Received: 5 October 2020 | Revised: 18 February 2021 | Accepted: 19 February 2021

DOI: 10.1002/mgg3.1647

CLINICAL REPORT

WILEY

Deletion of conserved non-coding sequences downstream from *NKX2-1*: A novel disease-causing mechanism for benign hereditary chorea

Jun Liao[1,2] | **Keith A. Coffman**[3] | **Joseph Locker**[4] | **Quasar S. Padiath**[2] | **Bruce Nmezi**[2] | **Robyn A. Filipink**[3] | **Jie Hu**[1,5] | **Malini Sathanoori**[1,4,5] | **Suneeta Madan-Khetarpal**[3] | **Marianne McGuire**[3] | **Allison Schreiber**[6] | **Rocio Moran**[6] | **Neil Friedman**[7] | **Lori Hoffner**[8] | **Aleksandar Rajkovic**[2,4,5,8] | **Svetlana A. Yatsenko**[1,2,4,5,8] | **Urvashi Surti**[1,2,4,5,8]

[1]Pittsburgh Cytogenetics Laboratory, Magee-Womens Hospital of UPMC, Pittsburgh, PA, USA

[2]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

[3]Department of Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

[4]Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

[5]Department of Obstetrics, Gynecology and Reproductive Sciences, University of Pittsburgh, Pittsburgh, PA, USA

[6]Genomic Medicine Institute, Cleveland Clinic, Cleveland, OH, USA

[7]Center for Pediatric Neurology, Cleveland Clinic, Cleveland, OH, USA

[8]Magee Womens Research Institute, University of Pittsburgh Medical Center, Pittsburgh, PA, USA

Correspondence
Svetlana A. Yatsenko, Pittsburgh Cytogenetics Laboratory, Magee-Womens Hospital, 300 Halket Street, Pittsburgh, PA 15213, USA.
Email: yatsenkosa@upmc.edu

Present address
Jun Liao, Department of Pathology and Cell Biology, Columbia University Irving Medical Center, New York, NY, USA
Keith A. Coffman, Tourette Syndrome Center of Excellence, Children's Mercy Hospital, Kansas City, MS, USA

Abstract

Background: Benign hereditary chorea (BHC) is an autosomal dominant disorder characterized by early-onset non-progressive involuntary movements. Although *NKX2-1* mutations or deletions are the cause of BHC, some BHC families do not have pathogenic alterations in the *NKX2-1* gene, indicating that mutations of non-coding regulatory elements of *NKX2-1* may also play a role.

Methods and Results: By using whole-genome microarray analysis, we identified a 117 Kb founder deletion in three apparently unrelated BHC families that were negative for *NKX2-1* sequence variants. Targeted next generation sequencing analysis confirmed the deletion and showed that it was part of a complex local genomic rearrangement. In addition, we also detected a 648 Kb *de novo* deletion in an isolated BHC case. Both deletions are located downstream from *NKX2-1* on chromosome 14q13.2-q13.3 and share a 33 Kb smallest region of overlap with six previously reported cases. This region has no gene but contains multiple evolutionarily highly conserved non-coding sequences.

Conclusion: We propose that the deletion of potential regulatory elements necessary for *NKX2-1* expression in this critical region is responsible for BHC phenotype in these patients, and this is a novel disease-causing mechanism for BHC.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2021 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals LLC.

KEYWORDS
benign hereditary chorea, chromosome 14q13.2-q13.3, copy number variations, *NKX2-1*, non-coding regulatory elements

1 | INTRODUCTION

Benign hereditary chorea (BHC, MIM# 118700) is a rare autosomal dominant movement disorder with early onset in infancy or childhood. It is characterized by slightly progressive or non-progressive chorea, without cognitive decline or dementia. The choreic movements can involve any parts of the body and generally worsen with stressful situations and cease during sleep. Other associated but variable neurological features include developmental delay in motor skills, subnormal intelligence, memory deficits, learning difficulties, dysarthria, dystonia, hypotonia, ataxia, and tremor (Inzelberg et al., 2011; Kleiner-Fisman & Lang, 2007). Some BHC patients also have congenital hypothyroidism and respiratory diseases, including neonatal respiratory distress, recurrent pulmonary infections, chronic interstitial lung disease, and lung carcinoma, which together with the chorea phenotype, are commonly classified as choreoathetosis and congenital hypothyroidism with or without pulmonary dysfunction or brain-lung-thyroid syndrome (MIM# 610978; Willemsen et al., 2005).

An association between BHC and *NKX2-1* (MIM# 600635) mutations was first reported in 2002 (Breedveld, van Dongen, et al., 2002; Krude et al., 2002). To date, more than 140 sequence variants and deletions involving the *NKX2-1* gene have been detected in BHC families worldwide. *NKX2-1* is located on human chromosome 14q13 and encodes a member of the homeodomain-containing NKX2 family of transcription factors (Guazzi et al., 1990). It is expressed in brain, lung, and thyroid tissues both in embryonic and adult stages (Kimura et al., 1996; Krude et al., 2002; Lazzaro et al., 1991; Magno et al., 2009; Trueba et al., 2005). Consistent with its expression pattern, *Nkx2-1* homozygous knockout mice have extensive defects in these three organs and die at birth (Kimura et al., 1996; Sussel et al., 1999), while heterozygous mice show poor motor coordination and hyperthyrotropinemia (Pohlenz et al., 2002). Although all this evidence supports the pathogenic role of *NKX2-1* in BHC, there are several clinically diagnosed BHC families that fail to show detectable *NKX2-1* mutations (Bauer et al., 2006; Breedveld, Percy, et al., 2002; Shimohata et al., 2007), indicating the possibility of functional alterations in non-coding *NKX2-1* sequences or regulatory elements within the vicinity of *NKX2-1*, or that they may be due to defects in other gene(s). This hypothesis was further supported by a few reported deletions located in a close proximity to *NKX2-1* in patients/families with BHC or brain-lung-thyroid syndrome (Barnett et al., 2012; Dale et al., 2012; Invernizzi et al., 2018; Kharbanda et al., 2017; Thorwarth et al., 2014).

Non-coding regulatory elements, such as enhancers and silencers, are required for precise control of the spatiotemporal expression of many genes, especially those encoding transcription factors that regulate embryonic development. Although these elements are located at various distances from their target genes, sometimes even in introns of unrelated neighboring genes, they are characterized by strong evolutionary conservation between distant species due to their important biological function. Consequently, disruption of these elements by sequence variants, copy number variations (CNVs), or genomic balanced structural rearrangement may lead to human congenital defects, such as aniridia, sex reversal, holoprosencephaly, or digital malformations (Kleinjan & van Heyningen, 2005; Klopocki & Mundlos, 2011). However, because current genetic testing strategy generally focuses on the protein-coding regions (exons), which comprise less than 2% of the human genome, the contribution of disrupted non-coding regulatory elements to human disease is likely underestimated.

Here, we report a series of BHC patients with 14q13 deletions proximal to *NKX2-1*. Further genetic analysis suggests the presence of potential non-coding regulatory elements in the overlapping region shared by these deletions.

2 | CLINICAL REPORT

2.1 | Family 1

The family includes three affected individuals in two generations (Figure 1). Consanguinity is denied and the family history is otherwise unremarkable. The family's ethnic background is English.

The proband (II-1) is a 14-year-old boy who was born following an uncomplicated term pregnancy and vaginal delivery. With a birth weight of 2.15 kg, he was appeared healthy and did not require resuscitation. However, all of his developmental milestones were delayed. He was noted to have choreiform movements at 8–9 years of age, and these became progressively worse over time. His last neurological examination, at 13 years of age, revealed moderate choreoathetosis that was distal predominant in the upper and lower extremities. He had significant difficulties with eating and handwriting and his walking was complicated by intermittent falls. He also developed spasticity of the lower extremities

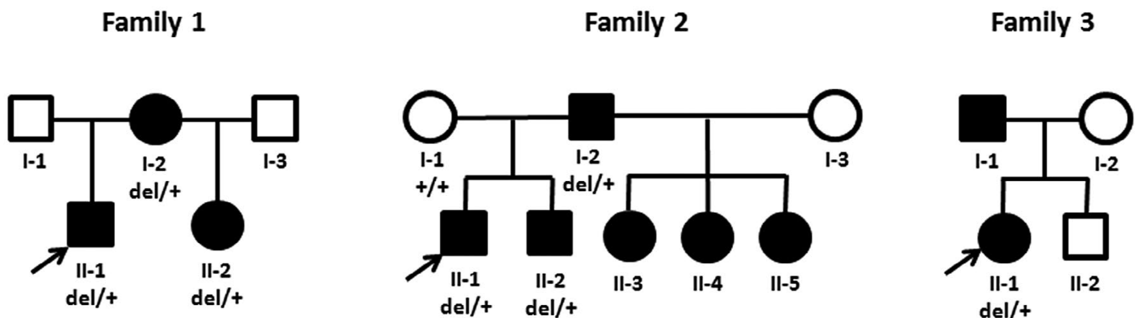


FIGURE 1 Pedigrees of three BHC families reported in this paper. Affected individuals are represented by closed symbols and unaffected individuals by open symbols. Probands are indicated by arrows. Genotypes of all tested individuals are listed below their symbols. +/+, normal microarray result; del/+, a heterozygous 14q13.3 deletion was detected

and required bilateral Achilles tendon transfer in his ankles. His family reported that he was hyperactive, impulsive and easily distractible. He had a poor academic performance at school and received speech and language services throughout his education for a significant stutter and various articulation problems. A comprehensive neuropsychological evaluation at 13 years of age indicated that his overall intellectual abilities were within the average range. However, he demonstrated specific and pervasive deficits of attention, concentration, processing speed, and memory, and was diagnosed with attention deficit hyperactivity disorder (ADHD), predominantly inattentive type. He has no history of pulmonary diseases or thyroid abnormalities. His brain MRI was unremarkable. Laboratory studies, including serum amino acids, urine organic acids, complete blood count (CBC), peripheral smear, serum lactate, and CSF neurotransmitters, were all normal. He also had a normal *NKX2-1* gene sequence.

The second patient (II-2) in this family is the maternal half-sister of the proband. She is an 8-year-old girl born at term after an uncomplicated pregnancy by cesarean section because of failure to progress. Although her birth weight was only 2.47 kg, she was well at birth and did not require intubation or resuscitation. Her development was normal except for falling problems during walking and running, and she was subsequently diagnosed with BHC at 5 years of age. Her choreoathetosis was much more significant in the lower than in the upper extremities and progressed to intermittent difficulties with ambulation. She also exhibited spasticity and nystagmus. After treatment with clonazepam at a dose of 0.125 mg twice a day, her symptoms improved remarkably. She walked better and her adventitious movements were under much better control. Behaviorally, she was diagnosed with disruptive behavior disorder, not otherwise specified, at 6 years of age because of her highly aggressive and mean behavior. She has no pulmonary or thyroid symptoms. Her brain MRI was normal. CAG repeat analysis of *ATXN3* and *NKX2-1* sequences were both normal.

The mother of both patients (I-2) had also choreiform movements as long as she could remember and had been diagnosed with "cerebral palsy" in her childhood. Her choreoathetosis was milder than that of either of her children. She reported that she was unable to sit still and always had some difficulties with walking.

2.2 | Family 2

This family includes six affected individuals in two generations (Figure 1). Consanguinity is denied and the family history is otherwise unremarkable. The family's ethnic background is German.

The proband (II-1) is a 16-year-old boy, delivered at 38 weeks' gestation by planned cesarean section secondary to placenta previa, with a birth weight of 3.69 kg. His developmental milestones were delayed, with sitting independently at 12 months and walking independently at 24 months. His choreic movements, which occurred in his trunk and all four extremities, were apparent at birth but remained stable over years. He also exhibited dystonia, ataxia, dysarthria, and tremor. His choreoathetosis was exacerbated when he was stressed, tired, or sick, and disappeared when he was asleep. His IQ was within the average range. He was also diagnosed with ADHD, intermittent explosive disorder, conduct disorder, and depression, and currently lives in a residential treatment facility because of his behavioral problems. He had a history of inflammatory bowel disease and gastroesophageal reflux, but lungs and thyroid were normal. Extensive work-up, including newborn screen, brain and spine MRI, immunoglobulin studies, thyroid function studies, urine amino acid, urine organic acids, comprehensive metabolic profile, calcium panel, uric acid, GGT, prolactin, CSF glucose, CSF neurotransmitters, lupus testing (ANA, C3/C4, anti-double stranded DNA), DNA sequencing of *NKX2-1*, *DYT1*, *HGPRT*, *GCH1*, *POLG1*, Aprataxin, Friedreich ataxia, and spinocerebellar atrophy (*SCA1*, *2*, *3*, *6*, *7*, *8*, *10*, *11*, *13*, *14*, *17*, *22*, *DRPLA*, and *FRDA*) were all normal.

The second patient (II-2) in this family is the full brother of the proband. He is a 12-year-old boy delivered at 38 weeks'

gestation by planned cesarean section, with a birth weight of 3.97 kg. He had global developmental delay, most severely in motor skills. He sat at 10 months, crawled at 15 months, and walked at 24 months. Similarly to the proband, he also had choreic movements with dystonia and ataxia, which remained unchanged for many years. However, his choreoathetosis was milder than that of the proband and did not cause any difficulties with gross motor movements. He also had a history of asthma, encopresis, ankyloglossia, otitis media, and pneumonia. Brain and spine MRI was normal.

The father (I-2) of both patients had the same movement disorder and a history of learning difficulties, explosive disorder, and depression. He also had three daughters (II-3, II-4, II-5) with another partner. All of them were reported to have choreoathetosis but clinical details were not available.

2.3 | Family 3

The proband (II-1) is a 13-year-old girl born after a normal term pregnancy, with a birth weight of 3.30 kg. She had a history of developmental delay, mainly in gross motor skills. She sat and crawled at one year of age, started to walk at age three, and learned to walk independently at age four. She had mild-to-moderate choreic movements in all four extremities and hypotonia in her lower extremities. She had normal cognitive abilities and performed well academically at school. She did not have any other neurological or behavioral problems and her brain MRI was unremarkable. Tests for CBC, CMP, CPK, very long chain fatty acids, acylcarnitine profile, and fragile X were all normal. The *NKX2-1* gene sequence was normal. Her father (I-1) was also reported to have choreoathetosis but clinical details were not available. The family history is otherwise unremarkable (Figure 1).

2.4 | The isolated case

This patient is an 8-year-old girl, born at 38 weeks by unassisted vaginal delivery, with a birth weight of 3.41 kg. She had a history of motor developmental delay. She did not sit independently until 15 months, crawled at 19 months, stood at 33 months, and walked at 36 months. She has constant involuntary choreiform movements of her arms and legs that only stop when she is asleep and become worse when she is tired or sick. She also has ataxia and an unsteady and uncoordinated gait with poor balance, causing persistent difficulties with standing, walking, and climbing stairs. She is hyperkinetic, with normal intelligence, and is an A student at school. There is no history of pulmonary disease or thyroid abnormality. TSH and free T4 were within the normal ranges. She had no dysmorphic features, neurocutaneous stigmata, hepatosplenomegaly, orthopedic deformities, or scoliosis. Brain MRI, EEG, and EMG were all normal. Metabolic studies and genetic tests, including tests for fragile X, spinal muscular atrophy, mitochondrial myopathy panel, *DYT1*, and a complete ataxia panel (*SCA1*, *SCA2*, *SCA3*, *SCA5*, *SCA6*, *SCA7*, *SCA8*, *SCA10*, *SCA14*, *SCA17*, *DRPLA*, *FRDA*, *POLG1*, *SETX*, *SIL1*, *TTPA*, and *KCNC3*) were all normal.

There is no history of neurological or neuromuscular disease in the family except for a maternal second cousin with Duchenne muscular dystrophy. Any known consanguinity was denied. The family ethnic background is mixed Caucasian.

3 | METHODS

3.1 | Ethical compliance

This study was approved by the University of Pittsburgh IRB (STUDY21010129).

3.2 | Array comparative genomic hybridization (array) analysis

Oligonucleotide-based whole-genome array comparative genomic hybridization (CGH) was performed at Pittsburgh Cytogenetics Laboratory (except the isolated case which was processed at Signature Genomic Laboratories) as previously described (Liao et al., 2011). NimbleGen 135 K oligonucleotide array, SignatureChip Oligo Solution, version 2.0, which was custom-designed by Signature Genomic Laboratories (Spokane) and made by Roche NimbleGen (Madison), was used. Results were displayed by custom oligonucleotide array CGH analysis software Genoglyphix, version 2.6 (Signature Genomic Laboratories).

3.3 | Fluorescence *in situ* hybridization analysis

Fluorescence *in situ* hybridization (FISH) analysis was performed on metaphase spreads of cultured peripheral blood lymphocytes from patients and their parents, using standard procedures. FISH probes RP11-363G18 and RP11-74F2 were made from the RPCI-11 bacterial artificial chromosome (BAC) library (Invitrogen), and labeled using a Nick Translation Kit (Abbott Molecular). Images were captured using Isis FISH Imaging System v5.3 software (MetaSystems).

3.4 | Haplotype analysis

Four microsatellite loci were chosen for haplotyping. All the microsatellites were di-nucleotide repeat loci. The genomic

location, details of the repeats, and PCR primers are described in Table S1. PCR products were fluorescently labeled using a nested PCR approach and a FAM labeled M13 primer (Sigma-Aldrich, St. Louis, Missouri, USA) as described previously (Schuelke, 2000). PCR products were diluted 1:5 and then run on an ABI 3130xl automated DNA sequencer using a LIZ 500 size standard (Applied Biosystems, Carlsbad, California, USA) as per the manufacturer's instructions. Fragment sizes were estimated using PeakScanner software (Applied Biosystems, Carlsbad, California, USA). The allele frequencies at these loci were determined from control individuals from the same geographic area as the patients. Heterozygosity indices were calculated using the formula $1-\Sigma pi^2$ where pi represents the frequency of the i'th allele. Fisher's exact tests were used to determine significance of association between alleles in patients and controls.

3.5 | Targeted next generation sequencing

DNA library generation, amplification, target enrichment, and sequencing were carried out at the Epigenomics Shared Facility of Albert Einstein College of Medicine, using a SeqCap EZ Choice Library (NimbleGen) to capture a 3.55 Mb interval, chr14:35,178,000–38,727,000 (hg19). An indexed library was generated using Illumina TruSeq reagents, captured, and subjected to paired-end sequencing on an Illumina HiSeq 2000 sequencing system. Analyzed sequences had an average depth of 192. DNA copy number analysis was carried out with the GenPlay genotype analyzer (Lajugie & Bouhassira, 2011), SNP/Indel analysis used the Genome Analysis Toolkit (GATK; McKenna et al., 2010), and chromosome breakpoints were detected with Pindel (Ye et al., 2009) and the Integrative Genomics Viewer (IGV; Thorvaldsdottir et al., 2013).

4 | RESULTS

4.1 | Molecular analysis of patients

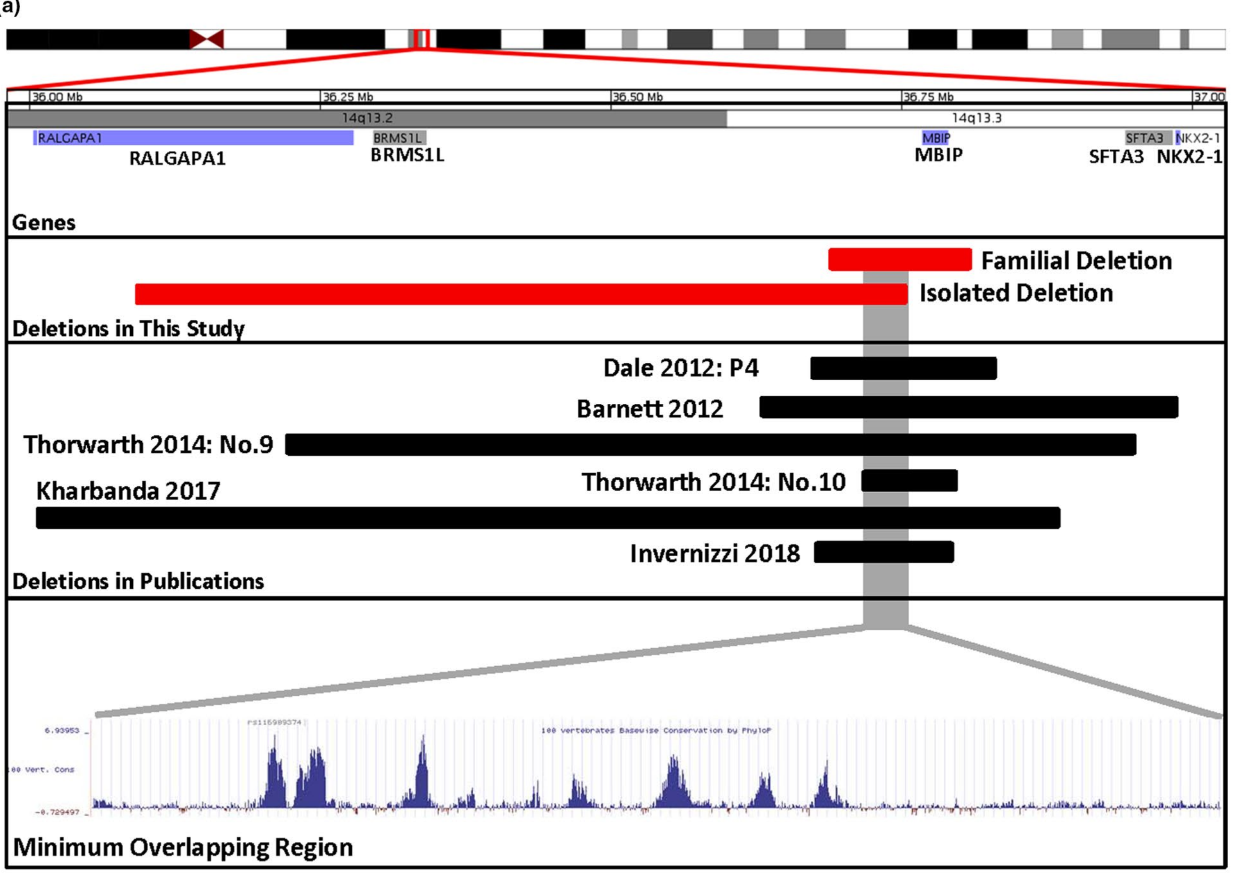
DNA specimens from eight individuals from the three families (I-2, II-1, II-2 in Family 1; I-1, I-2, II-1, II-2 in Family 2; II-1 in Family 3) were available for whole-genome array CGH analysis. An interstitial single-copy loss of chromosome 14q13.3 region (chr14:36693910-36810554, hg19) was detected in all seven tested BHC patients but not in the unaffected member I-1 in Family 2 (Figure 1). The deletion is ~117 Kb in size and contains a single gene, *MBIP* (Figure 2a). Metaphase FISH analysis was performed in probands of Families 1 and 2 by using BAC probes RP11-363G18 and RP11-74F2, respectively, which overlapped but were larger than the deleted region. FISH results showed a diminished signal in one of the chromosome 14 homologues, confirming the presence of a heterozygous deletion (Figure 2b).

The deletions in these three families had exactly the same sizes and breakpoints. In addition, although to the best of our knowledge, none of the families are related, they all reside within a short distance of each other in a very rural area of north central Pennsylvania, raising the possibility that they may share a common ancestor. In order to test the possible founder effect, affected individuals and normal controls from the same geographic region were genotyped using microsatellite markers flanking the deleted region. Of the 4 loci studied, only 2 (US1 and US3), were found to be polymorphic. US1 was about 136 Kb centromeric to the putative deletion junction, while US3 was about 41 Kb telomeric to the deletion junction. Chromosomes carrying the deletion from patients in all three affected families shared the same alleles at these loci (Table 1). The 229–195 (US1-US3) haplotype was found to be significantly associated with patient chromosomes and was not found in the normal chromosomes analyzed ($p = 6 \times 10^{-4}$), indicating that these families do share a common founder haplotype in the deleted region.

In the isolated case, a longer heterozygous deletion (chr14:36116503-36764947, hg19) on chromosome 14q13.2-q13.3 region was identified by array CGH analysis. The deletion was ~648 Kb in size and contained at least two known genes, *RALGAPA1* and *BRMS1L* (Figure 2a). Deletion was confirmed in 28 of 30 cells examined by FISH analysis using BAC probe RP11-679C15. In two cells, two normal FISH signals were observed, indicating a possible mosaicism or a false-negative for a deletion hybridization pattern. It was also determined to be a *de novo* event, based on normal FISH results in both parents, which is consistent with the negative family history of movement disorders. No other CNVs with potential clinical relevance were found in any of the above individuals tested by whole-genome array CGH analysis.

4.2 | Comparative genomic analysis of the smallest region of overlap

Three probands (II-1 s in three families) as well as the patient II-2 in Family 1 have been tested by sequence analysis for *NKX2-1*. All patients showed normal results. Interestingly, although detected deletions do not involve *NKX2-1*, their distal breakpoints are located just 175 and 221 Kb downstream from *NKX2-1* (Figure 2a). By comparing these two deletions with six previously reported deletions also outside *NKX2-1*, we defined a 33 Kb non-coding segment (chr14: 36732190–36764947, hg19) as a smallest region of overlap (Figure 2a). It raises the possibility that potential cis-acting regulatory elements of *NKX2-1* may be present in this critical interval. Because such regulatory elements are often evolutionarily conserved, we performed 100 Vertebrates Basewise



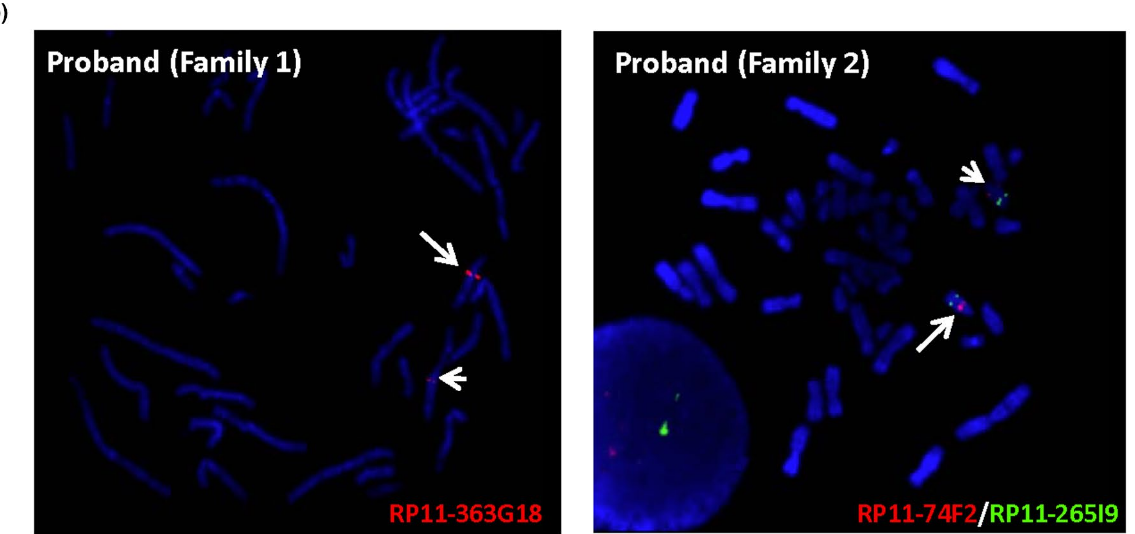


FIGURE 2 (a) Schematic illustration of chromosome 14q13.2-q13.3 region showing deletions in three BHC families and an isolated case in this study, six deletions reported in previous studies, and PhyloP scores from 100 Vertebrates Basewise Conservation Analysis in the minimum overlapping region. (b) Metaphase FISH results for probands of Families 1 and 2 by using BAC probes RP11-363G18 (red) and RP11-74F2 (red) respectively. BAC probe RP11-265I9 (green) is used as a control. Normal signals are indicated by arrows and diminished ones are indicated by arrowheads

Conservation Analysis by PhyloP for this region in the USCS Genome Browser (https://genome.ucsc.edu/). PhyloP scores clearly showed multiple evolutionarily conserved non-coding sequences in this region (Figure 2a), strongly suggesting the functional importance of the region and its potential role in gene regulation. Multiple enhancers are mapped inside of the deleted segments, including enhancers GH14 J036269 and GH14 J036272, located in the 33 Kb smallest region of

TABLE 1 Haplotype analysis of BHC families with 14q13.3 deletion

Locus	Heterozygosity index (1-Σpi^2)	Allele	Patient (frequency)	Controls (frequency)	Fisher's test *p* value
US1	0.72	229	3/3 (1)	1/20 (0.05)	.07
US3	0.75	195	3/3 (1)	7/20 (0.35)	.002
229–195 haplotype	—	—	3/3 (1)	0/20 (0)	.0006

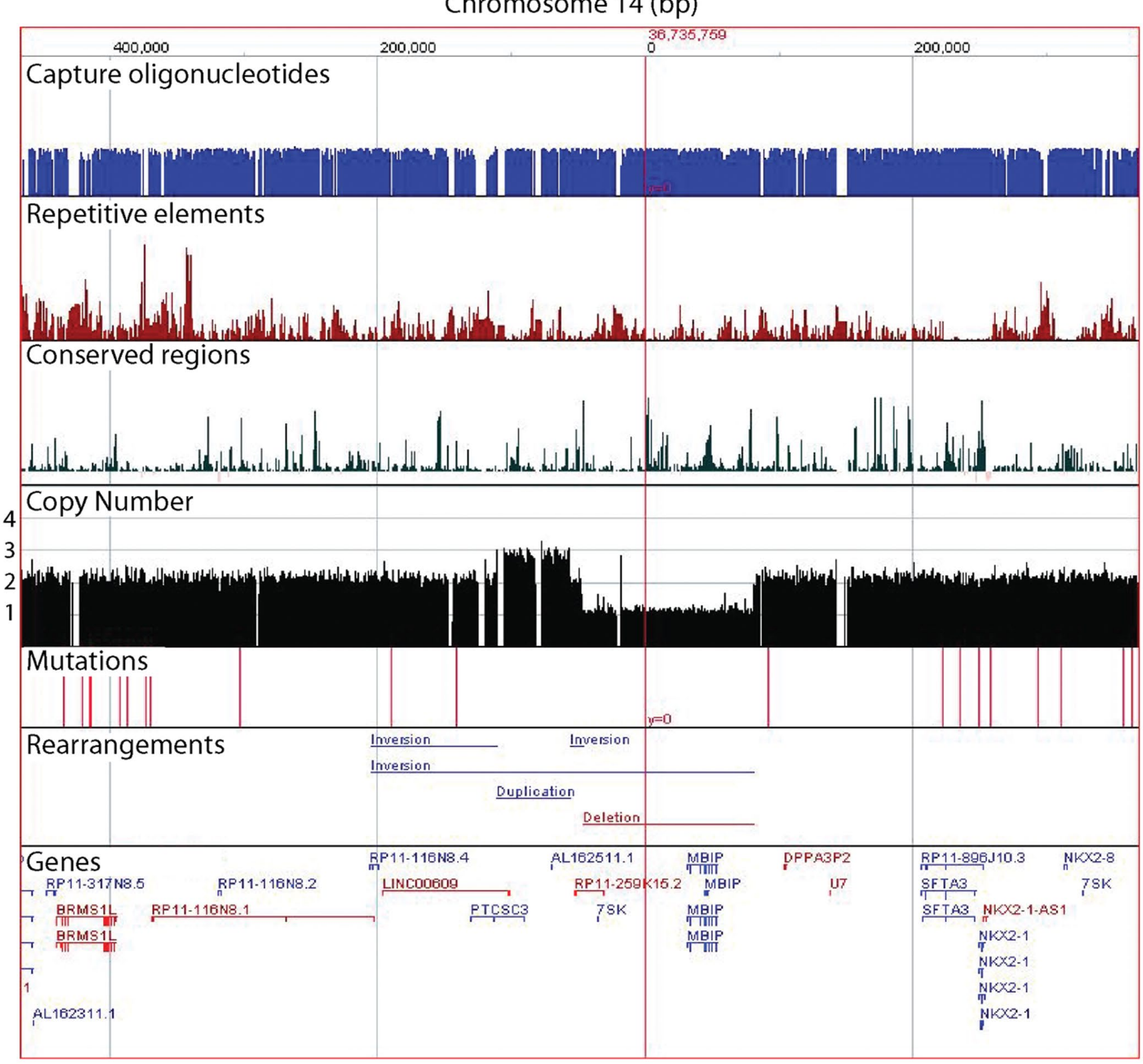


FIGURE 3 Next generation sequence analysis of the genomic region encompassing 14q13 deletions

overlap (Fishilevich et al., 2017). In order to provide additional evidence for this hypothesis, the expression level of *NKX2-1* in the leukocytes of patients and several healthy control individuals was analyzed by quantitative RT-PCR method. Unfortunately, *NKX2-1* expression in blood samples was too low to be detected.

4.3 | Next generation sequencing of the deleted region

In order to further characterize the genomic region encompassing the deletions, a 3.55 Mb interval of chromosome 14 (chr14:35178000–38727000, hg19) was captured from

TABLE 2 Summary of BHC patients with 14q13 deletions proximal to NKX2-1 in this and previous studies

Patient	Family 1 II−1	Family 1 II−2	Family 2 II−1	Family 2 II−2	Family 3 II−1	The Isolated Case	Dale et al. (2012)	Barnett et al. (2012)
Gender	M	M	M	M	F	F	NA	F
Age	14 y	8 y	16 y	12 y	13 y	8 y	1.2 y	2 y
Genetics								
Deletion size	117 Kb					648 Kb	155 Kb	3.54 Mb
Coordinates (hg19)	chr14: 36693910–36810554					chr14: 36116503–36764947	chr14: 36660897–36815991	chr14: 36612756–36966361
Involved gene	*MBIP*					*RALGAPA1*, *BRMS1L*	*MBIP*	16 genes including *RALGAPA1*, *BRMS1L*, and *MBIP*
Inheritance	Maternal		Paternal		Paternal	De novo	Maternal	NA
Clinical features								
Chorea	+	+	+	+	+	+	+	+
Motor DD	+	+	+	+	+	+	+	+
Hypothyroidism	−	−	−	−	−	−	NA	+
Lung diseases	−	−	−	−	−	−	NA	+
Other related findings	Spasticity, ADHD	Spasticity, Nystagmus, DBD	ADHD, IED, CD, Depression	Dystonia, Ataxia	Hypotonia	Hyperkinesis, Ataxia, Poor Balance	NA	NA

Abbreviations: +, present; −, absent; ADHD, attention deficit hyperactivity disorder; CD, conduct disorder; DBD, disruptive behavior disorder; DD, developmental delay; F, female; IED, intermittent explosive disorder; M, male; NA, not available.

the DNA sample of patient II-1 in Family 2, using a custom oligonucleotide library. After paired-end sequencing, the reads were normalized to control DNA for CNV analysis, which demonstrated a 0.13 Mb region of allelic deletion (chr14:36680180-36817162, hg19), confirming the result by array CGH analysis (Figure 3). However, there was also a nearby 0.06 Mb region of allelic duplication, suggesting a gene rearrangement more complex than simple deletion. Direct sequence analysis confirmed this complexity by showing several rearrangements in the region including deletions, duplications, and inversions (Figure 3 and Table S2), suggesting that this complex intrachromosomal rearrangement has resulted from multiple fork slippage and template switching events. Further computational analysis for detecting point mutations, indels, and known SNPs, identified several novel mutations (Table S3), but none corresponded to coding regions, splicing elements, exon sequences of lncRNAs, or highly conserved nontranscribed regions. There were no additional breakpoints or significant mutations within a 1 Mb interval centered on the deletion, which extended well past *NKX2-1*; this seems to rule out a mutation within an extended region that co-segregates with the deletion and might have included *NKX2-1*.

5 | DISCUSSION

In this study, we report a series of familial and isolated BHC cases with no detectable mutations in the coding region of the *NKX2-1* gene but deletions downstream from it. In three BHC families, a 117 Kb founder deletion was co-segregated with BHC phenotype among multiple family members. While a 648 Kb *de novo* overlapping deletion was found in a BHC patient without family history. In addition, six more deletions in the same chromosomal region with intact *NKX2-1* genes have been reported in patients/families with BHC or brain-lung-thyroid syndrome in the literature (Barnett et al., 2012; Dale et al., 2012; Invernizzi et al., 2018; Kharbanda et al., 2017; Thorwarth et al., 2014). By comparing microarray results from these patients, we defined a 33 Kb smallest region of overlap that contains no gene but does include multiple evolutionarily conserved non-coding sequences, suggesting their role in regulating *NKX2-1* expression. In contrast, we reviewed microarray results in more than 3,000 patients referred to our laboratory for other congenital anomalies and did not find any CNVs overlapping with these deletions. We also checked the Database of Genomic Variants (DGV, http://projects.tcag.ca/variation), a publicly available

Thorwarth et al. (2014)		Kharbanda et al. (2017)	Invernizzi et al. (2018)					
No. 9	No. 10		I−1	II−3	II−5	III−1	III−2	III−3
F	M	F	F	M	M	M	F	F
NA	NA	7 y	NA	NA	NA	NA	NA	NA
727 Kb	77 Kb	870 Kb	115 Kb					
chr14: 36187694–36914876	chr14: 36732190–36808778	chr14: 36000450–36870812	chr14: 36675857–36790795					
RALGAPA1, BRMS1L, MBIP	*MBIP*	*RALGAPA1, BRMS1L, MBIP*	*MBIP*					
De novo	Familial	De novo	Familia					
+	+	+	+	+	+	+	+	−
NA	NA	+	NA	NA	NA	+	+	+
+	−	+	−	−	−	−	−	−
−	−	−	−	+	+	+	+	−
NA	NA	Hypotonia, Myoclonus-Dystonia	NA	NA	Minor Cognitive and Psychiatric Involvement	Hypotonia, Myoclonic Jerks, Mild Cognitive Impairment	Hypotonia, Myoclonic Jerks	Unbalanced Gait

database of CNVs in the general population and found no benign deletions reported in this region. Altogether, collective evidence suggests that deletions downstream of *NKX2-1* are likely to be pathogenic and associated with BHC phenotype even without mutations in the coding region.

Clinical and genetic features of six patients with available clinical details in this study and six previously reported patients/families are summarized and compared in Table 2. Most patients had choreiform movements of variable degree and a history of developmental delay in motor skills. Other associated neurological symptoms found in our patients include spasticity, nystagmus, dystonia, hypotonia, ataxia, poor balance, hyperkinesis, and ADHD. Interestingly, at least three patients in our BHC families (II-2 in Family 1, I-2 and II-1 in Family 2) and one patient from a previous report (II-5 in Invernizzi et al., 2018) also had significant psychobehavioral problems, including disruptive behavior disorder, intermittent explosive disorder, conduct disorder, and depression. To the best of our knowledge, only two BHC patients with *NKX2-1* mutations have been reported to show psychiatric features, one with schizophrenia (Glik et al., 2008) and the other with postpartum psychosis (Salvatore et al., 2010).

The high incidence of neurobehavioral findings in our BHC families may be due to the *MBIP* gene deletion. As the only gene in the deleted region shared by these patients, *MBIP*, which encodes a protein kinase inhibitor with unknown biological function (Fukuyama et al., 2000), has been shown to be co-expressed with *NKX2-1* in the developing forebrain, particularly in the medial ganglionic eminence (Tucker et al., 2008). MBIP is also a component of histone acetyltransferase activity complex, which has the ability to modify chromatin. Deletions involving MBIP or its regulatory elements may affect the intra- and inter-chromosomal interactions between the co-expressed genes or result in a gene silencing secondary to chromatin modifications. Alternatively, other unknown genetic modifiers shared by these three families may also play a role.

In conclusion, we identified a 117 Kb founder deletion caused by a complex genomic rearrangement in three BHC families without detectable *NKX2-1* gene mutation, and a 648 Kb *de novo* mosaic deletion in an isolated case. Consistent with previously reported deletions in this region, both of them are located proximal to *NKX2-1* in chromosome 14q13.2-q13.3 and they share a 33 Kb smallest region of overlap containing no gene but multiple evolutionarily

highly conserved non-coding sequences. We propose that deletions of potential regulatory elements necessary for *NKX2-1* expression in this critical region are responsible for BHC phenotype in these patients. This is a novel disease-causing mechanism for BHC, and we suggest that more BHC patients with normal *NKX2-1* sequences should be screened for deletions or mutations in the surrounding regulatory regions. It would also be interesting to perform haplotype analysis on other BHC families with deletions in this region to determine whether they are also from the same founder event or indicate the presence of a mutation hotspot in the region. In addition, further studies are needed to fine-map these regulatory elements and determine their role in BHC, NKX2-1-related conditions, and intrachromosomal gene interactions, which will provide further insight into the genetic pathways regulating *NKX2-1* expression. Finally, in this era of personalized genomics, high-resolution genome-wide technologies such as genomic microarray and whole-genome sequencing are becoming routine diagnostic tools, which will inevitably result in the detection of many more pathogenic mutations in non-coding regulatory sequences. This study exemplifies the importance of adopting a more extensive "functionome" view, when we make genetic diagnoses, to include this new type of mutation, especially in genomic regions near known disease-causing or key developmental genes.

ACKNOWLEDGMENTS

We thank the staff at the Pittsburgh Cytogenetics Laboratory, Magee-Womens Hospital of UPMC for their technical support.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Jun Liao, Urvashi Surti, Svetlana A. Yatsenko, Jie Hu, Malini Sathanoori, and Aleksandar Rajkovic performed genetic analysis. Jun Liao, Urvashi Surti, Svetlana A. Yatsenko, and Lori Hoffner wrote the manuscript. Keith A. Coffman, Robyn A. Filipink, Suneeta Madan-Khetarpal, Marianne McGuire, Allison Schreiber, Rocio Moran, and Neil Friedman collected the clinical data and conducted the clinical evaluations. Joseph Locker performed the NGS experiment. Quasar S. Padiath and Bruce Nmezi performed the haplotype analysis. All authors approved the final manuscript.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Jun Liao https://orcid.org/0000-0002-9806-2722
Quasar S. Padiath https://orcid.org/0000-0002-2468-6364

REFERENCES

Barnett, C. P., Mencel, J. J., Gecz, J., Waters, W., Kirwin, S. M., Vinette, K. M. B., Uppill, M., & Nicholl, J. (2012). Choreoathetosis, congenital hypothyroidism and neonatal respiratory distress syndrome with intact NKX2-1. *American Journal of Medical Genetics Part A*, *158A*(12), 3168–3173. https://doi.org/10.1002/ajmg.a.35456

Bauer, P., Kreuz, F. R., Burk, K., Saft, C., Andrich, J., Heilemann, H., Riess, O., & Schols, L. (2006). Mutations in TITF1 are not relevant to sporadic and familial chorea of unknown cause. *Movement Disorders*, *21*(10), 1734–1737. https://doi.org/10.1002/mds.21031

Breedveld, G. J., Percy, A. K., MacDonald, M. E., de Vries, B. B., Yapijakis, C., Dure, L. S., Ippel, E. F., Sandkuijl, L.A., Heutink, P., & Arts, W. F. (2002). Clinical and genetic heterogeneity in benign hereditary chorea. *Neurology*, *59*(4), 579–584. https://doi.org/10.1212/WNL.59.4.579

Breedveld, G. J., van Dongen, J. W., Danesino, C., Guala, A., Percy, A. K., Dure, L. S., Harper, P., Arts, W. F., Lazarou, L. P., Van Der Linde, H., de Vries, B., Joosse, M., & Heutink, P. (2002). Mutations in TITF-1 are associated with benign hereditary chorea. *Human Molecular Genetics*, *11*(8), 971–979. https://doi.org/10.1093/hmg/11.8.971

Dale, R. C., Grattan-Smith, P., Nicholson, M., & Peters, G. B. (2012). Microdeletions detected using chromosome microarray in children with suspected genetic movement disorders: a single-centre study. *Developmental Medicine & Child Neurology*, *54*(7), 618–623. https://doi.org/10.1111/j.1469-8749.2012.04287.x

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., Lancet, D., & Cohen, D. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. *Database (Oxford)*, *2017*, 1–17. https://doi.org/10.1093/database/bax028

Fukuyama, K., Yoshida, M., Yamashita, A., Deyama, T., Baba, M., Suzuki, A., Mohri, H., Ikezawa, Z., Nakajima, H., Hirai, S.-I., & Ohno, S. (2000). MAPK upstream kinase (MUK)-binding inhibitory protein, a negative regulator of MUK/dual leucine zipper-bearing kinase/leucine zipper protein kinase. *Journal of Biological Chemistry*, *275*(28), 21247–21254. https://doi.org/10.1074/jbc.M001488200

Glik, A., Vuillaume, I., Devos, D., & Inzelberg, R. (2008). Psychosis, short stature in benign hereditary chorea: a novel thyroid transcription factor-1 mutation. *Movement Disorders*, *23*(12), 1744–1747. https://doi.org/10.1002/mds.22215

Guazzi, S., Price, M., De Felice, M., Damante, G., Mattei, M. G., & Di Lauro, R. (1990). Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity. *EMBO Journal*, *9*(11), 3631–3639. https://doi.org/10.1002/j.1460-2075.1990.tb07574.x

Invernizzi, F., Zorzi, G., Legati, A., Coppola, G., D'Adamo, P., Nardocci, N., Garavaglia, B., & Ghezzi, D. (2018). Benign hereditary chorea and deletions outside NKX2-1: What's the role of MBIP? *European Journal of Medical Genetics*, *61*(10), 581–584. https://doi.org/10.1016/j.ejmg.2018.03.011

Inzelberg, R., Weinberger, M., & Gak, E. (2011). Benign hereditary chorea: An update. *Parkinsonism & Related Disorders*, *17*(5), 301–307. https://doi.org/10.1016/j.parkreldis.2011.01.002

Kharbanda, M., Hermanns, P., Jones, J., Pohlenz, J., Horrocks, I., & Donaldson, M. (2017). A further case of brain-lung-thyroid syndrome with deletion proximal to NKX2-1. *European Journal of Medical Genetics*, *60*(5), 257–260. https://doi.org/10.1016/j.ejmg.2017.03.001

Kimura, S., Hara, Y., Pineau, T., Fernandez-Salguero, P., Fox, C. H., Ward, J. M., & Gonzalez, F. J. (1996). The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. *Genes & Development*, *10*(1), 60–69. https://doi.org/10.1101/gad.10.1.60

Kleiner-Fisman, G., & Lang, A. E. (2007). Benign hereditary chorea revisited: a journey to understanding. *Movement Disorders*, *22*(16), 2297–2305. https://doi.org/10.1002/mds.21644

Kleinjan, D. A., & van Heyningen, V. (2005). Long-range control of gene expression: emerging mechanisms and disruption in disease. *The American Journal of Human Genetics*, *76*(1), 8–32. https://doi.org/10.1086/426833

Klopocki, E., & Mundlos, S. (2011). Copy-number variations, noncoding sequences, and human phenotypes. *Annual Review of Genomics and Human Genetics*, *12*(1), 53–72. https://doi.org/10.1146/annurev-genom-082410-101404

Krude, H., Schutz, B., Biebermann, H., von Moers, A., Schnabel, D., Neitzel, H., Tönnies, H., Weise, D., Lafferty, A., Schwarz, S., DeFelice, M., von Deimling, A., van Landeghem, F., DiLauro, R., & Grüters, A. (2002). Choreoathetosis, hypothyroidism, and pulmonary alterations due to human NKX2-1 haploinsufficiency. *Journal of Clinical Investigation*, *109*(4), 475–480. https://doi.org/10.1172/JCI0214341

Lajugie, J., & Bouhassira, E. E. (2011). GenPlay, a multipurpose genome analyzer and browser. *Bioinformatics*, *27*(14), 1889–1893. https://doi.org/10.1093/bioinformatics/btr309

Lazzaro, D., Price, M., de Felice, M., & Di Lauro, R. (1991). The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain. *Development*, *113*(4), 1093–1104.

Liao, J., DeWard, S. J., Madan-Khetarpal, S., Surti, U., & Hu, J. (2011). A small homozygous microdeletion of 15q13.3 including the CHRNA7 gene in a girl with a spectrum of severe neurodevelopmental features. *American Journal of Medical Genetics Part A*, *155A*(11), 2795–2800.

Magno, L., Catanzariti, V., Nitsch, R., Krude, H., & Naumann, T. (2009). Ongoing expression of Nkx2.1 in the postnatal mouse forebrain: Potential for understanding NKX2.1 haploinsufficiency in humans? *Brain Research*, *1304*, 164–186.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Pohlenz, J., Dumitrescu, A., Zundel, D., Martine, U., Schonberger, W., Koo, E., Weiss, R. E., Cohen, R. N., Kimura, S., & Refetoff, S. (2002). Partial deficiency of thyroid transcription factor 1 produces predominantly neurological defects in humans and mice. *Journal of Clinical Investigation*, *109*(4), 469–473. https://doi.org/10.1172/JCI0214192

Salvatore, E., Di Maio, L., Filla, A., Ferrara, A. M., Rinaldi, C., Sacca, F., Peluso, S., Macchia, P. E., Pappatà, S., & De Michele, G. (2010). Benign hereditary chorea: clinical and neuroimaging features in an Italian family. *Movement Disorders*, *25*(10), 1491–1496. https://doi.org/10.1002/mds.23065

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, *18*(2), 233–234. https://doi.org/10.1038/72708

Shimohata, T., Hara, K., Sanpei, K., Nunomura, J., Maeda, T., Kawachi, I., Kanazawa, M., Kasuga, K., Miyashita, A., Kuwano, R., & Hirota, K. (2007). Novel locus for benign hereditary chorea with adult onset maps to chromosome 8q21.3 q23.3. *Brain*, *130*(Pt 9), 2302–2309.

Sussel, L., Marin, O., Kimura, S., & Rubenstein, J. L. (1999). Loss of Nkx2.1 homeobox gene function results in a ventral to dorsal molecular respecification within the basal telencephalon: evidence for a transformation of the pallidum into the striatum. *Development*, *126*(15), 3359–3370.

Thorvaldsdottir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192. https://doi.org/10.1093/bib/bbs017

Thorwarth, A., Schnittert-Hübener, S., Schrumpf, P., Müller, I., Jyrch, S., Dame, C., Biebermann, H., Kleinau, G., Katchanov, J., Schuelke, M., Ebert, G., Steininger, A., Bönnemann, C., Brockmann, K., Christen, H.-J., Crock, P., deZegher, F., Griese, M., Hewitt, J., … Krude, H. (2014). Comprehensive genotyping and clinical characterisation reveal 27 novel NKX2-1 mutations and expand the phenotypic spectrum. *Journal of Medical Genetics*, *51*(6), 375–387. https://doi.org/10.1136/jmedgenet-2013-102248

Trueba, S. S., Auge, J., Mattei, G., Etchevers, H., Martinovic, J., Czernichow, P., Vekemans, M., Polak, M., & Attie-Bitach, T. (2005). PAX8, TITF1, and FOXE1 gene expression patterns during human development: new insights into human thyroid development and thyroid dysgenesis-associated malformations. *The Journal of Clinical Endocrinology & Metabolism*, *90*(1), 455–462.

Tucker, E. S., Segall, S., Gopalakrishna, D., Wu, Y., Vernon, M., Polleux, F., & Lamantia, A. S. (2008). Molecular specification and patterning of progenitor cells in the lateral and medial ganglionic eminences. *Journal of Neuroscience*, *28*(38), 9504–9518. https://doi.org/10.1523/JNEUROSCI.2341-08.2008

Willemsen, M. A., Breedveld, G. J., Wouda, S., Otten, B. J., Yntema, J. L., Lammens, M., & de Vries, B. B. (2005). Brain-Thyroid-Lung syndrome: a patient with a severe multi-system disorder due to a de novo mutation in the thyroid transcription factor 1 gene. *European Journal of Pediatrics*, *164*(1), 28–30. https://doi.org/10.1007/s00431-004-1559-x

Ye, K., Schulz, M. H., Long, Q., Apweiler, R., & Ning, Z. (2009). Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*, *25*(21), 2865–2871. https://doi.org/10.1093/bioinformatics/btp394

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Liao J, Coffman KA, Locker J, et al. Deletion of conserved non-coding sequences downstream from *NKX2-1*: A novel disease-causing mechanism for benign hereditary chorea. *Mol Genet Genomic Med*. 2021;9:e1647. https://doi.org/10.1002/mgg3.1647