergic neurons reveals essential functions of Arx and Isl1 in prethalamic dopaminergic neuron development. *Dev Biol.* (2012) **369**:133–49. doi: 10.1016/j.ydbio.2012.06.010

62. Quillé ML, Carat S, Quéméner-Redon S, Hirchaud E, Baron D, Benech C, et al. High-throughput analysis of promoter occupancy reveals new targets for Arx, a gene mutated in mental retardation and interneuronopathies. *PLoS One* (2011) **6**:e25181. doi: 10.1371/journal.pone.0025181

63. Proud VK, Levine C, Carpenter NJ. New X-linked syndrome with seizures, acquired micrencephaly, and agenesis of the corpus callosum. *Am J Med Genet.* (1992) **43**:458–66. doi: 10.1002/ajmg.1320430169

64. Lau YC, Hinkley LB, Bukshpun P, Strominger ZA, Wakahiro ML, Baron-Cohen S, et al. Autism traits in individuals with agenesis of the corpus callosum. *J Autism Dev Disord.* (2013) **43**:1106–18. doi: 10.1007/s10803-012-1653-2

65. Siffredi V, Anderson V, Leventer RJ, Spencer-Smith MM. Neuropsychological profile of agenesis of the corpus callosum: a systematic review. *Dev Neuropsychol.* (2013) **38**:36–57. doi: 10.1080/87565641.2012.721421

66. Kobayashi M, Nishikawa K, Suzuki T, Yamamoto M. The homeobox protein Six3 interacts with the Groucho corepressor and acts as a transcriptional repressor in eye and forebrain formation. *Dev Biol.* (2001) **232**:315–26. doi: 10.1006/dbio.2001.0185

67. Roth M, Bonev B, Lindsay J, Lea R, Panagiotaki N, Houart C, et al. FoxG1 and TLE2 act cooperatively to regulate ventral telencephalon formation. *Development* (2010) **137**:1553–62. doi: 10.1242/dev.044909

68. Spiteri E, Konopka G, Coppola G, Bomar J, Oldham M, Ou J, et al. Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *Am J Hum Genet.* (2007) **81**:1144–57. doi: 10.1086/522237

69. Andrews W, Liapi A, Plachez C, Camurri L, Zhang J, Mori S, et al. Robo1 regulates the development of major axon tracts and interneuron migration in the forebrain. *Development* (2006) **133**:2243–52. doi: 10.1242/dev.02379

70. López-Bendito G, Flames N, Ma L, Fouquet C, Di Meglio T, Chedotal A, et al. Robo1 and Robo2 cooperate to control the guidance of major axonal tracts in the mammalian forebrain. *J Neurosci.* (2007) **27**:3395–407. doi: 10.1523/JNEUROSCI.4605-06.2007

71. Lamminmäki S, Massinen S, Nopola-Hemmi J, Kere J, Hari R. Human ROBO1 regulates interaural interaction in auditory pathways. *J Neurosci.* (2012) **32**:966–71. doi: 10.1523/JNEUROSCI.4007-11.2012

72. Wang R, Chen CC, Hara E, Rivas MV, Roulhac PL, Howard JT, et al. Convergent differential regulation of SLIT-ROBO axon guidance genes in the brains of vocal learners. *J Comp Neurol.* (2015) **523**:892–906. doi: 10.1002/cne.23719

73. Hannula-Jouppi K, Kaminen-Ahola N, Taipale M, Eklund R, Nopola-Hemmi J, Kääriäinen H, et al. The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. *PLoS Genet.* (2005) **1**:e50. doi: 10.1371/journal.pgen.0010050

74. Mascheretti S, Riva V, Giorda R, Beri S, Lanzoni LF, Cellino MR, et al. KIAA0319 and ROBO1: evidence on association with reading and pleiotropic effects on language and mathematics abilities in developmental dyslexia. *J Hum Genet.* (2014) **59**:189–97. doi: 10.1038/jhg.2013.141

75. Park JH, Pak HJ, Riew TR, Shin YJ, Lee MY. Increased expression of Slit2 and its receptors Robo1 and Robo4 in reactive astrocytes of the rat hippocampus after transient forebrain ischemia. *Brain Res.* (2016) **1634**:45–56. doi: 10.1016/j.brainres.2015.12.056

76. Marcos-Mondéjar P, Peregrín S, Li JY, Carlsson L, Tole S, López-Bendito G. The lhx2 transcription factor controls thalamocortical axonal guidance by specific regulation of robo1 and robo2 receptors. *J Neurosci.* (2012) **32**:4372–85. doi: 10.1523/JNEUROSCI.5851-11.2012

77. Sullivan PF, Fan C, Perou CM. Evaluating the comparability of gene expression in blood and brain. *Am J Med Genet B Neuropsychiatr Genet.* (2006) **141B**:261–8. doi: 10.1002/ajmg.b.30272

78. Rollins B, Martin MV, Morgan L, Vawter MP. Analysis of whole genome biomarker expression in blood and brain. *Am J Med Genet B Neuropsychiatr Genet.* (2010) **153B**:919–36. doi: 10.1002/ajmg.b.31062

79. Witt SH, Sommer WH, Hansson AC, Sticht C, Rietschel M, Witt CC. Comparison of gene expression profiles in the blood, hippocampus and prefrontal cortex of rats. *In Silico Pharmacol.* (2013) **1**:15. doi: 10.1186/2193-9616-1-15

80. Taylor CL, Rice ML, Christensen D, Blair E, Zubrick SR. Prenatal and perinatal risks for late language emergence in a population-level sample of twins at age 2. *BMC Pediatr.* (2018) **18**:41. doi: 10.1186/s12887-018-1035-9

81. Adane AA, Mishra GD, Tooth LR. Diabetes in pregnancy and childhood cognitive development: a systematic review. *Pediatrics* (2016) **137**:e20154234. doi: 10.1542/peds.2015-4234

82. Perna R, Loughan AR, Le J, Tyson K. Gestational diabetes: long-term central nervous system developmental and cognitive sequelae. *Appl Neuropsychol Child.* (2015) **4**:217–20. doi: 10.1080/21622965.2013.874951

83. Kellogg C, Ison JR, Miller RK. Prenatal diazepam exposure: effects on auditory temporal resolution in rats. *Psychopharmacology* (1983) **79**:332–7. doi: 10.1007/BF00433413

84. Odsbu I, Skurtveit S, Selmer R, Roth C, Hernandez-Diaz S, Handal M. Prenatal exposure to anxiolytics and hypnotics and language competence at 3 years of age. *Eur J Clin Pharmacol.* (2015) **71**:283–91. doi: 10.1007/s00228-014-1797-4