BRIEF REPORT

Human Mutation

OFFICIAL JOURNAL HGVS HUMAN GENOME VARIATION SOCIETY www.hgvs.org

GeneMatcher Aids in the Identification of a New Malformation Syndrome with Intellectual Disability, Unique Facial Dysmorphisms, and Skeletal and Connective Tissue Abnormalities Caused by De Novo Variants in *HNRNPK*

P. Y. Billie Au,[1†] Jing You,[2,3†] Oana Caluseriu,[4] Jeremy Schwartzentruber,[5] Jacek Majewski,[5] Francois P. Bernier,[1,6] Marcia Ferguson, Care for Rare Canada Consortium,[7] David Valle,[3,8,9,10] Jillian S. Parboosingh,[1,6] Nara Sobreira,[3] A. Micheil Innes,[1,6*‡] and Antonie D. Kline[7‡§]

[1]*Department of Medical Genetics, University of Calgary, Cumming School of Medicine, Alberta, Canada;* [2]*Predoctoral Training Program in Human Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205;* [3]*McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205;* [4]*Department of Medical Genetics, University of Alberta, Edmonton, Alberta, Canada;* [5]*Department of Human Genetics, McGill and Genome Quebec Innovation Center, McGill University, Quebec, Canada;* [6]*Alberta Children's Hospital, Research Institute for Child and Maternal Health, University of Calgary, Alberta, Canada;* [7]*Harvey Institute for Human Genetics, Department of Pediatrics, Greater Baltimore Medical Center, Baltimore, Maryland;* [8]*Center for Inherited Disease Research, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21224;* [9]*Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205;* [10]*McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205*

For the Matchmaker Exchange Special Issue

Received 7 April 2015; accepted revised manuscript 29 June 2015.
Published online 14 July 2015 in Wiley Online Library (www.wiley.com/humanmutation). DOI: 10.1002/humu.22837

ABSTRACT: We report a new syndrome due to loss-of-function variants in the heterogeneous nuclear ribonucleoprotein K gene (*HNRNPK*). We describe two probands: one with a de novo frameshift (NM_002140.3: c.953+1dup), and the other with a de novo splice donor site variant (NM_002140.3: c.257G>A). Both probands have intellectual disability, a shared unique craniofacial phenotype, and connective tissue and skeletal abnormalities. The identification of this syndrome was made possible by a new online tool, GeneMatcher, which facilitates connections between clinicians and researchers based on shared interest in candidate genes. This report demonstrates that new Web-based approaches can be effective in helping investigators solve exome sequencing projects, and also highlights the newer paradigm of "reverse phenotyping," where characterization of syndromic features follows the identification of genetic variants.

Hum Mutat 36:1009–1014, 2015. © 2015 Wiley Periodicals, Inc.

KEY WORDS: HNRNPK; reverse phenotyping; GeneMatcher; matchmaker exchange; WES

Whole-exome sequencing (WES) has rapidly become the primary method for the discovery of causative genes in rare disease. However, the ability to correctly identify responsible causative variants can be challenging, particularly with ultrarare disorders that frequently arise de novo. While in silico, in vitro, cell-based, and animal model systems are able to provide some validation in linking a gene variant to a phenotype, finding another proband with a similar phenotype and the same causative gene is invaluable. The need for improved connections between clinicians and researchers who have interest in the same genes and/or phenotypes has become increasingly apparent as more variants are identified with next-generation technologies, and has in part been addressed by several novel tools now available on the internet. Recently, Sobreira et al. (2015) have developed a Web-based tool called GeneMatcher (http://www.genematcher.org). Unlike other online tools such as PhenomeCentral, which match primarily based on phenotype, GeneMatcher facilitates connections between clinicians and researchers based on shared interest in a candidate gene. Herein, we illustrate how a Web-based match tool has played an instrumental role in supporting *HNRNPK* (MIM #600712) as the causative gene for a novel syndrome. We will also describe the distinct clinical features of the syndrome that is associated with *HNRNPK* haploinsufficiency.

HNRNPK encodes the heterogeneous nuclear ribonucleoprotein K (hnRNP K), a nucleic acid binding protein that has diverse roles in the regulation of chromatin remodeling, transcription, RNA stability and splicing, translation, and signal transduction [Bomsztyk et al., 2004; Barboro et al., 2014]. Many of these roles have been implicated in human disease, but none have been identified yet as causative genes for congenital malformation syndromes. Here, we describe two extensively investigated probands without a diagnosis who have been investigated by two separate research groups.

Additional Supporting Information may be found in the online version of this article.

†These authors contributed equally to this work.

‡These authors contributed equally to this work.

§Correspondence to: Antonie D. Kline. E-mail: akline@gbmc.org

*Correspondence to: Dr Antonie D Kline, Harvey institute for medical genetics, Greater baltimore medical center, 6701 N. Charles St, suite 2326, Baltimore, MD 21204. Dr A Micheil Innes, Department of Medical Genetics and Alberta Children's Hospital Research Institute, Alberta Childrens Hospital, 2888 Shaganappi Tr NW, Calgary Alberta Canada, T3B 6A8. E-mail: micheil.innes@albertahealthservices.ca

Contract grant sponsor: Genome Canada; Canadian Institutes of Health Research; Ontario Genomics Institute; Ontario Research Fund; Genome Quebec and Children's Hospital of Eastern Ontario Research Foundation; National Human Genome Research Institute (NIH, 1U54HG006542).

WES independently identified de novo putative loss-of-function variants in *HNRNPK* in each proband. Although both groups were convinced that *HNRNPK* was the correct gene based on familial segregation and in silico analyses, neither group was able to independently confirm the gene as causative. Ultimately, confirmation of pathogenicity was made through a GeneMatcher connection. Both probands with the loss-of-function *HNRNPK* variants have the same distinct clinical syndrome, consisting of intellectual disability associated with characteristic craniofacial features, skeletal, connective tissue, and cardiac anomalies. Additionally, two individuals with deletions of 9q21 encompassing *HNRNPK* have been recently reported in the literature with phenotypes that significantly overlap with our probands [Pua et al., 2014; Hancarova et al., 2015], lending further support that haploinsufficiency of *HNRNPK* leads to a congenital malformation syndrome associated with intellectual disability.

Proband 1 was born to a 33-year-old G3P2 mother and 33-year-old father after an uncomplicated healthy pregnancy. However, prenatal ultrasound identified a two vessel cord and a cardiac septal defect with pericardial effusion. Amniocentesis revealed a normal male karyotype. He was delivered via caesarean for breech presentation at 37 weeks gestation due to preterm labor, and had respiratory distress for 2–3 days. Birth weight was 3.03 kg (50^{th} centile), length 47 cm (10–50^{th} centile), and head circumference 32.5 cm (10–50^{th} centile).

Dysmorphic features were observed at birth. These features included plagiocephaly, shallow orbits with proptotic eyes, abnormal ears, broad nasal root with a dimple at the nasal tip, poorly defined philtrum, excess nuchal skin, cryptorchidism, and a sacral dimple with a coccygeal appendage. He was hypotonic. Postnatal echocardiogram showed two small ventricular septal defects. There was feeding difficulty and failure to thrive in early infancy. At 6 months, he had a prominent metopic suture and bitemporal narrowing. He had surgery for sagittal suture craniosynostosis, and repair of a nasal sinus tract connecting to an inclusion cyst. There were persistent middle ear effusions during childhood, and he currently has mixed conductive and sensorineural hearing loss. At age 12, he had surgery for progressive scoliosis. Spine MRI showed multiple segmentation and fusion defects in the thoracic spine, and extra lumbar vertebral bodies (Fig. 1D). Syrinx was observed at T7–T9 and at T12, with a low lying conus and a lipomyelomeningocele at the terminal thecal sac. Brain MRI was normal.

He was assessed by Medical Genetics at 13 years of age in Calgary. Weight was 36 kg (10^{th} centile), height was 147 cm (above 10^{th} centile), and head circumference was 54.5 cm (just above 50^{th} centile). He appeared to have a dolichostenomelic body habitus. He had dolicocephaly, a long face, long and down-slanting palpebral fissures with proptotic eyes, and his lateral eyebrows were broad. He had a broad prominent nasal bridge and hypoplastic alae nasi (Fig. 1A). His philtrum was long, he had an open bite due to severe malocclusion, and micro and retrognathia. His palate was high arched and with a bifid uvula (Fig. 1C). His tongue had a prominent midline groove (Fig. 1E). His ears were low set with thick helices, and underdeveloped antihelix on the right side. His hands appeared slender with fifth finger clinodactyly (Fig. 1B) and he had planovalgus feet with broad halluces bilaterally, and crowding of toes. His chest was asymmetric due to scoliosis with widely spaced nipples, and he had a small midline abdominal hernia.

Global developmental delay was noticeable by 6 months of age. He walked independently at 4 years of age. Currently, at 17 years of age, he uses sign language and has verbal speech that is difficult to understand to an untrained ear. He requires assistance with activities of daily living such as hygiene and dressing. He was recently assessed by the Weschler Nonverbal Scale of Ability (Canadian) and the Adaptive Behaviour Assessment System (ABAS-II) with full-scale scores in the extremely low range at <0.1 percentile, consistent with moderate cognitive delay. However, he appears to socialize well at school.

On review of family history, he has two older siblings who are healthy with normal development. There is no history of intellectual disability, congenital anomalies, recurrent pregnancy loss, or early death in the family.

Array CGH (60K) showed a maternally inherited deletion of 308 kb at 8q12.3 overlapping a single gene, *ASPH*, which was thought to be noncontributory. Loeys–Dietz and Shprintzen–Goldberg syndromes were considered as possible diagnoses; however, sequencing for *TGFBR1*, *TGFBR2*, and *SKI* was negative.

Proband 2 was born to a healthy 28-year-old G2P1 mother and a 33-year-old father. The pregnancy was uncomplicated. Prenatal ultrasound at 18 weeks showed bilateral renal pelvis dilatation, ventriculomegaly, and nuchal thickening. Amniocentesis revealed a normal male karyotype, with incidental 9qh+. He was born by spontaneous vaginal delivery at over 38 weeks gestation, with no respiratory distress. Birth weight was 3,460 g (90^{th} centile), with a length of 54 cm (>95^{th} centile), and head circumference of 34.5 cm (75^{th} centile).

Multiple anomalies were recognized at birth, including dysmorphic facies with a broad nasal ridge and bifid nasal tip, increased nuchal skin, postaxial polydactyly, sacral dimple, right cryptorchidism, and hypotonia. Postnatal echocardiogram showed a bicuspid aortic valve, mild branch pulmonic stenosis, and patent ductus arteriosus. Renal investigations showed hydronephrosis and vesicoureteral reflux.

He was reassessed at 2 years and again at 8, 9, and 11 years of age in Baltimore. Follow-up echocardiogram revealed resolution of pulmonic stenosis and development of aortic root dilation. He developed osteoporosis in early childhood, and had transient hypothyroidism. He also has severe gastroesophageal reflux, gastrointestinal dysmotility, and requires a jejunostomy feeding tube. He has a neurogenic bladder and has a continent stoma. He has hypertension secondary to chronic kidney disease. He had asthma and nighttime hypoxia in early childhood. Hearing has been normal. Ophthalmology examination showed an optic nerve pit, hyperopia, and megalocornea, and he had corneal abrasions secondary to exposure from lagophthalmos. He had a ridged metopic suture and possible craniosynostosis. He developed mild scoliosis in childhood, and had developmental hip dysplasia, and later coxa valga. Spine MRI showed the presence of extra lumbar vertebrae. He has symptoms of dysautonomia. Brain MRI showed hypomyelination and prominent vasculature suggesting abnormal venous drainage.

On examination at 11 years of age, weight was 31.5 kg (50^{th} centile), with a length of 135.5 cm (10–25^{th} centile), and head circumference of 54.4 cm (75^{th} centile). Metopic suture and occipital shelf were prominent. His eyebrows were sparse laterally. He had long palpebral fissures and mild ptosis. He had a short nose with broad nasal bridge, notched alae nasi, and overhanging columnella. His mouth was held open and was downturned, palate was high with normal uvula, and tongue was tethered and bifid with a prominent midline groove (Fig. 1H and I). His ears were underdeveloped. He had small inverted nipples and a supernumerary nipple. He had pectus excavatum. He was hyperextensible, with restricted extension in some proximal interphalangeal joints. His hands and feet were narrow, and he had caudal displacement of the third digits of his feet.

He has developmental delay and attention-deficit hyperactivity disorder. In early childhood, he had frequent self-stimulatory behavior. At age 11, he uses a device or computer to communicate, and has a few words, but has hundreds of signs. Psychometric testing

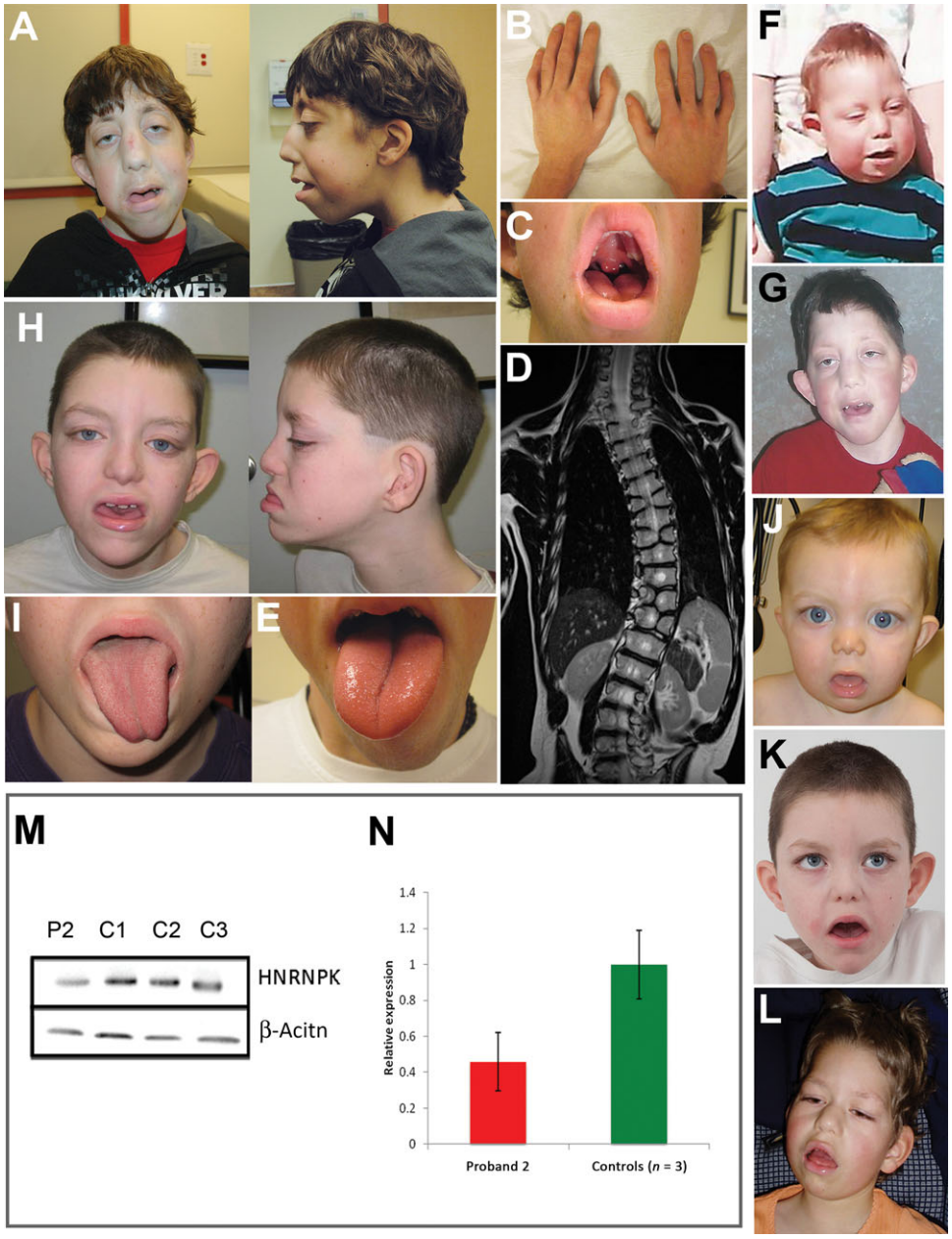

Figure 1. Clinical features of probands with HNRNPK loss-of-function variants. Proband 1 (A–G) and proband 2 (H–K). **A**: Proband 1 at 13 years. **B**: Hands of proband 1, with fifth finger clinodactyly, prominent joints, and shorter distal phalanges. **C**: High palate and bifid uvula in proband 1. **D**: T2 coronal MRI of proband 1 at 12 years, demonstrating multiple vertebral segmentation defects and thoracolumbar scoliosis. **H**: Proband 2 at 11 years. **E** and **I**: Bifid tongues with prominent midline groove in proband 2 (I), and prominent midline groove of tongue in proband 1 (E). **F** and **G**: Proband 1 at ages 14 months and 8 years. **J** and **K**: Proband 2 at 1 year and 8 years. **L**: Individual described by Hancarova et al. (2015) at age 6 years. **M**: Representative western blot demonstrating reduced hnRNPK protein expression in proband 2 fibroblasts (P2) versus control fibroblasts (C1–3). **N**: Comparison of mean hnRNPK protein expression levels determined from three independent western blots, $P = 0.008426$. Republished with permission.

has placed him at approximately a 4–5-year level for cognition. He is not toilet trained and requires assistance with walking.

On review of family history, he has an older brother with attention-deficit disorder and autism with low normal range IQ who is otherwise nondysmorphic and healthy. The family history is otherwise noncontributory.

Investigations have included a normal BAC array in 2007 and a normal SNP array CGH. Kabuki syndrome and TGFβ pathway disorders were considered; however, sequencing for *KMT2D*, *TGFBR1*, and *TGFBR2* were normal. *FLNB* sequencing was also normal.

Families of both proband 1 and proband 2 have provided informed consent, and exome sequencing studies for both families have complied with the ethical regulations of the University of Calgary and Johns Hopkins University institutional review boards, respectively. Exome sequencing of proband 1 and his parents was performed through the Care for Rare Canada consortium. The

Agilent SureSelect 50 Mb capture kit (V3) was used for target enrichment. The Illumina HiSeq platform was used for sequencing, and read alignment, variant calling, and annotation were done as per previous FORGE and Care for Rare Canada projects [Bernier et al., 2012]. Variants were filtered for nonsynonymous variants, rare variants not present in dbSNP, and for de novo variants not present in either parent. Two de novo variants were identified. One variant was a nonframeshift insertion in *NEFH*. The second variant was a frameshift insertion in *HNRNPK*, NM_002140.3:c.953+1dup, which adds a cytosine residue in between the +1 and +2 splice sites, and is predicted to alter gene expression either through nonsense-mediated decay or a frameshift that results in protein truncation, p.(Gly319Argfs*6). The *HNRNPK* variant was confirmed to be present in the proband and absent in the parents by Sanger sequencing. Variants have been submitted to ClinVar: SCV000223813 (c.257G>A) and SCV000223814 (c.953+1dup).

Exome sequencing for proband 2 and his parents was performed through the Baylor Hopkins Center for Mendelian Genomics (BHCMG). The Agilent SureSelect Human All Exon V4 51 Mb Kit was used for target capture. The Illumina HiSeq2000 platform was used for sequencing, read alignment, and variant calling, using methodology that has been previously reported [Hoover-Fong et al., 2014]. We identified potential causal variants by standard filtering criteria as formerly described, followed by the use of the PhenoDB Variant Analysis Tool [Sobreira et al., 2015] to design the prioritization strategy. We prioritized rare functional variants that were de novo, homozygous, compound heterozygous, or X-linked in the proband, and excluded variants with a minor allele frequency >0.01 in dbSNP, the Exome Variant Server (release ESP6500SI-V2), or 1000 Genomes Project [Abecasis et al., 2012], and excluded all variants found in our in-house controls. In the *de novo* variant analysis, we identified a variant in *HNRNPK*, NM_002140.3:c.257G>A. While this sequence change is predicted to alter an amino acid, p.(Arg86His), it occurs in the last codon of exon 5 and is predicted to abolish the splice donor site. This variant was also confirmed to be absent in parents by Sanger sequencing.

There are no loss-of-function variants in *HNRNPK* in the ExAC database (Exome Aggregation Consortium [ExAC], Cambridge, MA [URL: http://exac.broadinstitute.org] [May 2015]), nor in the EVS database (Exome Variant Server, NHLBI GO Exome Sequencing Project [ESP], Seattle, WA [URL: http://evs.gs.washington.edu/EVS/] [May 2015]). No variants have been reported at either of the two residues that have been changed in these two probands, and no deletions involving this gene have been reported in DECIPHER.

Using western blot analysis, we assessed protein expression of hnRNPK in proband 2 versus control fibroblasts (Supp. Methods). The hnRNPK protein was significantly decreased in Proband 2, 48.5% of control levels in fibroblasts ($P < 0.01$, Welch's *t*-test; Fig. 1M and N), supporting the predicted impact of the c.257G>A splice variant on *HNRNPK* gene expression.

There are three alternatively spliced isoforms of hnRNP K described in GenBank, and four isoforms have been reported in the literature and in UniProt [Kimura et al., 2010]. These isoforms either differ in the 5′UTR, have an alternative basic or acidic C-terminus end consisting of either ^{459}SGKFF^{463} or ^{459}ADVEGF^{464}, respectively, or missing amino acid residues between 111 and 134 due to alternative splicing of exon 8. The splice-site variant in proband 2 (c.257G>A) at the junction of exon 6 and intron 6, and the frameshift truncating variant in proband 1 (c.953+1dup, p.(G319Rfs*6)) located at the 5′ end of exon 12, are both expected to disrupt all of these known isoforms based on their location.

HNRNPK was submitted as a candidate gene to Genematcher (https://genematcher.org) independently by research groups in Calgary and Baltimore, in an attempt to find additional probands with variants in *HNRNPK*. GeneMatcher is a Web-based program that was developed with support from the Baylor-Hopkins Center for Mendelian Genomics as part of the Centers for Mendelian Genomics. Candidate genes of interest are posted by investigators, and Genematcher connects investigators sharing a gene of interest via e-mail. Genematcher was successful in facilitating a connection in this case. Subsequent sharing of proband characteristics and careful reappraisal of phenotype, including photographs, revealed that both probands possessed striking similarities, allowing identification of a novel syndrome and *HNRNPK* as the causative gene.

Here, we report the discovery of a new distinct syndrome that is due to loss-of-function variants in the *HNRNPK* gene. Both probands share a common facial phenotype, characterized by long palpebral fissures, ptosis, a broad prominent nasal bridge, hypoplastic alae nasi, an open downturned mouth with a cupid's bow-shaped upper vermilion and full lower lip, ears with underdeveloped and thick helices, and a unique tongue with a prominent median crease (Fig. 1; Table 1). Both probands also share a connective tissue and skeletal phenotype with features that partially overlap with TGFβ pathway-related disorders such as Loeys–Dietz or Shprintzen–Goldberg syndrome. Clinicians in both centers had pursued testing for TGFβ-related syndromes (*TGFBR1*, *TGFBR2* in both probands, and *SKI* in proband 1), although neither proband was completely typical for these disorders [Loeys et al., 2005; Au et al., 2014]. Relevant features included a high palate (with bifid uvula in proband 1), scoliosis with vertebral segment defects and extra lumbar vertebrae, hip dysplasia, hyperextensibility, and craniosynostosis. Proband 2 also had aortic dilation and pectus excavatum. Other interesting features shared by these probands include decreased sweating and mild oligodontia (Table 1).

Two individuals have been reported in the literature with deletions of chromosome 9q21 that encompass the *HNRNPK* gene. Pua et al. (2014) describe a de novo 2.6-Mb deletion of 9q21.32q21.33 encompassing 12 genes, in a delayed hypotonic infant who died at 14 months of age. This infant had facial dysmorphisms, cleft of the soft palate, atrial septal defect and aortic coarctation, bilateral hip dislocation, hindfoot deformity, and delayed myelination on brain MRI [Pua et al., 2014]. More recently, Hancarova et al. (2015) reported a 13-year-old female with a de novo 2-Mb deletion of 9q21.32q21.33 encompassing at least eight genes, with seven genes in the region of overlap with the infant described by Pua et al. (2014), with a very similar facial gestalt to the two probands described in this paper (Fig. 1L). This individual also had atrial septal defect, congenital hip dysplasia, equinovarus foot deformity, severe delay, and hypotonia that later developed into spastic quadraparesis [Hancarova et al., 2015]. Hancarova et al. (2015) had highlighted *NTRK2* as a possible candidate-causative gene for this microdeletion, given that the deletion described by Pua et al. (2014) also deleted *NTRK2* and the role of *NTRK2* in neurodevelopment. While some of the differences observed in these 9q21 microdeletion patients are likely attributed to haploinsufficiency of other genes within the region, there is considerable phenotypic overlap with these microdeletion patients and our two probands with *HNRNPK* putative loss of function variants, particularly the female described by Hancarova et al. (2015). This lends additional support that haploinsufficiency of *HNRNPK* has a critical role in development.

HNRNPK encodes heterogeneous nuclear ribonucleoprotein K, and was initially identified as a component of hnRNP particles. hnRNP K is involved with RNA and/or single strand DNA binding

Table 1. Clinical Features of Probands with Haploinsufficiency of *HNRNPK*

	Proband 1 (male)	Proband 2 (male)	Pua et al. (2014) (female)	Hancarova et al. (2015) (female)
	HNRNPK c.953+1dup p.Gly319Argfs*6	*HNRNPK* c.257G>A p.Arg85His, splice-site variant	Del 9q21.32-q21.33 (2.6 Mb) (85504717-88069314) [HG19]	Del 9q21.32-q21.33 (2Mb) (86595071-88357495) [HG19)]
Growth parameters	At birth: Wt (50 percentile), L (10–50 percentile), HC (10–50 percentile) At 13 years: Wt (10 percentile), Ht (>10 percentile), HC (>50 percentile)	At birth: Wt (90 percentile), L (>95 percentile), HC (75 percentile) At 9 years: Wt (50 percentile), Ht (50–75 percentile), HC (75 percentile)	At birth: Wt (<50 percentile), L (10 percentile), HC (25 percentile) At 6 months: HC <3 percentile	At birth: Wt (<50 percentile), L (25 percentile), HC (<10 percentile) At 13 years: all parameters <3 percentile
Craniofacial	Ridged metopic suture, dolicocephaly sagittal and lambdoid craniosynostosis, long face	Ridged metopic suture, turricephaly long face	Broad face, low anterior hairline	Narrow forehead, low posterior hairline
Eyes	Long downslanting palpebral fissures, proptosis, ptosis, broad lateral eyebrows	Hyperopia, long palpebral fissures, ptosis, optic nerve pit, megalocornea, lagophthalmos, sparse lateral eyebrows	Epicanthal folds	Hypermetropia, long downslanting palpebral fissures
Ears	Underdeveloped helices, hearing loss (conductive and sensorineural)	Underdeveloped thick helices	Low set, cupped right ear	Large, low set ears
Nose	Wide nasal ridge, hypoplastic alae nasi	Wide nasal ridge, cleft of alae nasi	Depressed nasal tip	Wide nasal ridge
Mouth	Open bite, downturned mouth, high palate, bifid uvula, prominent midline groove of tongue, missing molar	Open, downturned mouth, high palate, normal uvula, prominent midline groove/bifid tongue, missing molar and incisor	Cleft soft palate	Open bite, large downturned mouth
Chest	Widely spaced nipples	Pectus excavatum, inverted nipples, supernumerary nipples	NR	NR
CVS/Resp	Two small VSDs	Bicuspid aortic valve, aortic root dilation	Large ASD respiratory difficulty	AVSD, R atrial hemangioma
Gastrointestinal	Constipation in early childhood	GERD, cyclic vomiting, constipation, GI dysmotility, G-tube fed		
Genitourinary	Cryptorchidism	Cryptorchidism, VUR, neurogenic bladder, hydronephrosis	Uterine didelphys	VUR
Skeletal	Hip dysplasia, scoliosis, extra lumbar vertebrae and multiple vertebral segmentation defects, elbow contractures	Hip dysplasia, scoliosis, extra lumbar vertebrae, hyperextensible	Hip dysplasia, proximal upper extremity and distal lower extremity shortening	Hip dysplasia
Hands/feet	Planovalgus feet, large hallux, crowded toes, decreased creases on feet	Postaxial polydactyly, overlapping toes	Hindfoot deformity	Single palmar crease, talipes equinovarus, overlapping toes
Skin	Sacral dimple with coccygeal appendage, decreased sweating, intermittent facial rash	Sacral dimple, decreased sweating, intermittent rash	NR	NR
Neuro	Hypotonia, hyporeflexia, high pain tolerance	Hypotonia, hyporeflexia, high pain tolerance, migraine	Hypotonia	Hypotonia, later spasticity
Development	Mild intellectual disability, walks independently, communicates with many words, short phrases, and uses signs and devices	Mild-moderate intellectual disability, ADHD, walks with assistance, communicates with few words, many signs, and uses devices	Severe delay	Severe intellectual disability
MRI findings	Normal brain, syrinx T7–T9,T12, terminal lipomyelomeningocele	Hypomyelination	Thalamostriate vasculopathy, thin corpus callosum, hypomyelination	NR

NR, not reported; ASD, atrial septal defect; VSD, ventricular septal defect; AVSD, atrioventricular septal defect; VUR, vesicoureteral reflux; GERD, gastroesophageal reflux.

through its three KH domains. HnRNP K also shuttles between the nucleus and the cytoplasm, and has a K-protein interactive domain that mediates interactions with various other proteins. HnRNP K likely acts as a docking platform to allow interaction of kinases and other signal transduction factors with nucleic acid-related cellular activities, and has roles in chromatin remodeling, transcription, RNA stability and splicing, translation, and signal transduction [Bomsztyk et al., 2004; Barboro et al., 2014]. The renin-angiotensin pathway appears to be regulated by hnRNP K at several levels, including hnRNP K regulation of renin mRNA stability [Skalweit et al., 2003], and post-transcriptional control of collagen synthesis downstream of the angiotensin II type I receptor [Thiele et al., 2004]. More recently, hnRNP K has also been implicated in synaptic plasticity through its effects on ERK kinase cascade activation [Folci et al., 2014], and has a role in translational regulation of proteins involved in axonogenesis [Hutchins and Szaro, 2013]. The role of hnRNP K in multiple pathways may explain its effects on multiple organ systems in our probands. Its functions in angiotensin signaling and synaptic plasticity are particularly interesting given the connective tissue and neurodevelopmental features in these probands, as well as the hypertension in patient 2.

The confirmation of *HNRNPK* as the causative gene in both probands would not have been possible without the ability of the investigators to connect with each other. In proband 1, *HNRNPK* had

been identified as a possible candidate gene for more than 2 years, and more traditional methods, such as presentation of the proband at conferences and other attempts at finding a phenotypic match for the proband, had not been successful. Finding cohorts of patients with a very specific phenotype is highly useful when attempting to identify an underlying genetic etiology, and has been a powerful and effective approach in many exome studies aimed at gene discovery. Some examples include the identification of *EZH2* as the causative gene for Weaver syndrome, and identification of *SRCAP* for Floating–Harbor syndrome, which both relied greatly on accurate phenotyping in a cohort of individuals with previously well-recognized "classic" dysmorphic syndromes [Gibson et al., 2012; Hood et al., 2012]. However, this strategy can be difficult if the genetic disease affecting the proband is very rare and/or not previously described, or if the phenotype is subtle or heterogeneous. Furthermore, online tools that depend on matching by phenotype, such as PhenomeCentral [Girdea et al., 2013], can be limited by clinician interpretation of what features are most significant, which may lead to discrepancies in descriptions of phenotypes that are actually similar, particularly if features are nonspecific. This may affect the ability to match, even if candidate genes are incorporated into the PhenomeCentral match algorithm. GeneMatcher allows a different approach, where investigators can search for potential matches based on the gene without information about phenotype, circumventing some of the bias and limitations of looking for matches based primarily on descriptions of phenotype. Further analysis of phenotype after matching can subsequently be done in reverse to determine whether the match is appropriate. However, it would be important to be aware of possible pitfalls from this approach as well. For example, certain genes may be implicated in multiple distinct phenotypes depending on the nature of the variant, and these scenarios may complicate gene matching and reverse phenotyping. Additionally, it is possible that phenotyping exclusively in reverse based on a candidate gene may lead to unintentional bias or a clinical assessment that is overly focused. Therefore, patients would ideally be well phenotyped initially prior to candidate gene identification, with further assessment of shared characteristics to be performed after a gene match is made.

Nevertheless, there is an imminent need for effective ways to elucidate unsolved exomes as increasing numbers of patients undergo WES, and as exome sequencing moves further into clinical practice. Therefore, the ability to make matches based on gene and/or phenotype to validate causative variants is becoming extremely important. We have demonstrated the value and utility of this new approach to matching through the identification of a new congenital malformation syndrome associated with intellectual disability due to haploinsufficiency of *HNRNPK*. Our probands have vertebral defects, connective tissue abnormalities, structural congenital heart defects, and genitourinary abnormalities. Finally, the facial gestalt is particularly distinct, characterized by long palpebral fissures, ptosis, a broad prominent nasal bridge, prominent underdeveloped ears, and an open downturned mouth with a prominent median crease to the tongue.

Acknowledgments

We are grateful to the families for participating in this project. The authors acknowledge intellectual contributions from all members of the Baylor-Hopkins Center for Mendelian Genomics (BHCMG). This work was performed under the Care4Rare Canada Consortium. We acknowledge the contribution of the high-throughput sequencing platform of the McGill University and Genome Quebec Innovation Centre, Montreal, Canada. We would like to thank Taila Hartley (Clinical Coordinator) and Chandree Beaulieu (Project Manager) at the Children's Hospital of Eastern Ontario Research Institute for their contribution to the infrastructure of Care4Rare. We would like to acknowledge Hancarova et al. (2015) for the use of their patient photograph.

References

Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. 2012. An integrated map of genetic variation from 1092 human genomes. Nature 491:56–65.

Au PY, Racher HE, Graham JM Jr, Kramer N, Lowry RB, Parboosingh JS, Innes AM. 2014. De novo exon 1 missense mutations of SKI and Shprintzen–Goldberg syndrome: two new cases and a clinical review. Am J Med Genet A 164A:676–684.

Barboro P, Ferrari N, Balbi C. 2014. Emerging roles of heterogeneous nuclear ribonucleoprotein K (hnRNP K) in cancer progression. Cancer Lett 352:152–159.

Bernier FP, Caluseriu O, Ng S, Schwartzentruber J, Buckingham KJ, Innes AM, Jabs EW, Innis JW, Schuette JL, Gorski JL, Byers PH, Andelfinger G, et al. 2012. Haploinsufficiency of SF3B4, a component of the pre-mRNA spliceosomal complex, causes Nager syndrome. Am J Hum Genet 90:925–933.

Bomsztyk K, Denisenko O, Ostrowski J. 2004. hnRNP K: one protein multiple processes. Bioessays 26:629–638.

Folci A, Mapelli L, Sassone J, Prestori F, D'Angelo E, Bassani S, Passafaro M. 2014. Loss of hnRNP K impairs synaptic plasticity in hippocampal neurons. J Neurosci 34:9088–9095.

Gibson WT, Hood RL, Zhan SH, Bulman DE, Fejes AP, Moore R, Mungall AJ, Eydoux P, Babul-Hirji R, An J, Marra MA, Chitayat D, et al. 2012. Mutations in EZH2 cause Weaver syndrome. Am J Hum Genet 90:110–118.

Girdea M, Dumitriu S, Fiume M, Bowdin S, Boycott KM, Chenier S, Chitayat D, Faghfoury H, Meyn MS, Ray PN, So J, Stavropoulos DJ, et al. 2013. PhenoTips: patient phenotyping software for clinical and research use. Hum Mutat 34:1057–1065.

Hancarova M, Puchmajerova A, Drabova J, Karaskova E, Vlckova M, Sedlacek Z. 2015. Deletions of 9q21.3 including NTRK2 are associated with severe phenotype. Am J Med Genet A 167:264–267.

Hood RL, Lines MA, Nikkel SM, Schwartzentruber J, Beaulieu C, Nowaczyk MJ, Allanson J, Kim CA, Wieczorek D, Moilanen JS, Lacombe D, Gillessen-Kaesbach G, et al. 2012. Mutations in SRCAP, encoding SNF2-related CREBBP activator protein, cause Floating-Harbor syndrome. Am J Hum Genet 90:308–313.

Hoover-Fong J, Sobreira N, Jurgens J, Modaff P, Blout C, Moser A, Kim OH, Cho TJ, Cho SY, Kim SJ, Jin DK, Kitoh H, et al. 2014. Mutations in PCYT1A, encoding a key regulator of phosphatidylcholine metabolism, cause spondylometaphyseal dysplasia with cone-rod dystrophy. Am J Hum Genet 94:105–112.

Hutchins EJ, Szaro BG. 2013. c-Jun N-terminal kinase phosphorylation of heterogeneous nuclear ribonucleoprotein K regulates vertebrate axon outgrowth via a posttranscriptional mechanism. J Neurosci 33:14666–14680.

Kimura Y, Nagata K, Suzuki N, Yokoyama R, Yamanaka Y, Kitamura H, Hirano H, Ohara O. 2010. Characterization of multiple alternative forms of heterogeneous nuclear ribonucleoprotein K by phosphate-affinity electrophoresis. Proteomics 10:3884–3895.

Loeys BL, Chen J, Neptune ER, Judge DP, Podowski M, Holm T, Meyers J, Leitch CC, Katsanis N, Sharifi N, Xu FL, Myers LA, et al. 2005. A syndrome of altered cardiovascular, craniofacial, neurocognitive and skeletal development caused by mutations in TGFBR1 or TGFBR2. Nat Genet 37:275–281.

Pua HH, Krishnamurthi S, Farrell J, Margeta M, Ursell PC, Powers M, Slavotinek AM, Jeng LJ. 2014. Novel interstitial 2.6 Mb deletion on 9q21 associated with multiple congenital anomalies. Am J Med Genet A 164A:237–242.

Skalweit A, Doller A, Huth A, Kahne T, Persson PB, Thiele BJ. 2003. Posttranscriptional control of renin synthesis: identification of proteins interacting with renin mRNA 3'-untranslated region. Circ Res 92:419–427.

Sobreira N, Schiettecatte F, Boehm C, Valle D, Hamosh A. 2015. New tools for Mendelian disease gene identification: PhenoDB variant analysis module; and GeneMatcher, a web-based tool for linking investigators with an interest in the same gene. Hum Mutat 36:425–431.

Thiele BJ, Doller A, Kahne T, Pregla R, Hetzer R, Regitz-Zagrosek V. 2004. RNA-binding proteins heterogeneous nuclear ribonucleoprotein A1, E1, and K are involved in post-transcriptional control of collagen I and III synthesis. Circ Res 95:1058–1066.