10.4.2 Genetic Information

Autosomal recessive inheritance was considered based on the presence of this syndrome in consanguineous families [11, 27, 31]. Three independent groups have performed homozygosity mapping and/or linkage analysis and each showed that the gene carbohydrate (*N*-acetylgalactosamine 4-*O*) sulfotransferase 14 (*CHST14*, NM_130468.3) was responsible for this syndrome [11, 27, 31].

The *CHST14* gene was first cloned by Evers et al. [13]. It contains one coding exon (1,131-bp open reading frame) and is localized at 15q15.1. This gene encodes D4ST1, a 376 amino acid type II transmembrane protein (molecular weight: 43 kDa), that is localized in the Golgi membrane. It transfers a sulfate group from 3′-phosphoadenosine 5′-phosphosulfate to position 4 of the GalNAc residues in dermatan to generate DS (Figs. 10.1a and 10.3a). Northern blotting revealed that *CHST14* is mainly expressed in heart, placenta, liver, and pancreas, and is weakly expressed in lung, skeletal muscle, and kidney [13].

To date, 11 pathogenic mutations of *CHST14* have been identified: p.Val49*, p.Lys69*, p.Arg135_Leu137delinsGlyThrGln, p.Phe209Ser, p.Arg213Pro, p.Lys226Alafs*16, p.Arg274Pro, p.Pro281Leu, p.Cys289Ser, p.Tyr293Cys, and p.Glu334Glyfs*107 [11, 27, 30, 31, 46, 48]. (p.Val48* was corrected to p.Val49*; Erratum in Am J Med Genet Part A 161A(2):403 (2013)) (p.Arg135Gly and p.Leu137Gln were originally reported by Dündar et al., but lately registered as c.403_410delCGCACCCTinsGGCACCCA, p.Arg135_Leu137delinsGlyThrGln in The Human Gene Mutation Database: https://portal.biobase-international.com/hgmd/pro/genesearch.php). Because these are protein truncation mutations and missense mutations, it seems likely that the mutations cause a loss of function.

10.4.3 Biochemical Information

Dündar et al. reported that DS-derived IdoA-GalNAc(4S) disaccharide was undetectable in fibroblasts derived from a patient with a homozygous p.Arg213Pro mutation. They also reported that GlcA-GalNAc(4S) content was greatly increased in the fibroblast extract and the culture media obtained from cultures of fibroblasts derived from this patient as compared with control fibroblasts [11]. It was also found that the amount of nonsulfated disaccharides (GlcA-GalNAc and IdoA-GalNAc) was increased in the cell extract and its media from the patient's fibroblast as compared with normal control fibroblasts. From these results, Dündar et al. proposed that epimerization of GlcA to IdoA by C5-carboxy epimerase is followed by sulfation of the C4 hydroxyl on the adjacent GalNAc residue by D4ST1. This process generates DS from dermatan and prevents back-epimerization from IdoA to GlcA [11, 28].

Miyake et al. measured the sulfotransferase activity of COS7 cells transfected with wild-type and mutant D4ST1 harboring the p.Lys69*, p.Pro281Leu, p.Cys289Ser, or p.Tyr293Cys mutations. The enzyme activity of the mutants was as low as that in mock transfected cells, suggesting that these missense mutations result in the loss of function [31]. The disaccharide composition of the decorin GAG chain isolated from the patient's fibroblasts consisted only of CS, without DS, while the chains isolated from normal fibroblasts consisted of CS/DS hybrid chains [31]. Furthermore, the level of nonsulfated dermatan was negligible in the patient's fibroblasts [31]. Thus, in this syndrome, the CS/DS chain is replaced with the CS chain, even though the core proteins are normal.

10.4.4 Pathology and Pathophysiology

Of the major DS proteoglycans in skin, decorin was a focus of research because it binds to collagen fibrils via its core protein and its GAG chains act as interfibrillar bridges [38, 39]. Three α collagen chains are self-assembled to generate tropocollagen, in the form of a triple helix. Tropocollagen then self-assembles to form collagen fibrils via decorin (Fig. 10.5a). Collagen fibrils are assembled into a collagen fiber, known as the collagen bundle, via the antiparallel com-

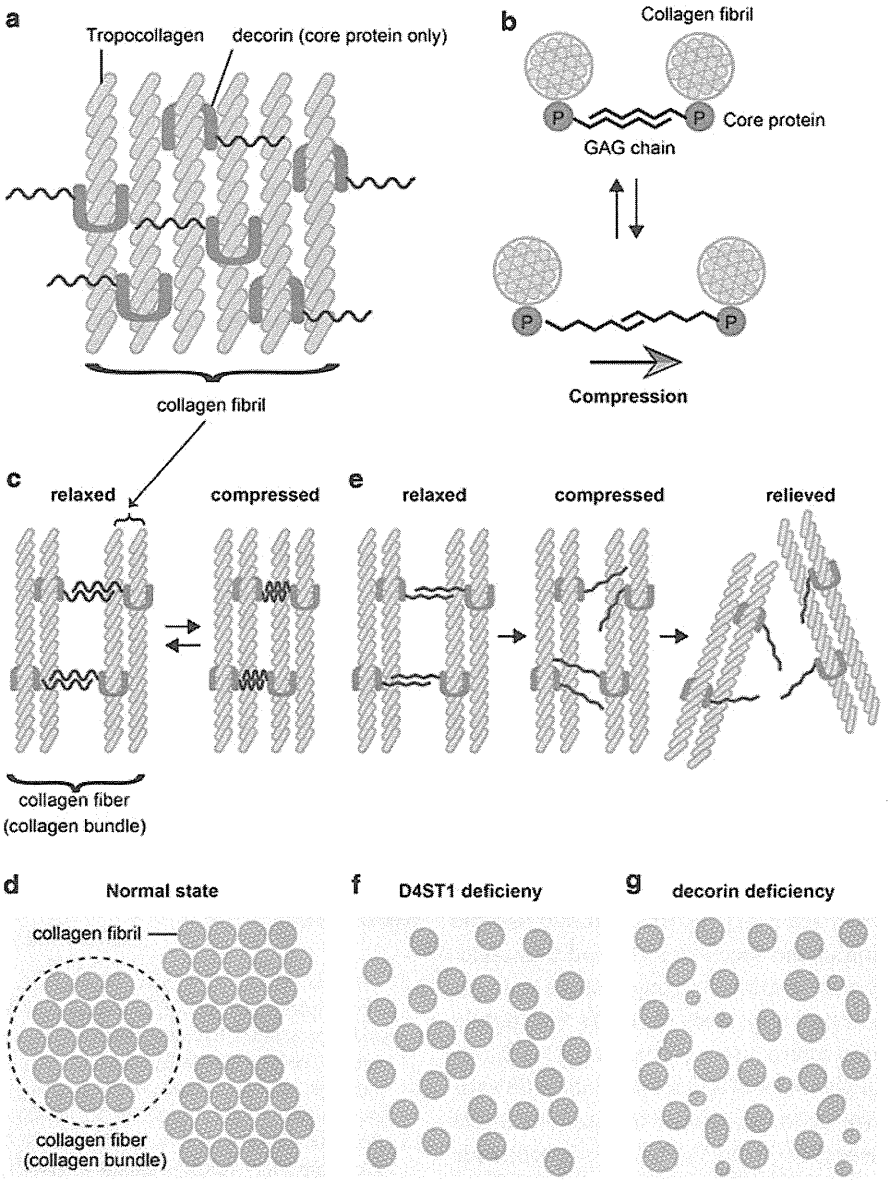

Fig. 10.5 Putative model of abnormal collagen bundle assembly in D4ST1-deficient EDS. (**a**) Tropocollagen directly binds to decorin and forms a collagen fibril. The *blue lines* represent the CS/DS hybrid chain. (**b**) Illustration of the sliding filament model showing reversible longitudinal slippage between the antiparallel GAG chains. The *black lines* represent unspecified GAGs. (**c**, **d**) In the normal state, the CS/DS chains can bend against the direction of mechanical compression and rebound to the original structure (**c**). Thus, the collagen bundles are refractory to compression stress (**d**). (**e**, **f**) In D4ST1-deficient EDS, the CS/DS chains are replaced with CS chains (*red lines*). These chains cannot resist mechanical compression, resulting in irreversible scattering of the collagen fibrils. (**g**) The size and shape of the collagen fibrils are highly variable in decorin-deficient mice

plex of the CS/DS hybrid GAG chains of decorin, which acts like a bridge to provide a space between individual fibrils and tighten the collagen fiber (Fig. 10.5c, d).

The GAGs span collagen fibrils in the extracellular matrix of skin and tendons, and the length of the GAG chain determines the width of the interfibrillar gap [35, 36]. Elasticity of the

extracellular matrix is explained by the sliding filament model, which allows reversible longitudinal slippage between the antiparallel GAG chains (Fig. 10.5b) [39]. Because tissue stability and elasticity depend on the structure of the GAG bridges, irreversible damage can occur if the bridges are inelastic [39].

Decorin is composed of a horseshoe-shaped core protein (molecular weight: ~45 kDa) and a single CS/DS hybrid chain on the N-terminal side (Fig. 10.5a) [22, 47]. Weber et al. reported that the model structure of decorin consists of an arch in which the inner concave surface is formed from a curved β-sheet and the outer convex surface is formed from α-helices. They also proposed that one tropocollagen fiber lies within the decorin convex and another interacts with one arm of the arch [47]. The IdoA:GlcA ratio in DS ranges from ~10 to >90 % depending on the tissue type [39]. Importantly, L-IdoA residues in DS can easily undergo conformational changes, unlike GlcA in CS [6, 7]. Thus, the IdoA:GlcA ratio should be higher in more flexible tissues [39].

Light microscopic investigation of skin specimens from two patients showed that fine collagen fibers were predominant in the reticular to papillary dermis and the number of thick collagen bundles was markedly reduced [31]. Electron microscopic examination of the specimens showed that collagen fibrils were dispersed throughout the reticular dermis, whereas they were regularly and tightly assembled in control tissue. Surprisingly, each collagen fibril was smooth and round, with little variation in size or shape, similar to the fibril in the control tissue (Fig. 10.5d, f) [31]. The disaccharide composition of the decorin GAG chain from a patient's fibroblasts only consisted of CS, without DS disaccharide, whereas control fibroblasts consisted of a CS/DS hybrid [31]. The transition of decorin from the CS/DS hybrid chain to a CS chain probably decreases the flexibility of the GAG chain. The sliding filament model proposes that mechanical compression might also work in the CS chain of D4ST1-deficient patients, but the inflexibility of the CS chain is unable to tolerate higher mechanical pressures or is too inelastic to maintain normal skin properties (Fig. 10.5e, f). This irreversible event could explain the progressive clinical course of this disease.

Interestingly, there were marked variations in the size and shape of dermal collagen fibrils in decorin-null mice (Fig. 10.5g) [8]. These findings suggest that the decorin core protein is important for collagen fibril formation, and that the CS/DS hybrid chain of decorin PG regulates the space between the collagen fibrils and form collagen bundles, as previously reported [37]. These findings suggest that the main pathological basis of this disorder could be insufficient assembly of collagen fibrils.

However, Dündar et al. reported that the light microscopic and electron microscopic findings of a patient's skin were unchanged compared to the normal control [11]. Malfait et al. reported that, in their patient, most collagen bundles were small diameter in size, and some were composed of collagen fibrils of varying diameter that were separated by irregular interfibrillar spaces [27]. In addition, the fibroblasts exhibited an elongated and/or dilated endoplasmic reticulum. So far, definitive histopathologic characteristics have not been established, so further studies are strongly encouraged to determine the major histological characteristics and underlying pathophysiology of this disorder.

References

1. Almeida R, Levery SB, Mandel U, Kresse H, Schwientek T, Bennett EP, Clausen H (1999) Cloning and expression of a proteoglycan UDP-galactose:beta-xylose beta1,4-galactosyltransferase I. A seventh member of the human beta4-galactosyltransferase gene family. J Biol Chem 274:26165–26171
2. Beighton P, De Paepe A, Steinmann B, Tsipouras P, Wenstrup RJ (1998) Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). Am J Med Genet 77:31–37
3. Bishop JR, Schuksz M, Esko JD (2007) Heparan sulphate proteoglycans fine-tune mammalian physiology. Nature 446:1030–1037
4. Bui C, Talhaoui I, Chabel M, Mulliert G, Coughtrie MW, Ouzzine M, Fournel-Gigleux S (2010) Molecular characterization of beta1,4-galactosyltransferase 7 genetic mutations linked to the progeroid form of Ehlers-Danlos syndrome (EDS). FEBS Lett 584:3962–3968
5. Bulow HE, Hobert O (2006) The molecular diversity of glycosaminoglycans shapes animal development. Ann Rev Cell Dev Biol 22:375–407
6. Casu B, Petitou M, Provasoli M, Sinay P (1988) Conformational flexibility: a new concept for explaining binding and biological properties of iduronic acid-

containing glycosaminoglycans. Trends Biochem Sci 13:221–225
7. Catlow KR, Deakin JA, Wei Z, Delehedde M, Fernig DG, Gherardi E, Gallagher JT, Pavao MS, Lyon M (2008) Interactions of hepatocyte growth factor/scatter factor with various glycosaminoglycans reveal an important interplay between the presence of iduronate and sulfate density. J Biol Chem 283:5235–5248
8. Danielson KG, Baribault H, Holmes DF, Graham H, Kadler KE, Iozzo RV (1997) Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. J Cell Biol 136:729–743
9. Dündar M, Demiryilmaz F, Demiryilmaz I, Kumandas S, Erkilic K, Kendirci M, Tuncel M, Ozyazgan I, Tolmie JL (1997) An autosomal recessive adducted thumb-club foot syndrome observed in Turkish cousins. Clin Genet 51:61–64
10. Dündar M, Kurtoglu S, Elmas B, Demiryilmaz F, Candemir Z, Ozkul Y, Durak AC (2001) A case with adducted thumb and club foot syndrome. Clin Dysmorphol 10:291–293
11. Dündar M, Müller T, Zhang Q, Pan J, Steinmann B, Vodopiutz J, Gruber R, Sonoda T, Krabichler B, Utermann G, others (2009) Loss of dermatan-4-sulfotransferase 1 function results in adducted thumb-clubfoot syndrome. Am J Hum Genet 85:873–882
12. Esko JD, Kimata K, Lindahl U (2009) Proteoglycans and sulfated glycosaminoglycans. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (eds) Essentials of glycobiology. Cold Spring Harbor, New York
13. Evers MR, Xia G, Kang HG, Schachner M, Baenziger JU (2001) Molecular cloning and characterization of a dermatan-specific N-acetylgalactosamine 4-O-sulfotransferase. J Biol Chem 276:36344–36353
14. Faiyaz-Ul-Haque M, Zaidi SH, Al-Ali M, Al-Mureikhi MS, Kennedy S, Al-Thani G, Tsui LC, Teebi AS (2004) A novel missense mutation in the galactosyltransferase-I (B4GALT7) gene in a family exhibiting facioskeletal anomalies and Ehlers-Danlos syndrome resembling the progeroid type. Am J Med Genet Part A 128A:39–45
15. Freeze HH (2006) Genetic defects in the human glycome. Nat Rev Genet 7:537–551
16. Hernandez A, Aguirre-Negrete MG, Gonzalez-Flores S, Reynoso-Luna MC, Fragoso R, Nazara Z, Tapia-Arizmendi G, Cantu JM (1986) Ehlers-Danlos features with progeroid facies and mild mental retardation. Further delineation of the syndrome. Clin Genet 30:456–461
17. Hernandez A, Aguirre-Negrete MG, Liparoli JC, Cantu JM (1981) Third case of a distinct variant of the Ehlers-Danlos Syndrome (EDS). Clin Genet 20:222–224
18. Hernandez A, Aguirre-Negrete MG, Ramirez-Soltero S, Gonzalez-Mendoza A, Martinez y Martinez R, Velazquez-Cabrera A, Cantu JM (1979) A distinct variant of the Ehlers-Danlos syndrome. Clin Genet 16:335–339
19. Jaeken J, Hennet T, Freeze HH, Matthijs G (2008) On the nomenclature of congenital disorders of glycosylation (CDG). J Inherit Metab Dis 31:669–672
20. Janecke AR, Baenziger JU, Müller T, Dündar M (2011) Loss of dermatan-4-sulfotransferase 1 (D4ST1/CHST14) function represents the first dermatan sulfate biosynthesis defect, "dermatan sulfate-deficient adducted thumb-club-foot syndrome". Hum Mutat 32:484–485
21. Janecke AR, Unsinn K, Kreczy A, Baldissera I, Gassner I, Neu N, Utermann G, Müller T (2001) Adducted thumb-club foot syndrome in sibs of a consanguineous Austrian family. J Med Genet 38:265–269
22. Kobe B, Deisenhofer J (1993) Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. Nature 366:751–756
23. Kosho T, Miyake N, Hatamochi A, Takahashi J, Kato H, Miyahara T, Igawa Y, Yasui H, Ishida T, Ono K, others (2010) A new Ehlers-Danlos syndrome with craniofacial characteristics, multiple congenital contractures, progressive joint and skin laxity, and multisystem fragility-related manifestations. Am J Med Genet Part A 152A:1333–1346
24. Kosho T, Miyake N, Mizumoto S, Hatamochi A, Fukushima Y, Yamada S, Sugahara K, Matsumoto N (2011) A response to: loss of dermatan-4-sulfotransferase 1 (D4ST1/CHST14) function represents the first dermatan sulfate biosynthesis defect, "dermatan sulfate-deficient Adducted Thumb-Clubfoot Syndrome". Which name is appropriate, "Adducted Thumb-Clubfoot Syndrome" or "Ehlers-Danlos syndrome"? Hum Mutat 32:1507–1509
25. Kosho T, Takahashi J, Ohashi H, Nishimura G, Kato H, Fukushima Y (2005) Ehlers-Danlos syndrome type VIB with characteristic facies, decreased curvatures of the spinal column, and joint contractures in two unrelated girls. Am J Med Genet Part A 138A:282–287
26. Kresse H, Rosthoj S, Quentin E, Hollmann J, Glossl J, Okada S, Tonnesen T (1987) Glycosaminoglycan-free small proteoglycan core protein is secreted by fibroblasts from a patient with a syndrome resembling progeroid. Am J Hum Genet 41:436–453
27. Malfait F, Syx D, Vlummens P, Symoens S, Nampoothiri S, Hermanns-Le T, Van Laer L, De Paepe A (2010) Musculocontractural Ehlers-Danlos Syndrome (former EDS type VIB) and adducted thumb clubfoot syndrome (ATCS) represent a single clinical entity caused by mutations in the dermatan-4-sulfotransferase 1 encoding CHST14 gene. Hum Mutat 31:1233–1239
28. Malmström A (1984) Biosynthesis of dermatan sulfate. II. Substrate specificity of the C-5 uronosyl epimerase. J Biol Chem 259:161–165
29. Mao JR, Bristow J (2001) The Ehlers-Danlos syndrome: on beyond collagens. J Clin Invest 107:1063–1069
30. Mendoza-Londono R, Chitayat D, Kahr WH, Hinek A, Blaser S, Dupuis L, Goh E, Badilla-Porras R, Howard A, Mittaz L, others (2012) Extracellular matrix and platelet function in patients with musculo-

contractural Ehlers-Danlos syndrome caused by mutations in the CHST14 gene. Am J Med Genet Part A 158A:1344–1354
31. Miyake N, Kosho T, Mizumoto S, Furuichi T, Hatamochi A, Nagashima Y, Arai E, Takahashi K, Kawamura R, Wakui K, others (2010) Loss-of-function mutations of CHST14 in a new type of Ehlers-Danlos syndrome. Hum Mutat 31:966–974
32. Nakajima M, Mizumoto S, Miyake N, Kogawa R, Iida A, Ito H, Kitoh H, Hirayama A, Mitsubuchi H, Miyazaki O, others (2013) Mutations in B3GALT6, which encodes a glycosaminoglycan linker region enzyme, cause a spectrum of skeletal and connective tissue disorders. Am J Hum Genet 92:927–934
33. Okajima T, Fukumoto S, Furukawa K, Urano T (1999) Molecular basis for the progeroid variant of Ehlers-Danlos syndrome. Identification and characterization of two mutations in galactosyltransferase I gene. J Biol Chem 274:28841–28844
34. Okajima T, Yoshida K, Kondo T, Furukawa K (1999) Human homolog of Caenorhabditis elegans sqv-3 gene is galactosyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. J Biol Chem 274:22915–22918
35. Scott JE (1988) Proteoglycan-fibrillar collagen interactions. Biochem J 252:313–323
36. Scott JE (1992) Morphometry of cupromeronic blue-stained proteoglycan molecules in animal corneas, versus that of purified proteoglycans stained in vitro, implies that tertiary structures contribute to corneal ultrastructure. J Anat 180(Pt 1):155–164
37. Scott JE (1995) Extracellular matrix, supramolecular organisation and shape. J Anat 187(Pt 2):259–269
38. Scott JE (1996) Proteodermatan and proteokeratan sulfate (decorin, lumican/fibromodulin) proteins are horseshoe shaped. Implications for their interactions with collagen. Biochemistry 35:8795–8799
39. Scott JE (2003) Elasticity in extracellular matrix 'shape modules' of tendon, cartilage, etc. A sliding proteoglycan-filament model. J Physiol 553:335–343
40. Seidler DG, Faiyaz-Ul-Haque M, Hansen U, Yip GW, Zaidi SH, Teebi AS, Kiesel L, Gotte M (2006) Defective glycosylation of decorin and biglycan, altered collagen structure, and abnormal phenotype of the skin fibroblasts of an Ehlers-Danlos syndrome patient carrying the novel Arg270Cys substitution in galactosyltransferase I (beta4GalT-7). J Mol Med (Berl) 84:583–594
41. Shimizu K, Okamoto N, Miyake N, Taira K, Sato Y, Matsuda K, Akimaru N, Ohashi H, Wakui K, Fukushima Y, others (2011) Delineation of dermatan 4-O-sulfotransferase 1 deficient Ehlers-Danlos syndrome: observation of two additional patients and comprehensive review of 20 reported patients. Am J Med Genet Part A 155A:1949–1958
42. Sisu E, Flangea C, Serb A, Zamfir AD (2011) Modern developments in mass spectrometry of chondroitin and dermatan sulfate glycosaminoglycans. Amino Acids 41:235–256
43. Sonoda T, Kouno K (2000) Two brothers with distal arthrogryposis, peculiar facial appearance, cleft palate, short stature, hydronephrosis, retentio testis, and normal intelligence: a new type of distal arthrogryposis? Am J Med Genet 91:280–285
44. Steinmann B, Royce PM, Superti-Furga A (2002) The Ehlers -Danlos syndrome. In: Royce PM and Steinmann B (eds) Connective tissue and its heritable disorders. (2nd edition) Wiley-Liss, Inc., New York
45. Sugahara K, Mikami T, Uyama T, Mizuguchi S, Nomura K, Kitagawa H (2003) Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. Curr Opin Struct Biol 13:612–620
46. Voermans NC, Kempers M, Lammens M, van Alfen N, Janssen MC, Bonnemann C, van Engelen BG, Hamel BC (2012) Myopathy in a 20-year-old female patient with D4ST-1 deficient Ehlers-Danlos syndrome due to a homozygous CHST14 mutation. Am J Med Genet Part A 158A:850–855
47. Weber IT, Harrison RW, Iozzo RV (1996) Model structure of decorin and implications for collagen fibrillogenesis. J Biol Chem 271:31767–31770
48. Winters KA, Jiang Z, Xu W, Li S, Ammous Z, Jayakar P, Wierenga KJ (2012) Re-assigned diagnosis of D4ST1-deficient Ehlers-Danlos syndrome (adducted thumb-clubfoot syndrome) after initial diagnosis of Marden-Walker syndrome. Am J Med Genet A 158A:2935–2940
49. Yasui H, Adachi Y, Minami T, Ishida T, Kato Y, Imai K (2003) Combination therapy of DDAVP and conjugated estrogens for a recurrent large subcutaneous hematoma in Ehlers-Danlos syndrome. Am J Hematol 72:71–72

REPORT

Identification of *KLHL41* Mutations Implicates BTB-Kelch-Mediated Ubiquitination as an Alternate Pathway to Myofibrillar Disruption in Nemaline Myopathy

Vandana A. Gupta,[1] Gianina Ravenscroft,[2] Ranad Shaheen,[3] Emily J. Todd,[2] Lindsay C. Swanson,[1] Masaaki Shiina,[4] Kazuhiro Ogata,[4] Cynthia Hsu,[1] Nigel F. Clarke,[5] Basil T. Darras,[6] Michelle A. Farrar,[7] Amal Hashem,[3] Nicholas D. Manton,[8] Francesco Muntoni,[9] Kathryn N. North,[10] Sarah A. Sandaradura,[5] Ichizo Nishino,[11] Yukiko K. Hayashi,[11] Caroline A. Sewry,[9] Elizabeth M. Thompson,[12,13] Kyle S. Yau,[2] Catherine A. Brownstein,[1] Timothy W. Yu,[1] Richard J.N. Allcock,[14] Mark R. Davis,[15] Carina Wallgren-Pettersson,[16] Naomichi Matsumoto,[17] Fowzan S. Alkuraya,[3] Nigel G. Laing,[2] and Alan H. Beggs[1,*]

Nemaline myopathy (NM) is a rare congenital muscle disorder primarily affecting skeletal muscles that results in neonatal death in severe cases as a result of associated respiratory insufficiency. NM is thought to be a disease of sarcomeric thin filaments as six of eight known genes whose mutation can cause NM encode components of that structure, however, recent discoveries of mutations in non-thin filament genes has called this model in question. We performed whole-exome sequencing and have identified recessive small deletions and missense changes in the Kelch-like family member 41 gene (*KLHL41*) in four individuals from unrelated NM families. Sanger sequencing of 116 unrelated individuals with NM identified compound heterozygous changes in *KLHL41* in a fifth family. Mutations in *KLHL41* showed a clear phenotype-genotype correlation: Frameshift mutations resulted in severe phenotypes with neonatal death, whereas missense changes resulted in impaired motor function with survival into late childhood and/or early adulthood. Functional studies in zebrafish showed that loss of Klhl41 results in highly diminished motor function and myofibrillar disorganization, with nemaline body formation, the pathological hallmark of NM. These studies expand the genetic heterogeneity of NM and implicate a critical role of BTB-Kelch family members in maintenance of sarcomeric integrity in NM.

Nemaline myopathy (NM) is a rare congenital disorder primarily affecting skeletal muscle function. Clinically, NM is a heterogeneous group of myopathies of variable severity.[1,2] The "severe" congenital form of NM presents with reduced or absent spontaneous movements in utero leading to severe contractures or fractures at birth and respiratory insufficiency leading to early mortality. Individuals with the "intermediate" congenital form of NM have antigravity movement and independent respiration at delivery but exhibit delayed motor milestones and require ventilatory support later in life. The "typical" congenital form of NM usually presents in the neonatal period or first year of life with hypotonia, weakness, and feeding difficulties with less prominent respiratory involvement. In these cases, the disease is usually static or very slowly progressive, and many individuals remain ambulant for much of their lives.[3] The defining diagnostic feature of all forms of NM, irrespective of genetic mutation, is the presence of numerous red-staining rods with Gomori trichrome stain that appear as rod-shaped electron-dense structures termed "nemaline bodies" by electron microscopy.[4] These nemaline bodies are most frequently cytoplasmic; however, the presence of intranuclear rods has also been reported.[5]

NM is a genetically heterogeneous condition, and mutations in eight different genes have been identified that are associated with dominant and/or recessive forms of this disease.[6–13] Mutations in these genes cause about 75%–80% of NM cases, suggesting the involvement of additional unidentified genes in disease etiology.

[1]Division of Genetics and Genomics, The Manton Center for Orphan Disease Research, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA; [2]Western Australian Institute for Medical Research and the Centre for Medical Research, University of Western Australia, Nedlands, Western Australia 6009, Australia; [3]Developmental Genetics Unit, King Faisal Specialist Hospital and Research Center, Riyadh 11211, Saudi Arabia; [4]Department of Biochemistry, Yokohama City University, Graduate School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004, Japan; [5]Institute for Neuroscience and Muscle Research, Children's Hospital at Westmead and Discipline of Paediatrics and Child Health, University of Sydney, Sydney 2145, Australia; [6]Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA; [7]Department of Neurology, Sydney Children's Hospital, Randwick NSW 2032, Australia; [8]Department of Surgical Pathology, SA Pathology at the Women's and Children's Hospital, North Adelaide, South Australia 5006; [9]Dubowitz Neuromuscular Centre, Institute of Child Health and Great Ormond Street Hospital, London WC1N 1EH, UK; [10]Murdoch Childrens Research Institute, The Royal Children's Hospital, Parkville, Victoria 3052, Australia; [11]Department of Neuromuscular Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo 187-8502, Japan; [12]Department of Paediatrics, University of Adelaide, Adelaide, South Australia 5000, Australia; [13]SA Clinical Genetics Service, SA Pathology at the Women's and Children's Hospital, North Adelaide, South Australia 5006, Australia; [14]Lotterywest State Biomedical Facility Genomics and School of Pathology and Laboratory Medicine, University of Western Australia, Perth, Western Australia 6000, Australia; [15]Department of Anatomical Pathology, Royal Perth Hospital, Perth, Western Australia 6000, Australia; [16]The Folkhälsan Institute of Genetics, Samfundet Folkhälsan, Biomedicum Helsinki, PB 63 (Haartmaninkatu 8), and Department of Medical Genetics, Haartman Institute, University of Helsinki, Helsinki 00014, Finland; [17]Department of Human Genetics, Yokohama City University, Graduate School of Medicine, 3-9 Fukuura, Kanazawa-ku, Yokohama 236-0004, Japan
*Correspondence: beggs@enders.tch.harvard.edu
http://dx.doi.org/10.1016/j.ajhg.2013.10.020.

Therefore, we performed whole-exome sequencing (WES) combined, when applicable, with autozygome analysis to identify mutations in novel genes that underlie the disease pathology in a cohort of individuals affected with NM with unknown genetic diagnosis. All subjects were enrolled following informed consent and research was conducted according to the protocols approved by the Institutional Review Boards of the respective institutions in which these individuals were recruited. Molecular screening was performed on genomic DNA isolated from blood samples following standard protocols.

We performed whole-exome or whole-genome sequencing on a cohort of 60 unrelated NM probands through Boston Children's Hospital Gene Partnership facility. Molecular screening was performed on genomic DNA isolated from blood samples with standard protocols. Whole-blood DNA was subjected to solution capture (SureSelect Human All Exon V4, Agilent Technologies) to generate barcoded whole-exome sequencing libraries. Libraries were sequenced on an Illumina HiSeq 2000, employing paired end reads (100 bp × 2) to a mean target coverage of 96.5% and a mean read depth of 71.6. Alignment, variant calling, and annotation were performed with a custom informatics pipeline employing BWA,[14] Picard, and ANNOVAR[15] focusing on rare (<3% in db SNP135, 1000 Genomes Project Database, and the [EVS] National Heart, Lung, and Blood Institute Exome Sequencing Project Exome Variant Server) protein affecting changes in known and novel human disease genes. Alternatively, probands for families 203 and 832 were sequenced to greater than 50× depth by Axeq Technologies on an Illumina HiSeq 2000 following Agilent SureSelect Exome enrichment with their standard Exome Sequencing service. Whole-exome sequencing identified homozygous mutations of *KLHL41* in two unrelated families, suggesting this gene to be a candidate for NM. All *KLHL41* mutations are numbered relative to the mRNA sequence NM_006063.2 (where position 1 is the first base of the initiating MET codon) and protein NP_006054.2. Family 1 is a nonconsanguineous family of Vietnamese origin. Proband 203-1 is a 16-year-old female with an intermediate form of NM with a high-arched palate, dysarthria, and scoliosis who has required ventilatory support since childhood. WES identified an apparently homozygous c.103T>C transition in exon 1 resulting in a p.Cys35Arg substitution in this individual (Figure 1A). This variant was present as heterozygous in the father and absent in the mother. Copy number analysis in the affected region showed a heterozygous deletion in the mother and the proband, c.(?_-77)_(*602_?)del. Therefore, individual 203-1 is compound heterozygous for a deletion involving a portion of *KLHL41* and a KLHL41 p.Cys35Arg missense change. The second proband (832-1), who is adopted of Russian origin, is ambulant at age 12 and exhibits the typical congenital form of NM. WES identified a homozygous deletion of one base and an insertion of four bases c.459delinsACTC in the proband resulting in a single amino acid insertion, p.Ser153_Ala154insLeu in the protein (Figure 1A).

Whole-exome sequencing in probands with severe NM in Australian and Saudi Arabian cohorts resulted in identification of *KLHL41* mutations in two further families. The first (6462) is a consanguineous family of Persian origin from Afghanistan with one child (D12-203) affected with severe NM and four unaffected children (see Figure S1 available online). Homozygosity mapping was performed on the proband with the Illumina HumanCytoSNP-12 array, and the only known NM loci found within homozygous regions were *CFL2* (MIM 601443) and *NEB* (MIM 161650); however, both were excluded following Sanger sequencing, as was *ACTA1* (MIM 102610), which is the most common cause of simplex NM cases. WES of DNA from proband D12-203 was performed at the Lotterywest Sate Biomedical Facility Genomics Node, Royal Perth Hospital, Western Australia.[13] WES identified 453 heterozygous or homozygous variants. Application of the homozygosity data to the list of candidates reduced this to seven candidate variants. Two of these seven candidate variants were in skeletal-muscle-specific genes and of these the most likely candidate was a homozygous deletion within *KLHL41* (chr2: 170382132–170382139; c.1748_1755delAAGGAAAT, p.Lys583Thrfs*7) (Figure 1A). The deletion was confirmed by Sanger sequencing. Both parents and two unaffected siblings were heterozygous for the deletion, and two further unaffected siblings were homozygous for the normal allele.

Family 12DG1177, from a Saudi Arabian cohort is consanguineous (Figure S1). The male proband (12DG1177-1) was a newborn with severe hypotonia, dislocation of hips and knees, and facial dysmorphism in the form of micrognathia and cleft palate. There was a positive family history of two previous sibs who died of unknown causes soon after birth, as well as three healthy living sibs. The proband died of cardiorespiratory arrest shortly after intubation at less than 24 hr of age. Exome capture was performed with TruSeq Exome Enrichment kit (Illumina) as described earlier.[16] Only novel coding and splicing homozygous variants within the autozygome of the affected individual were considered. After filtering, 8,653 homozygous, coding, or splice variants were present, and autozygosity mapping, dbSNP, and analysis of 240 control Saudi exomes finally led to the identification of 18 candidate variants. The only truncating change was a single base deletion in *KLHL41* (c.641delA). This deletion was present in the coding region of exon 1 of *KLHL41* resulting in the frameshift change p.Asn214Thrfs*14 (Figure 1A).

Subsequent screening for *KLHL41* mutations in 116 individuals affected with severe, intermediate, or typical congenital forms of NM in the Boston and Australian NM Cohorts by Sanger sequencing identified a further family (D10-236) with compound heterozygous mutation (c.581_583delAAG, p.Glu194del and c.1238C>T, p.Ser413Leu) in proband. This individual is of Chinese

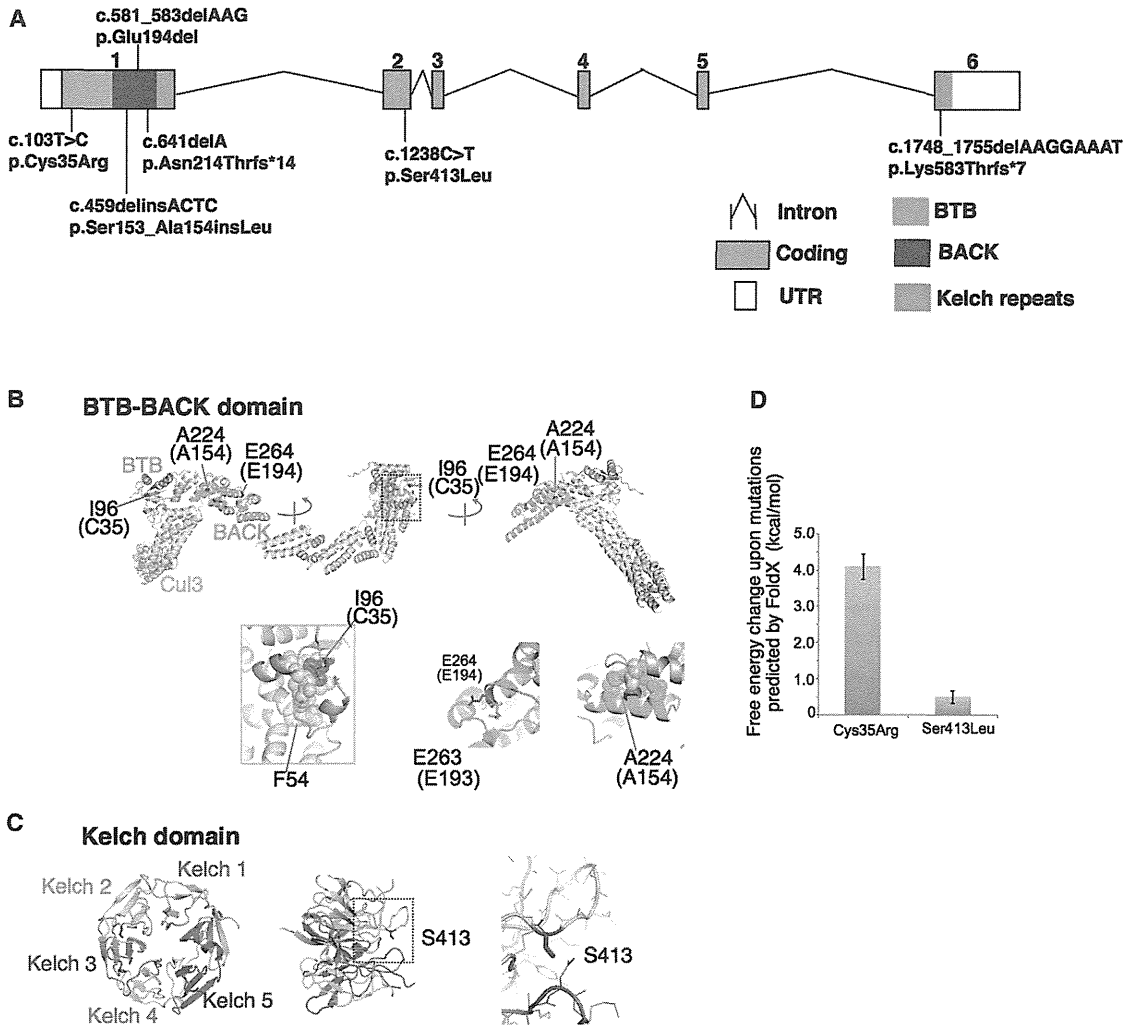

Figure 1. Overview of Mutations in *KLHL41* and Their Effect on Protein Structure

(A) Schematic representation of mutations in *KLHL41*. Boxes represent exons 1–6. Conserved domains of KLHL41 are indicated as follows: BTB (blue), BACK (red), and Kelch repeats (green). The BTB and BACK domains are encoded by exon 1 and the five Kelch repeats are encoded by exons 1–6.

(B and C) Crystal structures of the BTB-BACK domain of human Kelch-like protein (KLHL11) in complex with CUL3 (Protein Data Bank code 4AP2) (B) and the Kelch domain of rat KLHL41 (PDB code 2WOZ) (C). α helices, β strands, and loops are drawn as ribbons, arrows, and threads, respectively. The squared areas correspond to the close-up views in the insets. In (B), the BTB and BACK domains are colored pink and green, respectively, whereas CUL3 is colored yellow, except that Ile96, Ala224, and Glu264 (Cys35, Ala154, and Glu194 in human KLHL41, respectively) are colored red. The side chains of these residues and Glu263 (Glu193 in human KLHL41) are shown as sticks with the indications of amino acid numbers for human KLHL11 and those for human KLHL41 in parentheses. Side chains involved in hydrophobic cores around Ile96 and Ala224 are drawn in van der Waal's representation. In (C), the Kelch domain is color-coded to indicate each Kelch repeat, except that Ser413 is colored red. The side chain of Ser413 is shown as sticks. Molecular structures are drawn with PyMOL.

(D) Predicted free energy changes upon the substitutions of KLHL41 with FoldX software.

origin and exhibited the typical congenital form of NM. The detailed clinical features of affected individuals with mutations identified in *KLHL41* are presented in Table 1.

Overall, WES and Sanger sequencing resulted in identification of seven different mutations in Kelch-like family member 41 (*KLHL41*), previously known as *KBTBD10*, sarcosin, or *KRP1*, in affected NM individuals from five unrelated families (Figure 1A). Muscle histology was typical for NM: biopsies from probands of three different families (D12-203, 832-1, and 10-236) exhibited abnormal Gomori trichrome staining with presence of sarcoplasmic rods that varied from numerous small rods to fewer large rods in multiple myofibers (Figure 2A). No intranuclear rods or cores were seen. The missense changes identified in *KLHL41* are predicted to be pathogenic by polyphen, SIFT and pMUT and the mutated amino-acid residues are conserved in all representative species during evolution (Figure S2). The neighboring areas surrounding the sites of insertion or deletion are also relatively conserved, suggesting a structural or functional requirement for the altered amino acid residues (Figure S2). Sequencing of family members revealed that *KLHL41*

Table 1. Clinical Manifestations in Affected Individuals Harboring *KLHL41* Mutations

Proband ID	cDNA Change	Amino Acid Change	Clinical Category	Sex	Nationality	Pregnancy and Delivery	Alive at Age/Mobility/ Age at Death	Associated Features
203-1	c.103T>C c.(?_-77)_(*602_?)del	p.Cys35Arg Heterozygous p.0? Heterozygous	Intermediate	F	Vietnamese	Normal	16 yrs, uses wheelchair (ambulant 24–36 mo)	Ventilated 24 hr from 5 yrs. High-arched palate, dysarthria Scoliosis
832-1	c.459delinsACTC	p.Ser153_Ala154insLeu Homozygous	Other forms (grade of severity: mild)	M	Russian	No data	12 yrs, ambulant	Distal weakness > proximal distal contractures
D10-236	c.581_583delAAG c.1238C>T	p.Glu194del Heterozygous p.Ser413Leu Heterozygous	Typical form	M	Chinese	Normal - h 40	5 yrs, ambulant	VSD, finger contractures, focal renal echogenicity
D12-203	c.1748_1755del AAGGAAAT,	p.Lys583Thrfs*7 Homozygous	Fetal akinesia sequence	M	Persian	Polyhydramnios, breech presentation, emergency Cesarean section - h 31+2	Died at 3 mo (active support discontinued)	Arthrogryposis, macrocephaly, hypospadias No antigravity movements at birth
12DG1177	c.641delA	p.Asn214Thrfs*14 Homozygous	Severe form Fetal akinesia sequence	M	Saudi Arabian	Fetal movements weak, breech presentation	Died during 1st day of life	Dislocation of hips and knees, cleft palate, micrognathia, narrow chest

mutations showed a segregation pattern compatible with a recessive mode of inheritance in all families (Figure S1). Severe phenotypes associated with genetic null mutations and intermediate or typical congenital forms with mutations that should results in presence of residual protein, suggests a phenotype-genotype correlation in individuals affected with *KLHL41* mutations.

KLHL41 belongs to the family of BTB-Kelch domain-containing proteins.[17–20] Mutations in two other members of this family, *KBTBD13* (MIM 613727), and most recently *KLHL40* (MIM 615430), have been associated with a clinically distinct form of congenital myopathy exhibiting nemaline bodies, as well as multiminicores and severe NM, respectively.[12,13] To evaluate the impacts of the *KLHL41* mutations on the protein structure, we mapped them onto the crystal structures of the BTB-BACK domain of human KLHL11 in complex with human CUL3, a subunit of E3 ubiquitin ligases, (PDB code 4AP2)[21] and the Kelch domain of rat KLHL41 (PDB code 2WOZ),[22] analogous to those domains of human KLHL41. The Cys35 side chain is involved in a hydrophobic core of the BTB domain, which makes van der Waals contacts with Phe54 of Cul3 (Figure 1B). The p.Cys35Arg substitution present in affected individual 203-1 would likely destabilize the hydrophobic core and thereby impair the interaction with Cul3. This was supported by the FoldX result, in which free energy change upon the p.Cys35Arg substitution was predicted to be over 4 kcal/mol, which can be interpreted as considerable destabilization of a protein structure (Figure 1D; Figure S3).[23] In proband 832-1, a Leu residue is inserted between the amino acid positions 153 and 154 in the center of a helix, in which several residues are involved in a hydrophobic core of the BACK domain (Figure 1B). This amino acid insertion is likely to destabilize the BACK domain fold. In proband D10-236, the p.Ser413Leu substitution was mapped to a loop region, which is located near the substrate-binding region of the Kelch repeat 2 (Figure 1C; Figure S1B). A FoldX calculation predicted that the p.Ser413Leu substitution would have minimal effect on stability of the Kelch domain (Figure 1D). The effect of Glu194 deletion at the N-terminal end of an α helix can be compensated by the presence of Glu193 located in the loop (Figure 1B). Nonetheless, it cannot be excluded that the p.Ser413Leu and p.Glu194del changes alter the protein solubility or aggregate tendency and/or impair substrate binding. The conserved nature of the mutated KLHL41 domains, as well as the potential role of the mutations in disrupting those structural domains, supports the likely pathogenicity of these mutations.

The localization of KLHL41 in skeletal muscles was investigated by immunofluorescence of mouse FDB cultured myofibers and human skeletal muscle cryosections. Immunofluorescence with two different antibodies against N-terminal (Sigma, AV38732) and C-terminal parts of human KLHL41 (Abcam, ab66605) was performed, and z stacks were acquired by confocal microscopy as described

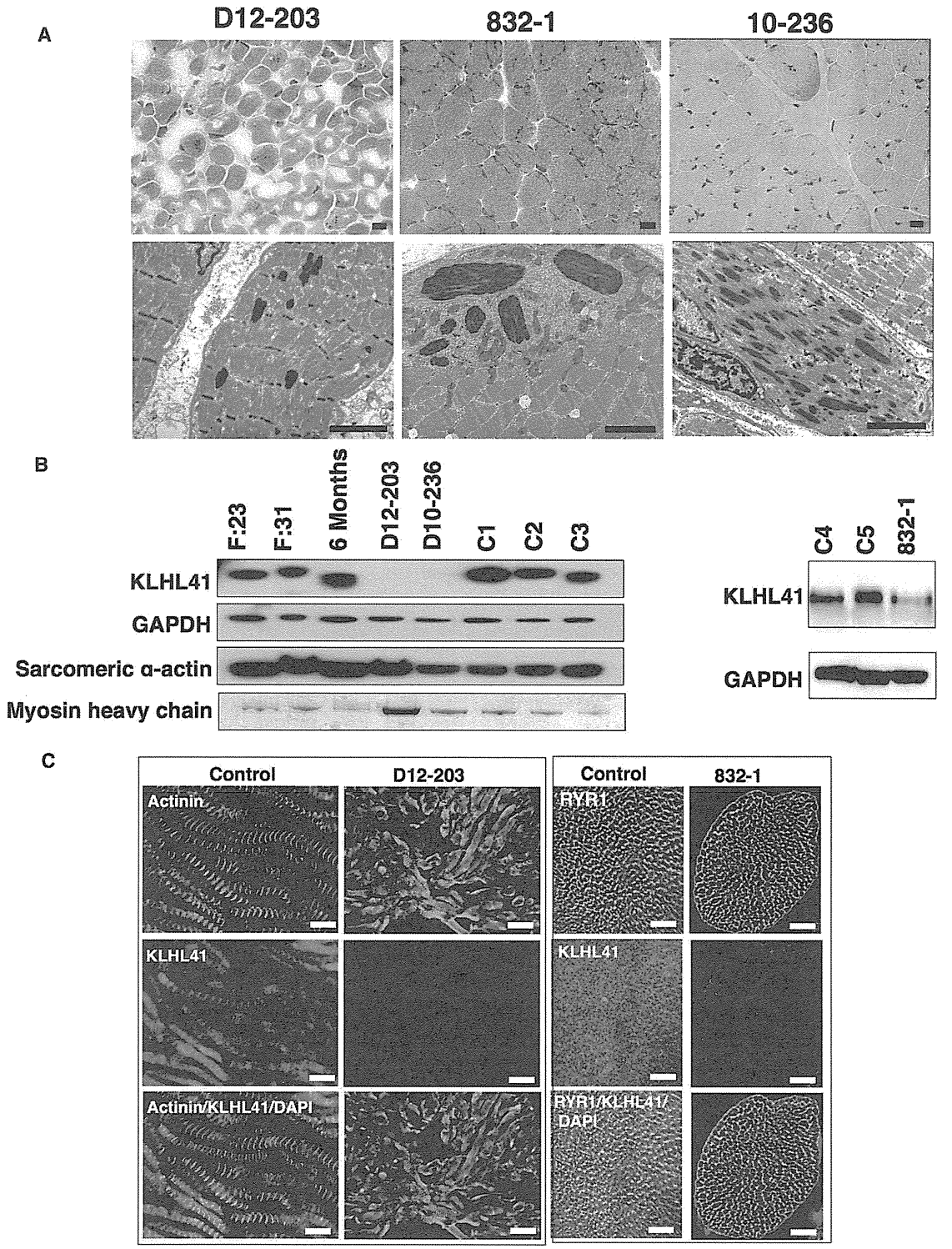

Figure 2. Muscle Pathology and Expression of KLHL41 Levels and Localization in Muscle of Affected Individuals
(A) Light microscopy of Gomori trichrome stained skeletal muscle from affected individuals with *KLHL41* mutations show cytoplasmic nemaline bodies (top panel). Electron microscopy of affected muscles reveals rods of variable frequency and size and severe myofibrillar disarray (bottom panels). (Scale bars represent 2 μm). Affected individuals' IDs are indicated at top.
(B) Immunoblotting analysis of KLHL41 levels in affected and unaffected muscles. A decrease in protein levels was observed in individuals with *KLHL41* mutations in comparison to normal control muscles. Immunoblotting with sarcomeric actin or Coomassie staining of myosin heavy chain showed no abnormal accumulation of sarcomeric proteins in affected muscles. Immunostaining for GAPDH was used for loading controls. Lanes: F:23, 23 week control fetus; F:30, 31 week control fetus; 6-month-old control baby, C1–C5 are normal age-matched control muscles.
(C) Immunofluorescence for KLHL41 in control and affected individual muscle biopsies showed highly reduced levels of KLHL41 in longitudinally oriented (left) or transverse sections (right) of skeletal muscles from affected individuals. Scale bars represent 50 μm.

previously.[24] Immunofluorescence with both antibodies resulted in similar staining patterns; however, due to lower background staining, the C-terminal antibody was used for further studies. Costaining with sarcomeric markers in longitudinal planes showed that KLHL41 staining predominated over the I-bands of the sarcomere and at perinuclear regions in human biopsies (Figure 2C) and murine cultured myofibers (Figure S4). Analysis of transverse sections of myofibers from control human biopsies revealed KLHL41 staining in a ring pattern around the myofibrils, generally colocalizing with ryanodine receptors (RYR1), which are a marker of the sarcoplasmic reticulum (Figure S5). Together, these observations suggest that KLHL41 localizes over (but not within) I bands, likely in association with the terminal cisternae of the sarcomplasmic reticulum (SR) and longitudinal vesicles of the SR present in the I-band area at the triadic regions (Figure S4). Colocalization studies with the ER marker protein disulfide isomerase (PDI) in myofibers and skeletal muscles further confirmed the localization of KLHL41 in SR-ER membranes (Figures S4). This overall localization pattern is most consistent with localization to the endoplasmic reticulum (ER) around myonuclei and to microdomains of the SR with ER characteristics.[25] Previous studies suggested that the closely related NM protein, KLHL40, localized at A-bands,[13] but double label immunofluorescence studies of both longitudinal and transverse sections here reveal that it appears colocalized with RYR1, around but not within the myofibrils in cultured myofibers and human skeletal muscles in a pattern overlapping, but not identical to, that of KLHL41 (Figures S4 and S5). These associations of proteins whose defects cause NM with the ER/SR contrasts with previously known NM proteins, all of which are sarcomeric thin filament components, with the exception of KBTBD13 whose localization is not well known.

In mouse tissues, immunoblotting detected KLHL41 in skeletal muscle and diaphragm (Figure S6). In cultured murine C2C12 cells, KLHL41 levels increased during differentiation to myotubes (Figure S6). Immunoblotting of affected skeletal muscle extracts revealed greatly reduced levels of KLHL41 in individuals with *KLHL41* mutations (Figure 2B) and immunofluorescence microscopy of affected individuals' skeletal muscles also showed that KLHL41 levels were greatly reduced in their myofibers (Figure 2C).

Cell culture studies have shown that KLHL41 interacts with nebulin, N-RAP (Nebulin-related anchoring protein), and actin in skeletal muscle and promotes the assembly of myofibrils.[26] KLHL41 regulates skeletal muscle differentiation as overexpression or knockdown inhibited C2C12 myoblast differentiation.[27] Knockdown of *Klhl41* in cultured cardiomyocytes resulted in sarcomeric disorganization with thickening of Z-lines as seen in NM.[28] However, the exact functions of KLHL41 in disease pathology are unknown. Recent studies have identified mutations in two other closely related family members *KBTBD13* and *KLHL40* as causes of NM suggesting the crucial requirement for several Kelch family proteins in skeletal muscle function.[12,13] To investigate the functional role of KLHL41 in vertebrate skeletal muscle development, we employed zebrafish as a model system. Zebrafish have two duplicated orthologs (*klhl41a* and *klhl41b*) that share ~80% similarity with *KLHL41*. Zebrafish whole-mount in situ hybridization was performed to study the spatio-temporal expression of these genes during zebrafish development as described previously.[29] Specifically, RNA probes specific for each Klhl41 gene were generated by amplification of the 3′ UTRs from a cDNA library of 2 day postfertilization (dpf) zebrafish embryos, followed by in vitro transcription to generate digoxigenin-labeled antisense transcripts (primer sequences are provided in Table S1). Whole-mount in situ hybridization showed ubiquitous expression of *klhl41a* during early development at 1 dpf, but by 2 dpf, *klhl41a* transcripts were virtually undetectable in the major axial skeletal muscles. In contract, *klhl41b* expression was predominantly seen in striated muscles, and strong expression in heart and skeletal muscles was observed throughout zebrafish development to at least 5 dpf (Figure 3A).

The effect of KLHL41 deficiency in zebrafish was studied by knocking down the Klhl41 genes with antisense morpholinos. Two independent morpholinos targeting an exon-intron spice site and translational start site were designed for both genes (morpholino sequences are provided in Table S2). As initial experiments with both morpholinos for each transcript resulted in similar phenotypes, we performed the remainder of our studies with the splice-site morpholinos (7 ng). *klhl41a* morphants exhibited leaner bodies, smaller eyes, and pericardial edema as seen in other myopathy models (n = 65–110) (Figure 3B).[30,31] Examination of 3 dpf morphants with polarized light showed reduced birefringence in axial skeletal muscles suggesting disorganized skeletal muscle structure (Figure 3B; Figure S7). Knockdown of *klhl41b* resulted in reduced birefringence without any other significant abnormalities (n = 82–132). Targeting both *klhl41a* and *klhl41b* (7 ng each) resulted in curved bodies with a 30% reduction in size along with small eyes and pericardial edema (n = 89–103), compared to fish injected with control morpholino (14ng). *klhl41a* morphant fish die by 3 dpf while *klhl41b* morphants typically did not survive past 5 dpf. Knockdown of both genes was lethal by 3 dpf. Double knockdown fish exhibited severely disorganized muscle (measured by reduced birefringence) compared to controls and either of the single knockdowns. RT-PCR and immunoblotting confirmed the knockdown of *klhl41a* and *klhl41b* transcripts and a reduction in protein levels (Figures 3C and 3D). Overexpression of human *KLHL41* mRNA in the double morphants resulted in a significant increase in the number of surviving fish with normal birefringence suggesting the specificity of morpholino injections and demonstrating the ability of this single evolutionary ortholog to complement both zebrafish genes (Figure 3E). Behavioral characterization of 3

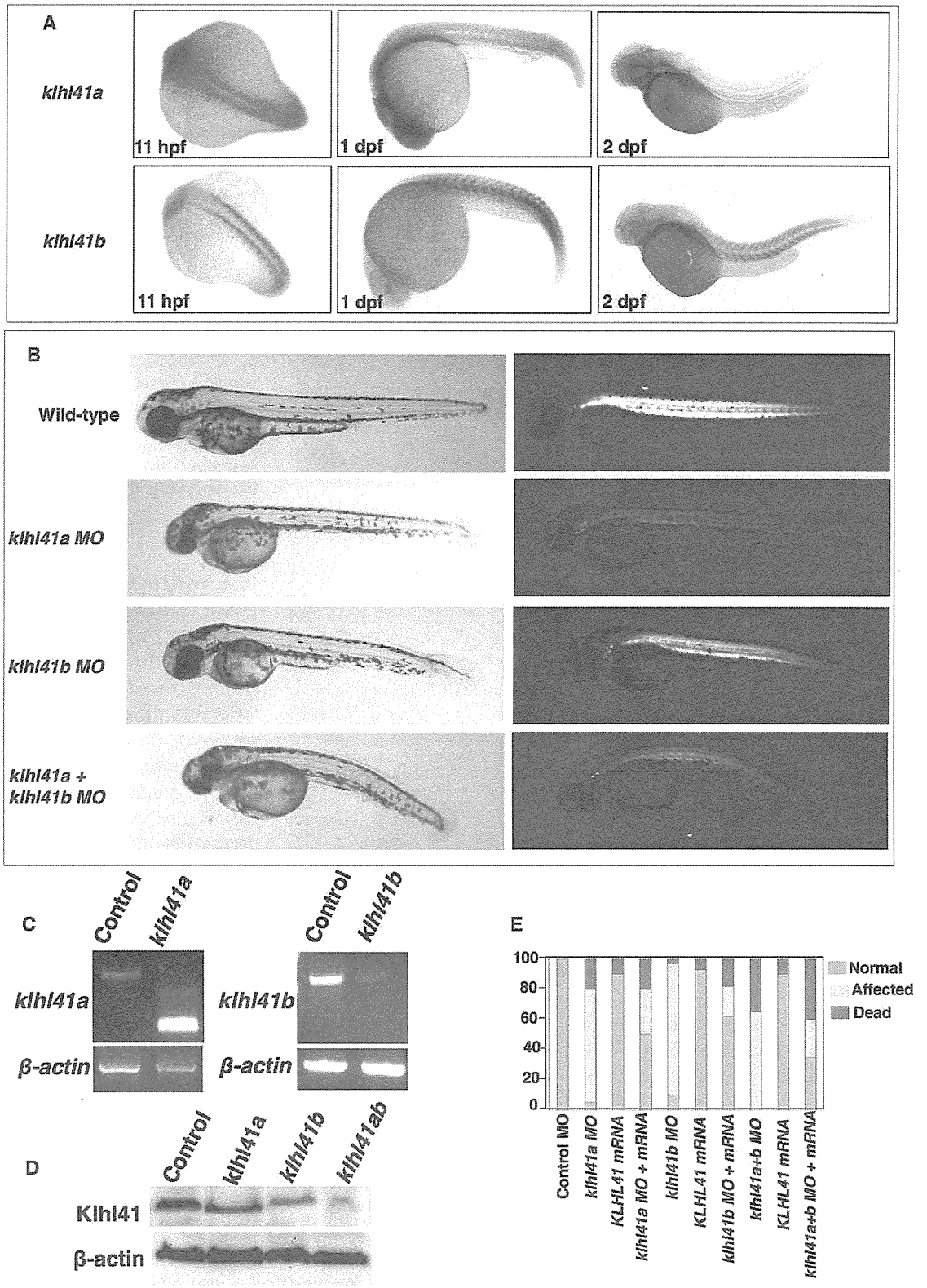

Figure 3. Characterization and Knockdown of Zebrafish Orthologs of *KLHL41*

(A) In situ hybridization of the zebrafish Klhl41 genes shows early expression during myogenesis in developing somites (11 hr after fertilization). *Klhl41a* is expressed in brain, eyes, and muscle at 1 dpf. Later in development expression is largely restricted to brain and heart (2 dpf), although low levels of expression in axial slow skeletal myofibers cannot be excluded due to limited sensitivity of the assay. *Klhl41b* expression is localized to skeletal muscle and heart at all developmental stages (1–2 dpf).

(B) Knockdown of Klhl41 genes in zebrafish using antisense morpholinos results in myopathic changes. Live microscopy of zebrafish embryos at 3 dpf reveals leaner and smaller bodies in comparison to wild-type (WT) fish. Under polarized microscopy, zebrafish embryos

(legend continued on next page)

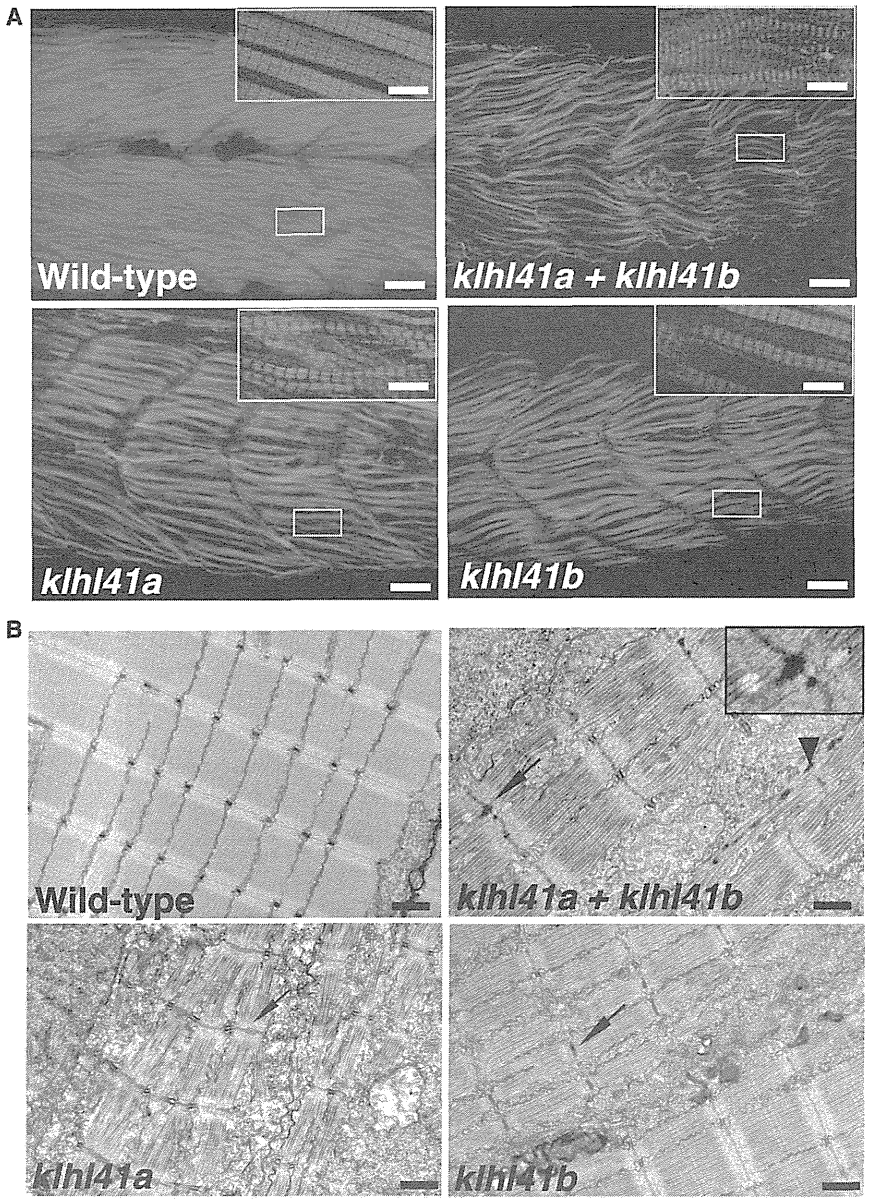

Figure 4. Loss of *klhl41* Function in Zebrafish Recapitulates the Disease Pathology of Human Nemaline Myopathies
(A) Whole-mount staining of 3 dpf zebrafish embryos with phalloidin showed extensive myofibrillar disarray of myofibers in *klhl41* morphant fish (scale bar represents 2 μm). Three dpf embryos fixed in 4% paraformaldehyde were incubated with phalloidin (Invitrogen, A12380, 1:40) overnight at 4°C. Skeletal muscles of klhl41-deficient embryos were smaller and exhibited an overall reduction of myofibrillar organization (inset, high magnification).
(B) Electron microscopy of klhl41-deficient skeletal muscle revealed thickened Z-lines in *klhl41a* or *klhl41b* morphants. In addition, skeletal muscle of double knockdown fish contained electron dense bodies, reminiscent of nascent nemaline rods (arrowhead, nemaline bodies like structures; arrow, thickened Z-lines) (scale bar represents 1 μm).

dpf morphant fish, knocked down for either or both Klhl41 genes, using the touch-evoked response assay showed significantly diminished motility in comparison to control fish (WT fish: 5.74 ± 0.98 cm/0.1 s; *klhl41a*: 1.32 ± 0.61 cm/0.1 s; *klhl41b*: 2.00 ± 0.49 cm/0.1 s; *klhl41ab*: 0.73 ± 0.39 cm/0.1 s), suggesting a significant degree of overall muscle weakness (Movies S1, S2, S3, and S4).[32] To visualize abnormalities in sarcomeric architecture, whole-mount staining of morphant fish and control zebrafish embryos was performed with phalloidin to stain the actin-thin filaments. Although well-organized myofibrillar striations (i.e., sarcomeres) were observed, the myofibrils in *klhl41* morphants tended to be thinner and were highly disorganized relative to control fish (Figure 4A). The myofibrilar disorganization in *klhl41* morphants was also evident by evaluation of ultrathin toluidine blue sections of control and morphant fish (Figure S7). The main diagnostic feature of NM is the presence of nemaline rods with or without Z-line streaming in skeletal muscle. Ultrastructural examination of zebrafish skeletal muscle by electron microscopy showed Z-line thickening in both *klhl41a* and *klhl41b* morphant fish (Figure 4B). Knockdown of both *klhla* and *klhl41b* resulted in the presence of numerous electron-dense structures, reminiscent of small or nascent nemaline bodies, in addition to Z-line thickening (Figure 4B). Given the differences in temporal expression of *klhl41a* (early embryogenesis) and *klhl41b*

exhibit a reduction in birefringence in morphant fish, quantified in ImageJ as described (WT controls: 100% ± 5.9% *klhl41a*: 23% ± 3.0%; *klhl41b*: 31% ± 8.2%; *klhl41ab*: 16% ± 4.2%). Double knockdown fish show a more severe skeletal muscle phenotype than single morphants.
(C) RT-PCR analysis showed knockdown of normal transcripts in the morphant fish.
(D) Immunoblot analysis showed reduction in Klhl41 levels in *klhl41a*, *klhl41b*, and *klhl41ab* fish. Klhl41 antibody recognizes both klhl41a and klhl41b and therefore show immunoreactibility to the other gene in the single morphants that is highly reduced in double morphants.
(E) Overexpression of human *KLHL41* mRNA restores the skeletal muscle phenotypes of *klhl41a/b* single and double morphants suggesting morpholino specificity. The mRNA concentration used to rescue were as follows: *klhl41a* (50 pg), *klhl41b* (75 pg), *klhl41a+b* (60 pg of each).

(maintained later in development), and the high degree of structural and functional conservation (both are rescued by the single human transcript), it is likely that increased severity of *klhl41a* morphants is due to this being the predominant embryonic isoform at the early stages targeted by morpholino injections.

Extensive skeletal muscle disorganization associated with sarcomeric abnormalities in morphant fish points toward a function of KLHL41 in skeletal muscle development and maintenance. Mutations affecting the closely related BTB-Kelch family member KLHL40 have recently also been reported to cause nemaline myopathy.[13] While *KLHL40* mutations resulted in a severe clinical presentation in most of the affected individuals, KLHL41 abnormalities are associated with a spectrum of phenotypes from severe with neonatal death, to survival into late childhood. However, no significant differences were seen in skeletal muscle pathology. KLHL40 contains a putative nuclear localization sequence (NLS) and is expressed throughout muscle differentiation, whereas KLHL41 lacks NLS and is expressed in late differentiation (Figure S8).[13] KLHL41 and many other BTB domain-containing Kelch family members are known to interact with Cul3 ubiquitin ligase to form functional ubiquitination complexes with proteins targeted for degradation.[21,33] KLHL41, which has been shown to interact with nebulin,[34] is now the third BTB-Kelch family member to be identified as a cause of NM when mutated. We hypothesize that improper surveillance and degradation of aberrant thin-filament proteins might explain the convergent pathological and clinical phenotypes associated with mutations of thin filament and BTB-Kelch family member genes in NM.

Supplemental Data

Supplemental Data include eight figures, two tables, and four movies and can be found with this article online at http://www.cell.com/AJHG/home.

Acknowledgments

We are grateful to the many NM affected individuals and their families, and to their treating physicians, for their participation in this research. Whole-exome sequencing was made possible through the generous support and assistance of David Margulies and the entire staff of The Gene Partnership Project at Boston Children's Hospital. We would like to thank Pankaj Agrawal and Wen-Hann Tan for many helpful discussions during the course of this work. We are thankful to Louise Trakimas of the electron microscope facility at Harvard Medical School for excellent help with zebrafish histology, and the Genotyping and Sequencing Core Facilities at KFSHRC for their technical help. V.A.G. is supported by K01 AR062601 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases of National Institute of Health. This work was also supported by the Muscular Dystrophy Association of USA (MDA201302), National Institutes of Health grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases R01 AR044345, the AUism Charitable Foundation, and A Foundation Building Strength (to A.H.B.); National Health and Medical Research Council of Australia Early Career Researcher Fellowship #1035955 (to G.R.); Research Fellowship APP1002147 and Project Grant APP1022707(to N.G.L.); the Association Française contre les Myopathies (#15734), Dubai-Harvard Foundation for Medical Research Collaborative Research Grant (to F.S.A.); a UWA Collaborative Research Award (G.R.); and the Great Ormond Street Hospital Children's Charity (to F.M.). E.J.T. and K.S.Y. are supported by University of Western Australia Postgraduate Awards. DNA sequencing was performed by the Boston Children's Hospital Genomics Program Molecular Genetics Core, and confocal microscopy was performed at Boston Children's Hospital Intellectual and Developmental Disability Research Center Imaging Core, both supported by National Institutes of Health grant P30 HD18655. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Received: August 9, 2013
Revised: October 15, 2013
Accepted: October 22, 2013
Published: November 21, 2013

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://browser.1000genomes.org
dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/
Picard, http://picard.sourceforge.net/
Pymol, http://www.pymol.org

References

1. Wallgren-Pettersson, C., Sewry, C.A., Nowak, K.J., and Laing, N.G. (2011). Nemaline myopathies. Semin. Pediatr. Neurol. *18*, 230–238.
2. Ryan, M.M., Schnell, C., Strickland, C.D., Shield, L.K., Morgan, G., Iannaccone, S.T., Laing, N.G., Beggs, A.H., and North, K.N. (2001). Nemaline myopathy: a clinical study of 143 cases. Ann. Neurol. *50*, 312–320.
3. Wallgren-Pettersson, C. (2002). Nemaline and myotubular myopathies. Semin. Pediatr. Neurol. *9*, 132–144.
4. Sewry, C.A. (2008). Pathological defects in congenital myopathies. J. Muscle Res. Cell Motil. *29*, 231–238.
5. Hutchinson, D.O., Charlton, A., Laing, N.G., Ilkovski, B., and North, K.N. (2006). Autosomal dominant nemaline myopathy with intranuclear rods due to mutation of the skeletal muscle ACTA1 gene: clinical and pathological variability within a kindred. Neuromuscul. Disord. *16*, 113–121.
6. Pelin, K., Hilpelä, P., Donner, K., Sewry, C., Akkari, P.A., Wilton, S.D., Wattanasirichaigoon, D., Bang, M.L., Centner, T., Hanefeld, F., et al. (1999). Mutations in the nebulin gene associated with autosomal recessive nemaline myopathy. Proc. Natl. Acad. Sci. USA *96*, 2305–2310.
7. Nowak, K.J., Wattanasirichaigoon, D., Goebel, H.H., Wilce, M., Pelin, K., Donner, K., Jacob, R.L., Hübner, C., Oexle, K., Anderson, J.R., et al. (1999). Mutations in the skeletal muscle

alpha-actin gene in patients with actin myopathy and nemaline myopathy. Nat. Genet. *23*, 208–212.
8. Laing, N.G., Wilton, S.D., Akkari, P.A., Dorosz, S., Boundy, K., Kneebone, C., Blumbergs, P., White, S., Watkins, H., Love, D.R., et al. (1995). A mutation in the alpha tropomyosin gene TPM3 associated with autosomal dominant nemaline myopathy. Nat. Genet. *9*, 75–79.
9. Tajsharghi, H., Ohlsson, M., Lindberg, C., and Oldfors, A. (2007). Congenital myopathy with nemaline rods and cap structures caused by a mutation in the beta-tropomyosin gene (TPM2). Arch. Neurol. *64*, 1334–1338.
10. Agrawal, P.B., Greenleaf, R.S., Tomczak, K.K., Lehtokari, V.L., Wallgren-Pettersson, C., Wallefeld, W., Laing, N.G., Darras, B.T., Maciver, S.K., Dormitzer, P.R., and Beggs, A.H. (2007). Nemaline myopathy with minicores caused by mutation of the CFL2 gene encoding the skeletal muscle actin-binding protein, cofilin-2. Am. J. Hum. Genet. *80*, 162–167.
11. Johnston, J.J., Kelley, R.I., Crawford, T.O., Morton, D.H., Agarwala, R., Koch, T., Schäffer, A.A., Francomano, C.A., and Biesecker, L.G. (2000). A novel nemaline myopathy in the Amish caused by a mutation in troponin T1. Am. J. Hum. Genet. *67*, 814–821.
12. Sambuughin, N., Yau, K.S., Olivé, M., Duff, R.M., Bayarsaikhan, M., Lu, S., Gonzalez-Mera, L., Sivadorai, P., Nowak, K.J., Ravenscroft, G., et al. (2010). Dominant mutations in KBTBD13, a member of the BTB/Kelch family, cause nemaline myopathy with cores. Am. J. Hum. Genet. *87*, 842–847.
13. Ravenscroft, G., Miyatake, S., Lehtokari, V.L., Todd, E.J., Vornanen, P., Yau, K.S., Hayashi, Y.K., Miyake, N., Tsurusaki, Y., Doi, H., et al. (2013). Mutations in KLHL40 are a frequent cause of severe autosomal-recessive nemaline myopathy. Am. J. Hum. Genet. *93*, 6–18.
14. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
15. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. *38*, e164.
16. Alkuraya, F.S. (2012). Discovery of rare homozygous mutations from studies of consanguineous pedigrees. Curr Protoc Hum Genet. *Chapter 6*, Unit 6, 12.
17. Adams, J., Kelso, R., and Cooley, L. (2000). The kelch repeat superfamily of proteins: propellers of cell function. Trends Cell Biol. *10*, 17–24.
18. Dhanoa, B.S., Cogliati, T., Satish, A.G., Bruford, E.A., and Friedman, J.S. (2013). Update on the Kelch-like (KLHL) gene family. Hum. Genomics *7*, 13.
19. du Puy, L., Beqqali, A., van Tol, H.T., Monshouwer-Kloots, J., Passier, R., Haagsman, H.P., and Roelen, B.A. (2012). Sarcosin (Krp1) in skeletal muscle differentiation: gene expression profiling and knockdown experiments. Int. J. Dev. Biol. *56*, 301–309.
20. Gray, C.H., McGarry, L.C., Spence, H.J., Riboldi-Tunnicliffe, A., and Ozanne, B.W. (2009). Novel beta-propeller of the BTB-Kelch protein Krp1 provides a binding site for Lasp-1 that is necessary for pseudopodial extension. J. Biol. Chem. *284*, 30498–30507.
21. Canning, P., Cooper, C.D., Krojer, T., Murray, J.W., Pike, A.C., Chaikuad, A., Keates, T., Thangaratnarajah, C., Hojzan, V., Marsden, B.D., et al. (2013). Structural basis for Cul3 protein assembly with the BTB-Kelch family of E3 ubiquitin ligases. J. Biol. Chem. *288*, 7803–7814.
22. Spence, H.J., Johnston, I., Ewart, K., Buchanan, S.J., Fitzgerald, U., and Ozanne, B.W. (2000). Krp1, a novel kelch related protein that is involved in pseudopod elongation in transformed cells. Oncogene *19*, 1266–1276.
23. Guerois, R., Nielsen, J.E., and Serrano, L. (2002). Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J. Mol. Biol. *320*, 369–387.
24. Lawlor, M.W., Alexander, M.S., Viola, M.G., Meng, H., Joubert, R., Gupta, V., Motohashi, N., Manfready, R.A., Hsu, C.P., Huang, P., et al. (2012). Myotubularin-deficient myoblasts display increased apoptosis, delayed proliferation, and poor cell engraftment. Am. J. Pathol. *181*, 961–968.
25. Kaisto, T., and Metsikkö, K. (2003). Distribution of the endoplasmic reticulum and its relationship with the sarcoplasmic reticulum in skeletal myofibers. Exp. Cell Res. *289*, 47–57.
26. Lu, S., Carroll, S.L., Herrera, A.H., Ozanne, B., and Horowits, R. (2003). New N-RAP-binding partners alpha-actinin, filamin and Krp1 detected by yeast two-hybrid screening: implications for myofibril assembly. J. Cell Sci. *116*, 2169–2178.
27. Paxton, C.W., Cosgrove, R.A., Drozd, A.C., Wiggins, E.L., Woodhouse, S., Watson, R.A., Spence, H.J., Ozanne, B.W., and Pell, J.M. (2011). BTB-Kelch protein Krp1 regulates proliferation and differentiation of myoblasts. Am. J. Physiol. Cell Physiol. *300*, C1345–C1355.
28. Greenberg, C.C., Connelly, P.S., Daniels, M.P., and Horowits, R. (2008). Krp1 (Sarcosin) promotes lateral fusion of myofibril assembly intermediates in cultured mouse cardiomyocytes. Exp. Cell Res. *314*, 1177–1191.
29. Gupta, V., Discenza, M., Guyon, J.R., Kunkel, L.M., and Beggs, A.H. (2012). α-Actinin-2 deficiency results in sarcomeric defects in zebrafish that cannot be rescued by α-actinin-3 revealing functional differences between sarcomeric isoforms. FASEB J. *26*, 1892–1908.
30. Dowling, J.J., Vreede, A.P., Low, S.E., Gibbs, E.M., Kuwada, J.Y., Bonnemann, C.G., and Feldman, E.L. (2009). Loss of myotubularin function results in T-tubule disorganization in zebrafish and human myotubular myopathy. PLoS Genet. *5*, e1000372.
31. Gupta, V.A., Kawahara, G., Myers, J.A., Chen, A.T., Hall, T.E., Manzini, M.C., Currie, P.D., Zhou, Y., Zon, L.I., Kunkel, L.M., and Beggs, A.H. (2012). A splice site mutation in laminin-α2 results in a severe muscular dystrophy and growth abnormalities in zebrafish. PLoS ONE *7*, e43794.
32. Smith, L.S., Beggs, A.H., and Gupta, V.A. (2013). Analysis of skeletal muscle defects in larval zebrafish by birefringence and touch-evoke escape response assays. J. Vis. Exp. *82*, e50925. http://dx.doi.org/10.3791/50925.
33. Zhang, D.D., Lo, S.C., Sun, Z., Habib, G.M., Lieberman, M.W., and Hannink, M. (2005). Ubiquitination of Keap1, a BTB-Kelch substrate adaptor protein for Cul3, targets Keap1 for degradation by a proteasome-independent pathway. J. Biol. Chem. *280*, 30091–30099.
34. Spence, H.J., McGarry, L., Chew, C.S., Carragher, N.O., Scott-Carragher, L.A., Yuan, Z., Croft, D.R., Olson, M.F., Frame, M., and Ozanne, B.W. (2006). AP-1 differentially expressed proteins Krp1 and fibronectin cooperatively enhance Rho-ROCK-independent mesenchymal invasion by altering the function, localization, and activity of nondifferentially expressed proteins. Mol. Cell. Biol. *26*, 1480–1495.

Whole genome sequencing in patients with retinitis pigmentosa reveals pathogenic DNA structural changes and *NEK2* as a new disease gene

Koji M. Nishiguchi[a,b], Richard G. Tearle[c], Yangfan P. Liu[d], Edwin C. Oh[d,e], Noriko Miyake[f], Paola Benaglio[a], Shyana Harper[g], Hanna Koskiniemi-Kuendig[a], Giulia Venturini[a], Dror Sharon[h], Robert K. Koenekoop[i], Makoto Nakamura[b], Mineo Kondo[b], Shinji Ueno[b], Tetsuhiro R. Yasuma[b], Jacques S. Beckmann[a,j,k], Shiro Ikegawa[l], Naomichi Matsumoto[f], Hiroko Terasaki[b], Eliot L. Berson[g], Nicholas Katsanis[d], and Carlo Rivolta[a,1]

[a]Department of Medical Genetics, University of Lausanne, 1005 Lausanne, Switzerland; [b]Department of Ophthalmology, Nagoya University School of Medicine, Nagoya 466-8550, Japan; [c]Complete Genomics, Inc., Mountain View, CA 94043; [d]Center for Human Disease Modeling and [e]Department of Neurology, Duke University, Durham, NC 27710; [f]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama 236-0004, Japan; [g]Berman-Gund Laboratory for the Study of Retinal Degenerations, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, MA 02114; [h]Department of Ophthalmology, Hadassah-Hebrew University Medical Center, Jerusalem 91120, Israel; [i]McGill Ocular Genetics Laboratory, McGill University Health Centre, Montreal, QC, Canada H3H 1P3; [j]Service of Medical Genetics, Lausanne University Hospital, 1011 Lausanne, Switzerland; [k]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland; and [l]Laboratory for Bone and Joint Diseases, Center for Genomic Medicine, RIKEN, Tokyo 108-8639, Japan

Edited by Jeremy Nathans, Johns Hopkins University, Baltimore, MD, and approved August 15, 2013 (received for review May 1, 2013)

We performed whole genome sequencing in 16 unrelated patients with autosomal recessive retinitis pigmentosa (ARRP), a disease characterized by progressive retinal degeneration and caused by mutations in over 50 genes, in search of pathogenic DNA variants. Eight patients were from North America, whereas eight were Japanese, a population for which ARRP seems to have different genetic drivers. Using a specific workflow, we assessed both the coding and noncoding regions of the human genome, including the evaluation of highly polymorphic SNPs, structural and copy number variations, as well as 69 control genomes sequenced by the same procedures. We detected homozygous or compound heterozygous mutations in 7 genes associated with ARRP (*USH2A*, *RDH12*, *CNGB1*, *EYS*, *PDE6B*, *DFNB31*, and *CERKL*) in eight patients, three Japanese and five Americans. Fourteen of the 16 mutant alleles identified were previously unknown. Among these, there was a 2.3-kb deletion in *USH2A* and an inverted duplication of ~446 kb in *EYS*, which would have likely escaped conventional screening techniques or exome sequencing. Moreover, in another Japanese patient, we identified a homozygous frameshift (p.L206fs), absent in more than 2,500 chromosomes from ethnically matched controls, in the ciliary gene *NEK2*, encoding a serine/threonine-protein kinase. Inactivation of this gene in zebrafish induced retinal photoreceptor defects that were rescued by human *NEK2* mRNA. In addition to identifying a previously undescribed ARRP gene, our study highlights the importance of rare structural DNA variations in Mendelian diseases and advocates the need for screening approaches that transcend the analysis of the coding sequences of the human genome.

medical genetics | ophthalmology | ciliopathy | retinal blindness

The identification of the genetic causes of rare Mendelian diseases is becoming increasingly important following some success with gene-based therapy, as recently reported for patients with a form of Leber congenital amaurosis (LCA), a severe autosomal recessive hereditary retinal dystrophy (1–3). The evidence that restoring a gene in the diseased retina could yield therapeutic effects has stimulated the pursuit of the genetic causes of other retinal dystrophies, including retinitis pigmentosa (RP).

RP is the name given to a group of hereditary retinal conditions in which degeneration of rod photoreceptors, responsible for vision under starlight or moonlight conditions, is more pronounced than that of cone photoreceptors, which mediate daylight vision. Individuals with RP typically experience night blindness at first, followed by progressive and unstoppable visual impairment in daytime conditions as well (4). Their visual fields become reduced gradually and sight is lost from the midperiphery to the periphery and then from the midperiphery to the center, resulting eventually in complete or near-complete blindness if left untreated. Most patients show intraretinal pigment in a bone spicule configuration around the fundus periphery, for which this condition was named. In addition, they typically show retinal arteriolar attenuation, elevated final dark adapted thresholds, and reduced and delayed electroretinograms (ERGs) (4). Vitamin A supplementation in combination with an omega-3 rich diet can slow the course of retinal degeneration and preserve visual acuity among adults with this condition (5, 6). Autosomal, recessively inherited RP (ARRP) is the most common form of hereditary retinal degeneration in humans. To date, over 50 genes have been associated with ARRP and allied disorders, among patients who are predominantly of European ancestry (RetNet; www.sph.uth.tmc.edu/retnet/home.htm). However, despite this high number of identified disease genes, ~40–50% of all diagnosed cases have no mutations in recognized loci (7). Furthermore, genetic defects in RP are also population specific. For example, a screening of 193 unrelated Japanese patients with isolate or autosomal recessive RP

Significance

Retinitis pigmentosa (RP) is a genetic disease that causes progressive blindness and that is caused by mutations in more than 50 genes. Conventional methods for identification of both RP mutations and novel RP genes involve the screening of DNA sequences spanning coding exons. In our work, we conversely test the use of whole genome sequencing, a technique that takes into account all variants from both the coding and noncoding regions of the human genome. In our approach, we identify a number of unique RP mutations, a previously undescribed disease gene, as well as pathogenic structural DNA rearrangements originating in introns.

Author contributions: K.M.N. and C.R. designed research; K.M.N., Y.P.L., E.C.O., N. Miyake, P.B., H.K.-K., and G.V. performed research; S.H., D.S., R.K.K., M.N., M.K., S.U., T.R.Y., S.I., N. Matsumoto, H.T., and E.L.B. contributed new reagents/analytic tools; K.M.N., R.G.T., Y.P.L., E.C.O., N. Miyake, P.B., H.K.-K., G.V., J.S.B., S.I., N. Matsumoto, N.K., and C.R. analyzed data; and K.M.N., E.C.O., J.S.B., E.L.B., N.K., and C.R. wrote the paper.

Conflict of interest statement: R.G.T. is an employee and shareholder of Complete Genomics, Inc.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: carlo.rivolta@unil.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1308243110/-/DCSupplemental.

MEDICAL SCIENCES

for 30 disease genes identified commonly within North American or European patients revealed candidate pathogenic mutations in only 14% of the cohort (8).

Recent advances in massively parallel sequencing have enabled the analysis of large amounts of sequences (genes) at reasonable costs, revolutionizing the traditional approach of exon-by-exon Sanger sequencing (9). The two major forms of sequencing strategies allowing large-scale analyses are whole genome sequencing (WGS) and whole exome sequencing (WES). The former reads the entire genome with no distinction between exons and non-exonic regions. It allows the detection of intergenic variants, copy number variations (CNVs), and other structural rearrangements, as well as unrecognized exonic sequences. The latter technique relies on targeted DNA capture and focuses on the analysis of the known exonic content of the genome, performed according to the genomic annotation available at a given point in time.

In this study, we performed WGS as a method for mutation discovery in a highly genetically heterogeneous Mendelian disease; to this end, we evaluated 16 unrelated RP patients from diverse ethnic backgrounds.

Results

Genome Sequencing. Genome sequencing in the 16 analyzed patients produced an average mapping yield of 200.8 ± 17.9 (mean ± SD) Gb and an average coverage of 66.1 ± 2.4 (mean ± SD) reads per base (*SI Appendix*, Table S1). This covered a genomic fraction of 0.968 ± 0.004, in which roughly 3.8 million putative variations were identified. Of these, ~7.7% were not reported in dbSNP build 131 and were classified as novel variants. Variations present within transcripts were classified further as synonymous and nonsynonymous, and analyzed separately for the North American and Japanese sets of patients. Scoring of large structural variations (SVs) could be achieved only for seven genomes, as the remaining DNA samples, possibly because of their older age, did not produce reliable mate pair information (*SI Appendix*, *Results S1*).

Assessment of pathogenic variants was performed by a series of filtering steps, summarized in Fig. 1.

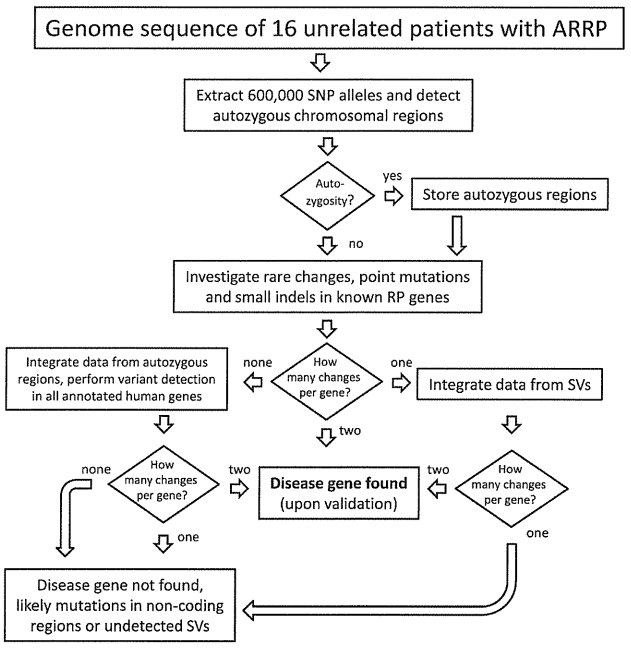

Fig. 1. Flowchart of the filtering process applied in this work.

Assessment of Autozygous Regions. Each genome was evaluated for known or undocumented parental consanguinity as well as for possible founder mutation events by extracting genotypes of known polymorphic SNPs and by searching for long intervals with high degrees of homozygosity (at least 500 consecutive SNP markers, or ~2.2 Mb on average), indicative of identity by descent (IBD). Significant genomic homozygosity was observed only in the five Japanese patients (individual IDs: R14, R15, R16, R18, and R19) who had documented parental consanguinity. The areas of IBD had essentially no overlap among these patients except for a 10-Mb interval on chromosome 1 shared by R15 and R19. Haplotype analysis indicated the shared intervals to be of different origins. No other patients carried genomic areas indicative of IBD; this was consistent with their family history reporting no parental consanguinity.

Sequence Analyses of Known RP Genes. We first focused our analyses on genes known to be associated with ARRP. We investigated both small variants (from 1 to 50 bp) from the mapping of short reads and, whenever possible, large SVs. Our results are summarized in *SI Appendix*, Table S2; detailed results are provided in *SI Appendix*, *Results S1*, Figs. S1 and S3, and Table S3.

In addition to point mutations and short indels (insertion/deletions), we detected pathogenic SVs in *USH2A* and *EYS* in patients 003–019 and R9, respectively, by combining information from sequence coverage and abnormal junctions/mate pair distance. In the genome of patient 003–019, we identified a ~2-kb deletion that removed exon 27 of *USH2A*, whereas patient R9 was found to carry a 446-kb head-to-head inverted duplication of the portion of chromosome 6 that included exons 23–29 of *EYS* (Fig. 2).

We found two pathogenic alleles, in either a homozygous or compound heterozygous state, in 8 of the 16 patients, 5 Americans and 3 Japanese, in seven different genes (*SI Appendix*, Table S2). Six patients carried mutations in one of the following genes: *USH2A*, *RDH12*, *CNGB1*, *EYS*, *PDE6B*, and *DFNB31*; 2 patients had mutations in *CERKL*. None of these mutations were found in the control cohorts of 95 healthy North American or 95 Japanese individuals. None of these mutations were reported previously, except p.R257X in *CERKL* and p.G76R in *RDH12* (10, 11). All mutations cosegregated with RP as recessive, pathogenic alleles in all family members of the index patients for whom DNA samples were available (Fig. 3).

Systematic Screening of All Genes. Based on the data from the analysis of known RP genes, we adopted a pipeline to perform a systematic analysis targeting all annotated genes in the genomes of patients with unsolved genetic etiology (*SI Appendix*, Fig. S2). With the aim of selecting a restricted number of candidate genes, more aggressive filtering was adopted with respect to the one used for the screening of known disease genes. The major differences in the analytical pipeline included removal of all entries in dbSNP. We safely applied this filtering because, given the low frequency of individual mutations in ARRP genes (including undetected ones), the risk of eliminating pathogenic DNA variants that could be fortuitously included in dbSNP build 131 is negligible. Further, to validate this approach, we applied it again retrospectively to the genomes for which mutations in RP genes were already detected. All of identified RP mutations were present in the final list of variants, supporting the sensitivity of the strategy. Detailed results are provided in *SI Appendix*, *Results S2* and are summarized in *SI Appendix*, Figs. S2 and S3 and Table S4.

In R19, in whom we did not find any clear-cut mutations in known ARRP genes, we found a homozygous frameshift variant (p.L206fs, c.617_624delTGTATGAGinsA) in the never in mitosis gene A (NIMA)-related kinase 2 (*NEK2*) gene. This variant was present within a highly homologous genomic stretch of 19.6 Mb of chromosome 1q32, predicted to be IBD (*SI Appendix*, Fig. S4).

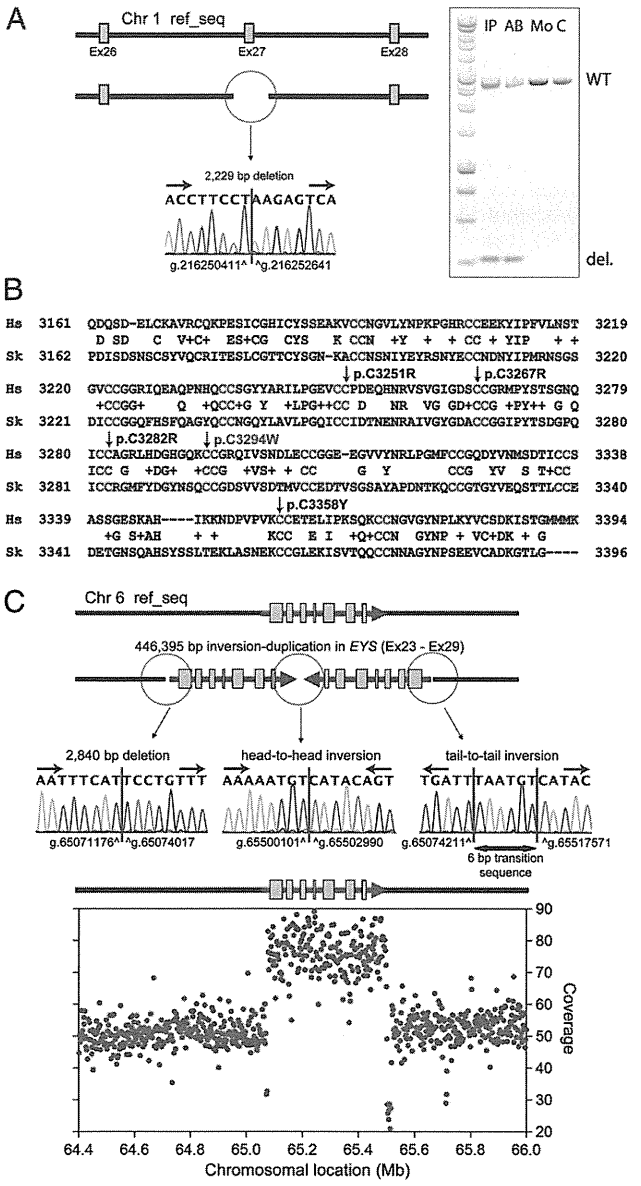

Fig. 2. Pathogenic structural variations identified. (*A*) Sequence of the heterozygous *USH2A* 2,229-bp deletion in patient 003–019 (*Left*) and electrophoresis of the PCR fragments showing a smaller fragment carrying the deletion in the index patient (IP) and her affected brother (AB) but not in her mother (Mo) or a control DNA (C). del, deleted; WT, wild type. (*B*) Alignment of the USH2A protein from *Homo sapiens* (Hs) and *Saccoglossus kowalevskii* (Sk, acorn worm) showing the conservation of 13 CC repeat motifs (red) and the location of the mutation p.C3294W, newly identified in patient 003–019 and her sister. Four previously reported disease-associated missense changes (p.C3251R, p.C3267R, p.C3282R, and p.C3358Y) also affect neighboring CC repeats. (*C*) Schematic representation and DNA sequence of the junctions characterizing the chromosomal rearrangement detected in patient R9 and involving the *EYS* gene. Integration of the information obtained by Sanger DNA sequencing and WGS coverage of the region allows identifying an inverted duplication encompassing exons 23–29.

Similar to most frameshifts producing a premature termination codon, p.L206fs is predicted to result in an mRNA allele that is subject to nonsense-mediated mRNA decay, and therefore in no protein product. Targeted DNA screening revealed that c.617_624delTGTATGAGinsA was absent from 1,273 Japanese and 95 North American control individuals. The entire coding sequence of the *NEK2* gene was then analyzed in a mixed cohort of 190 American patients with ARRP, in 64 Japanese patients with isolate RP, as well as in 13 patients found previously to show linkage between recessive retinal degeneration and the *NEK2* region. However, other than known polymorphisms (rs1056729, rs12031285, and rs45623136), we found only a few isolated heterozygous missense variants (p.R26Q, c.77G>A; p.V137I, c.409G>A; p.I265V, c.793A>G; p.N189S, c.566A>G; and p.K103E, c.307A>G; none were present in dbSNP) insufficient to account for ARRP. Notably, an additional Japanese male with ARRP was found to carry the same frameshift variant p.L206fs, but heterozygously, with no other variants in the *NEK2* coding sequence. This same patient (R51) was later found to carry the retinitis pigmentosa GTPase regulator (*RPGR*) mutation c.2405_2406delAG; p.E802fs (Human Gene Mutation Database entry: CD004115), described previously to be a sufficient cause of RP (12).

In light of a recent study reporting the involvement of noncoding RNA in the pathogenesis of retinal degeneration in mice (13), variants in noncoding RNA were also analyzed. After the removal of variants observed in 52 publicly available control genomes, only isolated heterozygous variants each with one entry per gene remained, insufficient to account for ARRP.

***nek2* Inactivation and Rescue in Zebrafish.** To validate the pathogenic role of *NEK2* deficiency in RP, we suppressed the sole ortholog of *NEK2* in zebrafish embryos and asked whether this manipulation might give rise to photoreceptor phenotypes. Upon injection of 6 ng of *nek2* splice-blocking morpholino, we observed gross ocular defects, including microphthalmia and enlarged eye sockets in 5-d postfertilization (dpf) morphant (MO) embryos (Fig. 4*A*). Whereas 63% of MO embryos displayed such phenotypes, only 21% of embryos expressing both MO and wild-type human *NEK2* mRNA did, suggesting that the ocular phenotypes are specific to the *nek2* suppression ($P < 0.001$) (Fig. 4*B*).

We next asked whether, in addition to overt structural abnormalities that may not directly inform the involvement of this gene to RP in humans, suppression of *nek2* might also give rise to photoreceptor defects consistent with those of patients with ARRP. We therefore embedded and paraffin sectioned control and MO embryos. In addition to the small eye phenotype, we detected alterations in the photoreceptor layer. Specifically, after

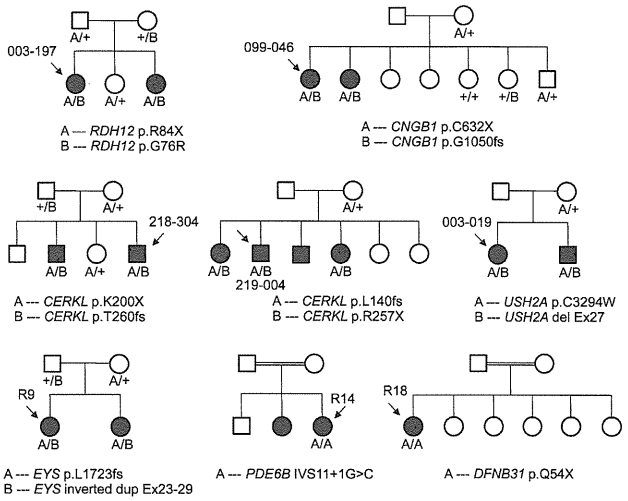

Fig. 3. Cosegregation analyses. All mutations analyzed cosegregated with the disease according to an autosomal recessive pattern of inheritance.

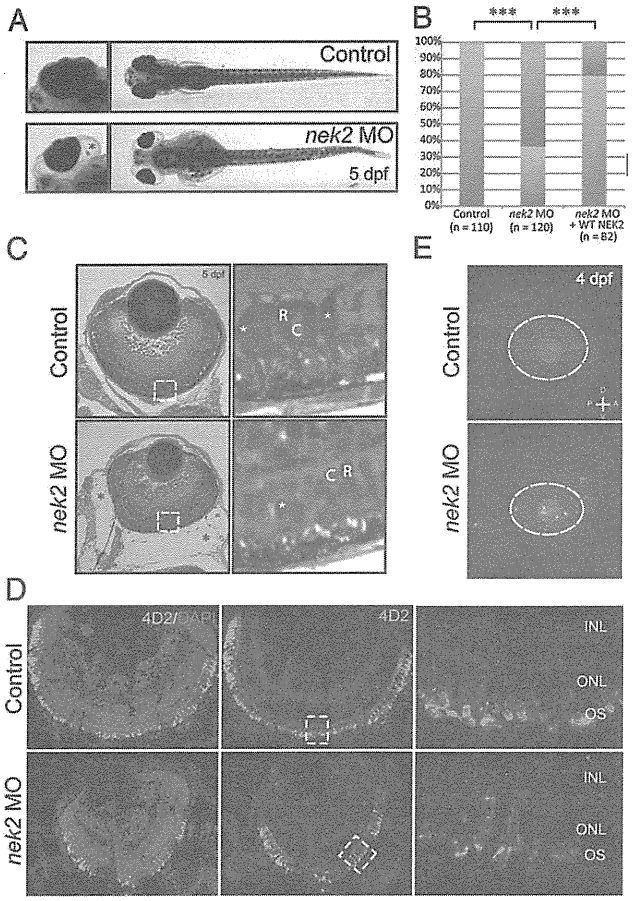

Fig. 4. In vivo functional evaluation of *nek2* loss in zebrafish. (*A*) Bright-field representation of 5-dpf control and *nek2* morphant zebrafish embryos. Magnified *Insets* highlight ocular phenotypes including microphthalmia and enlarged eye sockets (marked by the black asterisk). (*B*) Ocular phenotypes including microphthalmia and enlarged eye sockets vs. normal phenotypes (red bars and blue bars, respectively) are quantified in control and *nek2* morphant embryos, as well as in morphant animals rescued with human WT *NEK2* mRNA. Asterisks indicate statistically significant differences between groups ($P < 0.001$). (*C*) Histology of control and *nek2* morphant embryos also show enlarged eye sockets (marked by black asterisks) and microphthalmia. Magnified *Insets* show a decrease in the number of photoreceptors with apparent changes in domains of condensed chromatin (white asterisks). C, cones; R, rods. (*D*) Immunohistochemical analyses of retinal cryosections from control and *nek2* MO embryos, stained with DAPI (blue) and the 4D2 antibody against rhodopsin (green). Suppression of *nek2* results in the depletion of rods and in the mislocalization of rod opsin from the outer segment (OS) of photoreceptors. INL, inner nuclear layer; ONL, outer nuclear layer. (*E*) TUNEL immunofluorescent images of 4-dpf embryos, showing an increase in the number of apoptotic cells in *nek2* morphant embryos. The dotted ovals indicate the position of the eye. A, anterior; D, dorsal; P, posterior; V, ventral.

serial sectioning of 10–20 embryos injected with sham, MO, or MO + human *NEK2* mRNA, we observed a persistent decrease in the number of photoreceptors with large central domains of condensed chromatin. This phenotype was seen in all *nek2* MO embryos evaluated, but was absent from embryos injected with either sham or MO + human *NEK2* mRNA, suggesting a loss of rod photoreceptors specific to the suppression of *nek2* (Fig. 4*C* and *SI Appendix*, Fig. S5). To verify this observation, we used a rhodopsin (4D2) antibody to stain retinal cryosections from embryos injected with sham, MO, or MO + human *NEK2* mRNA (Fig. 4*D* and *SI Appendix*, Fig. S5). Immunohistochemical analyses of cross-sections from each condition demonstrated that the suppression of *nek2* resulted in the depletion of ~24% of 4D2-positive rod photoreceptors. In addition, mislocalization of rod opsin throughout the photoreceptor cells was evident in the central retina of *nek2* MO specimens, consistent with the hypothesis that nek2 is required for the appropriate trafficking of rhodopsin to the outer segments (Fig. 4*D*).

Further, to ask whether apoptosis, a major mechanism of photoreceptor loss in most known forms of RP (14), might account for some of the observed loss of photoreceptors, we performed TUNEL analysis. Masked scoring of embryos (~50 embryos per injection mixture) revealed a sevenfold increase in the number of TUNEL-positive cells in the eye and head region of *nek2* morphant embryos. By sharp contrast, we did not observe more than 1–10 TUNEL-positive cells in embryos injected with MO + *NEK2* mRNA (Fig. 4*E*).

Finally, we were intrigued by the discovery of a heterozygous frameshift variant p.L206fs in *NEK2* and the bona fide *RPGR* mutation p.E802fs in a patient with RP. We therefore asked whether the *RPGR* variant may interact genetically with the *NEK2* locus. To test this possibility, we coinjected subeffective doses of the *nek2* MO and *rpgr* MO and compared embryos with single or double MO ($n > 100$ at subeffective doses). Approximately 28% of embryos carrying subeffective doses of both *nek2* and *rpgr* MO revealed ocular ($P < 0.001$) and rod photoreceptor phenotypes (serial sectioning of 10 embryos per genotype) that exceeded the number of affected embryos induced by either *nek2* (3%) or *rpgr* (10%) MO alone, suggesting that the RPGR allele interacts *in trans* with the NEK2 locus to exacerbate photoreceptor defects (*SI Appendix*, Fig. S6).

Discussion

Massively parallel sequencing has proven to have a high potential to detect mutations in patients with rare Mendelian diseases (15). To date, most reports focus on monogenic conditions with no genetic heterogeneity, for which mutations can be recognized from benign variants since they invariantly affect the same gene in different patients.

In this study, we explored the efficacy of WGS in identifying mutations in unrelated patients from diverse ethnic backgrounds and presenting with a disease that is clinically the same but that has different genetic drivers. Whereas the small number of genomes analyzed in this study precludes an accurate analysis of quantitative measures, such as sensitivity of the WGS to detect mutations in known RP genes, we observed a few features that allowed us to make some valid comparisons between the different techniques currently available for genetic diagnosis. First, the majority of the pathogenic mutations identified were never reported before. This implies that tools that rely on systematic search for known pathogenic variants, both via mutation-centered resequencing and chip-based hybridization, may not be adequate for ARRP. Second, thanks to full-genome data, we detected complex structural variants whose junctions were located deep in noncoding regions. Because of their nature, these disease-causing variants would have been invisible to standard screening methods, or even to WES. Coverage-based analysis of CNV in exome sequencing has been attempted, with variable results. Limitations of this approach include the uneven efficiency of target DNA capture (and hence sequence coverage, on which assessment of number of copies is based) over different probes and, above all, the low probability of detecting junctions defining the SVs, which are more likely to be found in the nonexonic sequences composing ~98% of our genome. Unambiguous detection of abnormal junctions and mate pair information are crucial parameters in defining a SV; for instance, they allow distinguishing a tandem duplication from an inverted one. Third, because we had access to the full wealth of genomic information, we could integrate many sources of information

at once (e.g., SNP genotypes, phasing, etc.) that allowed us to accurately filter DNA variants that were related to the disease.

Genetic defects in *EYS* were proposed recently to be one of the major causes of ARRP in the Japanese population (16). We found that one of the pathogenic *EYS* alleles was a large SV (446 kb) with a complex genomic rearrangement. This finding supports the notion that SVs represent frequent pathogenic mutations in this gene (17). A homozygous nonsense mutation in exon 6 of *DFNB31* was identified in R18, a patient with nonsyndromic ARRP. The *DFNB31* gene encodes whirlin, a PDZ scaffold protein with expression in both hair cell stereocilia and retinal photoreceptor cells. Whirlin binds to the protein encoded by *USH2A* (18), a gene associated with both Usher syndrome type II (ARRP accompanied by hearing loss) and nonsyndromic ARRP (19). Whereas mutations in *DFNB31* have been reported as rare causes of Usher syndrome type II (20, 21), no DNA changes in its sequence have yet been associated with nonsyndromic ARRP. However, at the age of 66, the past medical history of this patient was significant for only hyperlipidemia and she did not report any hearing loss. We could not perform an auditory examination because she was no longer reachable.

In patients from consanguineous families, regions of IBD allowed restricting the search for pathogenic mutations to only a fraction of the genome. However, these same regions were susceptible to carrying other rare but nonpathogenic homozygous changes as well. Indeed, a higher number of candidate genes/mutations remained among Japanese patients with parental consanguinity compared with those without it (*SI Appendix*, Table S4). These results suggest that even if the analysis should be restricted to areas of IBD, genomes with high homozygosity do not necessarily offer an extra advantage in mutation detection, when comprehensive genomic sequencing in single individuals is performed.

In three patients we identified clear-cut pathogenic but heterozygous mutations in known ARRP genes that could not be associated directly with the disease. This was particularly evident for patient R14, who carried a heterozygous frameshift in *DFNB31* but was also homozygous for a mutation inactivating *PDE6B* (22). These findings are not surprising, given the elevated number of recessive ARRP mutations that are predicted to be present in the general population. Based both on theoretical assessments and on experimental data from control cohorts, we estimated that 1 in 3–7 individuals could be potential heterozygous carriers of an ARRP mutation (23, 24) or, as in the present case, 3 in 16.

The reasons why no candidate mutations of similar quality (i.e., two mutations, at least one of them being clearly deleterious in nature) to those revealed in known RP genes was uncovered in most of the unresolved genomes are unknown. Explanations for this observation may include the presence of variants or SVs that were undetected because of problems inherent in the mapping or sequencing procedure, or of less obvious pathogenic changes that alter splicing or transcription. These would include variants located in introns or in promoter regions, synonymous changes, or changes lying within important yet unannotated exons, genes, or genetic elements that have not been explored in the current study. Diseases caused by oligogenic modes of inheritance, or perhaps attributable to missense mutations for which efficient prioritization is difficult, is another possible explanation. De novo mutations in unknown dominant RP genes could also be evoked.

The search for mutations in unknown disease-causing genes revealed a number of genes with two nonsynonymous changes, which were mostly previously undescribed missenses. Application of more stringent filtering criteria by imposing the presence of at least one deleterious mutation followed by targeted annotation highlighted a single candidate, *NEK2*, in a Japanese patient who carried a homozygous frameshift in this gene. The serine/threonine-protein kinase NEK2 is known to play an important role in regulation of cell cycle progression through localization to the centrosomes and interaction with microtubules (25). The identified frameshift would result either in the creation of premature stop codon yielding a null allele or (less likely) a truncated protein lacking kinase activity and loss of microtubule binding. Importantly, defects in members of the Nek kinase family have been linked to impaired ciliogenesis and polycystic kidney disease (26). Recently, a role for *Nek2* in the left–right patterning of vital organs (a phenotype associated with ciliary function) was established in *Xenopus laevis* (27). In the same work, in situ hybridization revealed the expression of *nek2* transcripts in the eye (27). Furthermore, because NEK2 interacts with and can phosphorylate rootletin, a component of photoreceptor cilia (28, 29), *NEK2* was considered to be an important candidate for ARRP.

Our zebrafish studies showed that lack of Nek2 induces microphthalmia as a gross morphological phenotype. More importantly, in *nek2* morphants, we observed mistrafficking of rhodopsin, a hallmark of photoreceptor disease (30), and a reduced number of rod photoreceptors, likely via apoptotic processes. These phenotypes were rescued by injection of wild-type human *NEK2* mRNA, validating the specificity of the induced defects. Microphthalmia is a phenotype that is difficult to interpret in the present context but that is not uncommon to zebrafish models of RP (31, 32). Meanwhile, photoreceptor death, mistrafficking of rhodospsin, and reduction of the outer retinal layers are classical features of RP in both patients and animal models (7, 14, 33). Indeed, no microphthalmia was noted in patient R19.

Intriguingly, the *NEK2* frameshift identified in R19 was also present in R51, another patient with RP who had a deleterious mutation in *RPGR*. As the *RPGR* mutation in itself could explain the disease, an obvious question was whether the *NEK2* mutation might in fact represent a common benign allele. We therefore searched for this variant in 1,273 control Japanese individuals and found that none carried it (allele frequency $<3.9 \times 10^{-4}$). The p.L206fs mutation in *NEK2* is therefore exceedingly rare, such that its presence in a homozygous state in a patient is a strong argument in favor of its being an uncommon cause for ARRP. Although it is possible to attribute the presence of both *NEK2* and *RPGR* mutations in R51 to chance, a more parsimonious explanation is that mutations in these two genes, both expressed in the connecting cilium, act synergistically to define a severe RP phenotype, due to the established principles of mutational load and oligogenic interactions of pathogenic alleles (34). In turn, this would increase the likelihood for the patient of being examined at earlier ages and analyzed genetically. Multiple genetic modifier genes have been reported for cilia-encoding genes and especially for RPGR (35). These modifiers may account in part for the wide phenotypic spectrum associated with genetic defects in this gene, ranging from localized macular atrophy to retinitis pigmentosa of variable severity. To investigate the possibility of the cooperative effect between deficiencies in these two ciliary genes, we performed in vivo genetic interaction studies and showed that loss of Rpgr function can exacerbate Nek2 ocular phenotypes, including defects comprising the trapping of rhodopsin in the inner segment. Taken together, our genetic and functional data indicate that *NEK2* is a disease gene and that the retinal phenotype that results from its deficiency may represent a newly recognized ciliopathy.

To date, WGS has not been as widely explored as WES in the context of mutation detection. This can be attributed mainly to cost-related issues, because WGS is at least twice as expensive as WES procedures ensuring the same average coverage. We believe that the additional features displayed by WGS are worth the difference in price; however, this is a rather subjective matter that also depends on the disease that is being investigated. In the present case, WGS was essential to identify two pathogenic structural variations originating in introns. This is a significant finding, considering that only seven genomes could undergo SV analysis. Therefore, as a general rule, WGS is probably the strategy of