and c.694C>T (p.Arg232Cys) mutants showed mislocalization, whereas c.466G>A (p.Asp156Asn) and c.899G>C (p.Cys300Ser) mutants showed normal localization (Figure S1). To investigate whether the *B3GALT6* missense mutations affect the enzyme function, the GalT-II activities of soluble FLAG-tagged proteins for WT and mutant B3GALT6 proteins were assayed. The GalT-II activities of p.Ser65Gly-, p.Pro67Leu-, p.Asp156Asn-, p.Arg232Cys-, and p.Cys300Ser-B3GALT6 were significantly decreased compared with WT-B3GALT6 (Figure 2C), indicating that these mutations resulted in a loss of enzyme function. On the other hand, there were no significant differences in the GalT-II activities between WT-B3GALT6 and p.Glu174Asp-B3GALT6, a common polymorphism (rs12085009) in the public database (Figure 2C).

All SEMD-JL1 individuals with the *B3GALT6* mutation had the characteristic skeletal abnormalities, including platyspondyly, short ilia, and elbow malalignment (Table 1 and Figure S2); however, some had a range of extraskeletal and connective tissue abnormalities that overlapped with those seen in Ehlers-Danlos syndrome, progeroid form (EDS-PF [MIM 130070]). EDS-PF is an autosomal-recessive connective tissue disorder characterized by sparse hair, wrinkled skin, and defective wound healing with atrophic scars.[10] In addition, skeletal abnormalities so far reported in EDS-PF are limited to generalized osteopenia and radial head dislocation, which are in contrast with the severe generalized dysplasias of the axial and appendicular skeleton observed in SEMD-JL1. Thus, both disorders at first glance appear as separate clinical entities, although they share the clinical features of short stature, joint laxity and dislocation, and facial dysmorphism. In two families with individuals with EDS-PF, recessive mutations of *B4GALT7* (MIM 604327) have been found.[11,12] *B4GALT7* (RefSeq NM_007255.2) encodes an enzyme, xylosylated protein β-1,4-galactosyltransferase, that catalyzes the second step of the GAG linker region biosynthesis (Figure 1). Therefore, we speculated that B3GALT6 and B4GALT7 deficiencies might show similar phenotypes. We then examined *B3GALT6* in four additional individuals (P9–P12) who had phenotypes compatible with EDS-PF (Table 1 and Figure S3) but in whom no *B4GALT7* mutations had been found. Sanger sequencing of the EDS-PF-like subjects revealed that all were compound heterozygotes for *B3GALT6* mutations (Table 2). There were two frameshift mutations and one missense mutation (c.925T>A [p.Ser309Thr]) common in two families (F8 and F9). We investigated the enzyme function of the missense mutation by using the same assay for SEMD-JL1 missense mutations. The GalT-II activities of p.Ser309Thr-B3GALT6 were significantly decreased (Figure S4).

Collectively, 11 different mutations in individuals from 10 families were identified in *B3GALT6* by a combination of exome and targeted sequencing (Table 2 and Figure S5). None of these mutations were detected in more than 200 ethnicity-matched controls or in public databases, including the 1000 Genomes database, indicating that they are unlikely to be polymorphisms. SEMD-JL1 and EDS-PF-like individuals had no common mutations (Table 2). The individuals with *B3GALT6* mutations were short at birth and their short stature worsened with age. Their common clinical features were a flat face with prominent forehead and kyphoscoliosis (Table 1). Kyphoscoliosis was noticed in infancy in most cases and even in utero in severe cases. Although skeletal changes were essentially the same, craniofacial and skin abnormalities, joint laxities, and muscular hypotonia were variable among the individuals with *B3GALT6* mutations. Common radiographic features were platyspondyly that becomes less conspicuous with age, short ilia, and elbow malalignment (Table 1). Prominent lesser trochanters and metaphyseal flaring were seen in most cases. No individuals showed generalized osteoporosis. The disease phenotype was very variable between families (mutations), but in two familial cases, phenotypes were similar between the pair of the sibs. As a corollary, our results indicate that EDS-PF is genetically heterogeneous, with a proportion of cases being caused by mutations in *B4GALT7* and another in *B3GALT6*.

Diseases caused by defects in enzymes involved in the biosynthesis of the GAG linker region are categorized as the GAG linkeropathy. The first member of GAG linkeropathy has been identified to arise from an EDS-PF/B4GALT7 deficiency. *B4GALT7* mutations have been identified in homozygous c.808C>T (p.Arg270Cys)[12] and compound heterozygous (c.557C>A [p.Ala186Asp] and c.617T>C [p.Leu206Pro])[11] states. Another member of GAG linkeropathy manifests itself as Larsen-like syndrome, B3GAT3 type (MIM 245600). A family with individuals harboring a homozygous *B3GAT3* (MIM 606374; RefSeq NM_012200.3) mutation (c.830G>A [p.Arg227Gln]) has been identified. The clinical features of five affected individuals of the family are characterized by dislocation and laxity of joints and congenital heart defects.[11] The former considerably overlaps with the phenotypes of SEMD-JL1 and EDS-PF, two other GAG linkeropathies; however, the association of heart defects has critically differentiated this disease from the others (Figure 1).

Given that the linker region biosynthesis is nonparallel and that the defects in the three enzymes simply affect the amounts of the linker region available to form GAGs (CS, HS, DS), phenotypic similarities of the three diseases are quite understandable. The quantitative difference of the phenotypes (severity of the diseases) most probably results from the difference in the degree of enzyme defects resulting from mutations. On the other hand, qualitative differences of the three diseases (e.g., scoliosis caused by the *B3GALT6* mutation, heart disease caused by the *B3GAT3* mutation, etc.) suggest other explanations. Tissue expression patterns of the three genes do not entirely explain the differences. We examined their mRNA expression in various human tissues, including cartilage, bone, and connective tissues by quantitative real-time PCR (Figure S6). We detected strong expression of *B3GALT6* in

Table 3. The Amount of GAGs in the Lymphoblastoid Cells from Individuals with Spondyloepimetaphyseal Dysplasia with Joint Laxity Type 1

	GAG (Disaccharides/mg Acetone Powder)[a] [pmol]			
Subject	**CS/DS**	**CS**	**DS**	**HS**
Control	62	48	29	128
SEMD-JL1				
P1	313	295	118	15
P2	345	175	60	21
P3	270	162	28	20

[a]Calculated based on the peak area in chromatograms of digests with a mixture of chondroitinases ABC and AC-II (CS/DS), chondroitinases AC-I and AC-II (CS), chondroitinase B (DS), and heparinases I and III (HS).

cartilage and bone but only weak expression in skin, ligament, and tendon. *B4GALT7* expression was stronger in cartilage than *B3GALT6* and also weak in skin and ligament. *B3GAT3* expression was not specific to heart. The qualitative difference may result from the difference in the effects of the three genes on GAG formation.

To examine how *B3GALT6* mutations affects the products of GAGs in vivo, we measured the amounts of CS and HS chains at the surface of lymphoblastoid cells from the subjects by flow cytometry by using CS-stub and HS-stub antibodies as previously described.[13–15] In brief, purified GAG fractions were treated individually with a mixture of chondroitinases ABC and AC-II, a mixture of chondroitinases AC-I (EC 4.2.2.5) (Seikagaku Corp.) and AC-II (EC 4.2.2.5) (Seikagaku Corp.), chondroitinase B (EC 4.2.2.19) (IBEX Technologies), or a mixture of heparinases-I and -III (IBEX Technologies) for analyzing the disaccharide composition of CS/DS, CS, DS, and HS, respectively. The digests were labeled with a fluorophore 2-aminobenzamide (2AB) and aliquots of the 2AB derivatives of CS/DS/HS disaccharides were analyzed by anion-exchange HPLC on a PA-03 column (YMC Co.). The HS-stub antibody (3G10) showed a markedly reduced binding to the epitopes on the subjects' cells (Figure S7). The relative numbers of the HS chains presented as the mean fluorescence intensity (MFI) of the cell population stained with the antibody for P1, P2, and P3 were 26%, 56%, and 35% of the control, respectively. On the other hand, the CS-stub antibody (2B6) showed a similar binding to the epitopes on the subjects' cells relative to those of the control (Figure S7). The MFI for P1, P2, and P3 were 114%, 104%, and 106% of the control, respectively. Furthermore, we measured disaccharide of GAG chains from lymphoblastoid cells by using anion-exchange HPLC after digestion with chondroitinase and heparinase. The amounts of the disaccharide from HS chains were significantly decreased, whereas CS and DS chains were ~5 times higher than those in the control (Table 3).

Previous biochemical studies on EDS-PF with *B4GALT7* mutations show a reduction in the synthesis of DS chains.[16,17] The c.830G>A (p.Arg227Gln) mutation in *B3GAT3* causes a drastic reduction in GlcAT-I activity in fibroblasts of the individual with SEMD-JL1 and numbers of CS and HS chains on the core proteins at the surface of the fibroblasts are decreased to about half of the controls.[11] Cultured lymphoblastoid cells from individuals with a c.419C>T (p.Pro140Leu) mutation in *B3GAT3* show that defective synthesis is more pronounced for CS than for HS.[11] Taken together with our results, these findings suggest that the effects of the deficiencies of the three enzymes on GAG synthesis are not identical. A possible explanation for the qualitative phenotypic differences may be that the biosynthesis of the GAG linker region is not a simple step-by-step addition but involves parallel processing and/or alternative pathways. Other glycosyltransferases may have similar biochemical functions to these three enzymes and thus complement their deficient activities to variable degrees in cell- and/or tissue-specific manners, leading to differences in the amount of GAGs in the tissues. It is known that B3GALT6 and B4GALT7 have several homologs.[18] It must be noted that all biochemical studies so far have been performed in vitro or in cultured cells, and therefore there is a severe limitation to our understanding of the pathogenesis at tissue and organ levels.

By exome sequencing, we identified loss-of-function mutations in *B3GALT6* in 12 individuals from 10 families. The mutations produced a spectrum of connective tissue disorders characterized by lax skin, muscle hypotonia, joint dislocation, and skeletal dysplasia and deformity, which include phenotypes previously known as SEMD-JL1 and EDS-PF (Figures S1 and S2). The pleiotropic phenotypes of *B3GALT6* mutations indicate that B3GALT6 plays critical roles in development and homeostasis of various tissues, including skin, bone, cartilage, tendons, and ligaments. Biochemical studies that used lymphoblastoid cells of the individuals with *B3GALT6* mutations showed a decrease of HS and a paradoxical increase of CS and DS of the cell surface. Further clinical, genetic, and biological studies are necessary to understand the pathological mechanism of the diseases caused by enzyme defects involved in the biosynthesis of the GAG linker region and roles of the region in GAG metabolism and function.

Supplemental Data

Supplemental Data include seven figures and two tables and can be found with this article online at http://www.cell.com/AJHG/.

Acknowledgments

We thank the individuals with the disease and their family for their help to the study. We also thank the Japanese Skeletal Dysplasia Consortium. This study is supported by research grants from the Ministry of Health, Labor, and Welfare (23300101 to S.I. and N. Matsumoto; 23300201 to S.I.), by Grants-in-Aid for Young Scientists (23689052 to N. Miyake and 23790066 to S.M.) from the Japan Society for the Promotion of Science; by the Matching Program for Innovations in Future Drug Discovery and Medical Care

(K.S.); by The Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT); by a Grant-in-aid for Encouragement from the Akiyama Life Science Foundation (S.M.); by Swiss National Science Foundation Grants (31003A_141241 and 310030_132940); by The CoSMO-B project (Brazil and Switzerland); by the Leenaards Foundation (Switzerland); and by Research on intractable diseases, Health and Labour Sciences Research Grants, H23-Nanchi-Ippan-123 (S.I.).

Received: February 1, 2013
Revised: March 16, 2013
Accepted: April 5, 2013
Published: May 9, 2013

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, http://browser.1000genomes.org
ANNOVAR, http://www.openbioinformatics.org/annovar/
dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/
GATK, http://www.broadinstitute.org/gatk/
MutationTaster, http://www.mutationtaster.org/
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
Novoalign, http://www.novocraft.com/main/page.php?s=novoalign
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org/
Picard, http://picard.sourceforge.net/
PolyPhen, http://www.genetics.bwh.harvard.edu/pph2/
RefSeq, http://www.ncbi.nlm.nih.gov/RefSeq
SIFT, http://sift.bii.a-star.edu.sg/
UCSC Genome Browser, http://genome.ucsc.edu

References

1. Warman, M.L., Cormier-Daire, V., Hall, C., Krakow, D., Lachman, R., LeMerrer, M., Mortier, G., Mundlos, S., Nishimura, G., Rimoin, D.L., et al. (2011). Nosology and classification of genetic skeletal disorders: 2010 revision. Am. J. Med. Genet. A. *155A*, 943–968.
2. Beighton, P., Gericke, G., Kozlowski, K., and Grobler, L. (1984). The manifestations and natural history of spondylo-epimetaphyseal dysplasia with joint laxity. Clin. Genet. *26*, 308–317.
3. Nishimura, G., Satoh, M., Aihara, T., Aida, N., Yamamoto, T., and Ozono, K. (1998). A distinct subtype of "metatropic dysplasia variant" characterised by advanced carpal skeletal age and subluxation of the radial heads. Pediatr. Radiol. *28*, 120–125.
4. Boyden, E.D., Campos-Xavier, A.B., Kalamajski, S., Cameron, T.L., Suarez, P., Tanackovic, G., Andria, G., Ballhausen, D., Briggs, M.D., Hartley, C., et al. (2011). Recurrent dominant mutations affecting two adjacent residues in the motor domain of the monomeric kinesin KIF22 result in skeletal dysplasia and joint laxity. Am. J. Hum. Genet. *89*, 767–772.
5. Min, B.J., Kim, N., Chung, T., Kim, O.H., Nishimura, G., Chung, C.Y., Song, H.R., Kim, H.W., Lee, H.R., Kim, J., et al. (2011). Whole-exome sequencing identifies mutations of KIF22 in spondyloepimetaphyseal dysplasia with joint laxity, leptodactylic type. Am. J. Hum. Genet. *89*, 760–766.
6. Miyake, N., Elcioglu, N.H., Iida, A., Isguven, P., Dai, J., Murakami, N., Takamura, K., Cho, T.J., Kim, O.H., Hasegawa, T., et al. (2012). PAPSS2 mutations cause autosomal recessive brachyolmia. J. Med. Genet. *49*, 533–538.
7. Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., Kaname, T., Naritomi, K., Kawame, H., Wakui, K., et al. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat. Genet. *44*, 376–378.
8. Bai, X., Zhou, D., Brown, J.R., Crawford, B.E., Hennet, T., and Esko, J.D. (2001). Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the beta 1,3-galactosyltransferase family (beta 3GalT6). J. Biol. Chem. *276*, 48189–48195.
9. Saunders, C.J., Minassian, B.E., Chow, E.W., Zhao, W., and Vincent, J.B. (2009). Novel exon 1 mutations in MECP2 implicate isoform MeCP2_e1 in classical Rett syndrome. Am. J. Med. Genet. A. *149A*, 1019–1023.
10. Kresse, H., Rosthøj, S., Quentin, E., Hollmann, J., Glössl, J., Okada, S., and Tønnesen, 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.
11. Baasanjav, S., Al-Gazali, L., Hashiguchi, T., Mizumoto, S., Fischer, B., Horn, D., Seelow, D., Ali, B.R., Aziz, S.A., Langer, R., et al. (2011). Faulty initiation of proteoglycan synthesis causes cardiac and joint defects. Am. J. Hum. Genet. *89*, 15–27.
12. Faiyaz-Ul-Haque, M., Zaidi, S.H., Al-Ali, M., Al-Mureikhi, M.S., Kennedy, S., Al-Thani, G., Tsui, L.C., and Teebi, A.S. (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. A. *128A*, 39–45.
13. Kinoshita, A., and Sugahara, K. (1999). Microanalysis of glycosaminoglycan-derived oligosaccharides labeled with a fluorophore 2-aminobenzamide by high-performance liquid chromatography: application to disaccharide composition analysis and exosequencing of oligosaccharides. Anal. Biochem. *269*, 367–378.
14. Miyake, N., Kosho, T., Mizumoto, S., Furuichi, T., Hatamochi, A., Nagashima, Y., Arai, E., Takahashi, K., Kawamura, R., Wakui, K., et al. (2010). Loss-of-function mutations of CHST14 in a new type of Ehlers-Danlos syndrome. Hum. Mutat. *31*, 966–974.
15. Mizumoto, S., and Sugahara, K. (2012). Glycosaminoglycan chain analysis and characterization (Glycosylation/Epimerization). In Methods in Molecular Biology. In Proteoglycans: Methods and Protocols, F. Rédini, ed. (New York, USA: Humana Press, Springer), pp. 99–115.
16. Okajima, T., Fukumoto, S., Furukawa, K., and 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.
17. Quentin, E., Gladen, A., Rodén, L., and Kresse, H. (1990). A genetic defect in the biosynthesis of dermatan sulfate proteoglycan: galactosyltransferase I deficiency in fibroblasts from a patient with a progeroid syndrome. Proc. Natl. Acad. Sci. USA *87*, 1342–1346.
18. Togayachi, A., Sato, T., and Narimatsu, H. (2006). Comprehensive enzymatic characterization of glycosyltransferases with a beta3GT or beta4GT motif. Methods Enzymol. *416*, 91–102.

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.

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.

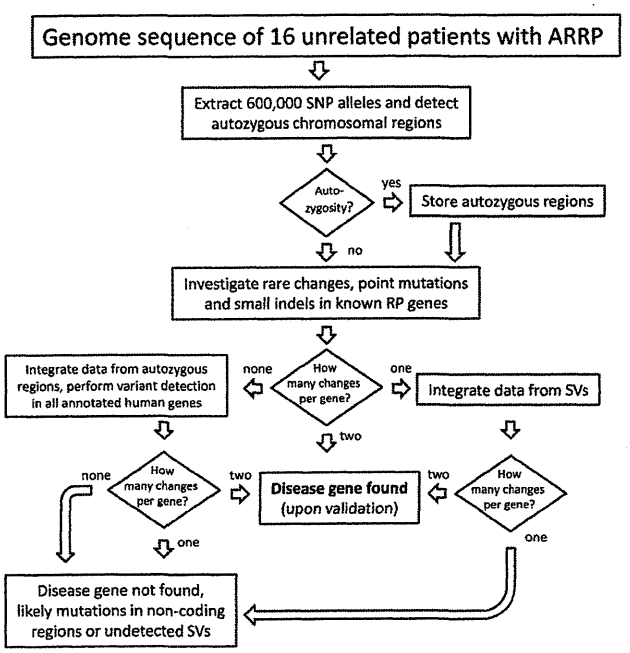

Fig. 1. Flowchart of the filtering process applied in this work.

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).

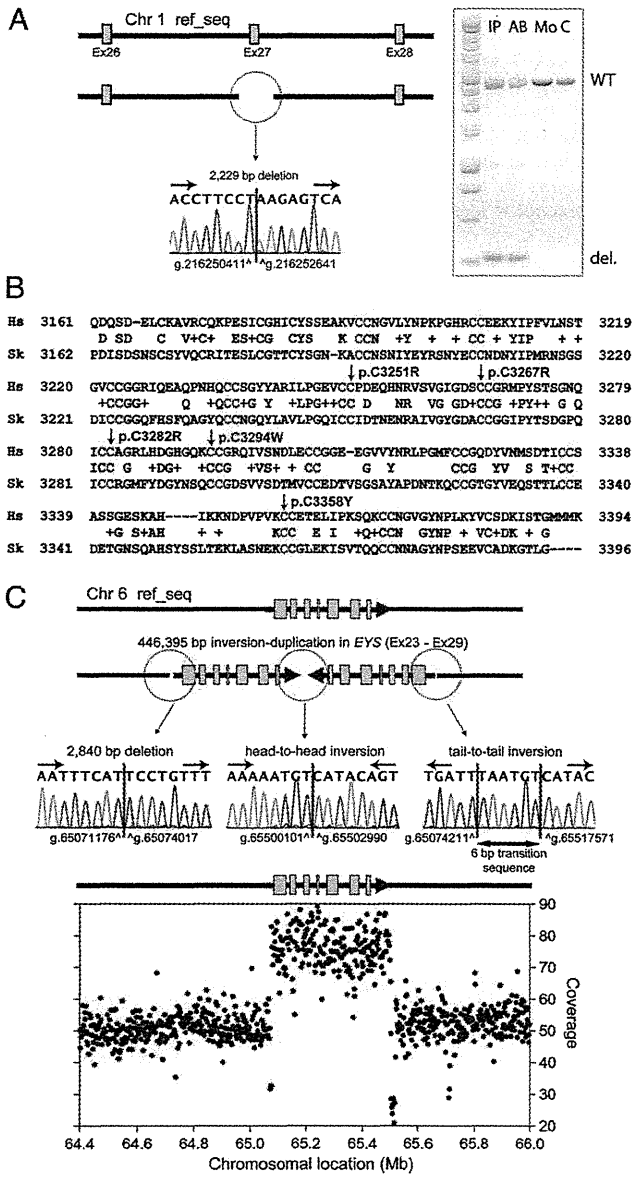

Fig. 2. Pathogenic structural variations identified. (*A*) Sequence of the heterozygous *USH2A* 2,229-bp deletion in patient 003–019 (*Left*) and electrophoresis of the PCR fragments showing a smaller fragment carrying the deletion in the index patient (IP) and her affected brother (AB) but not in her 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

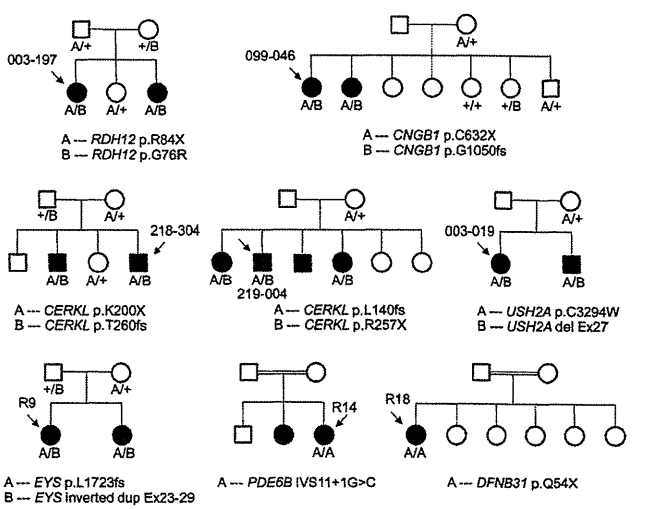

Fig. 3. Cosegregation analyses. All mutations analyzed cosegregated with the disease according to an autosomal recessive pattern of inheritance.

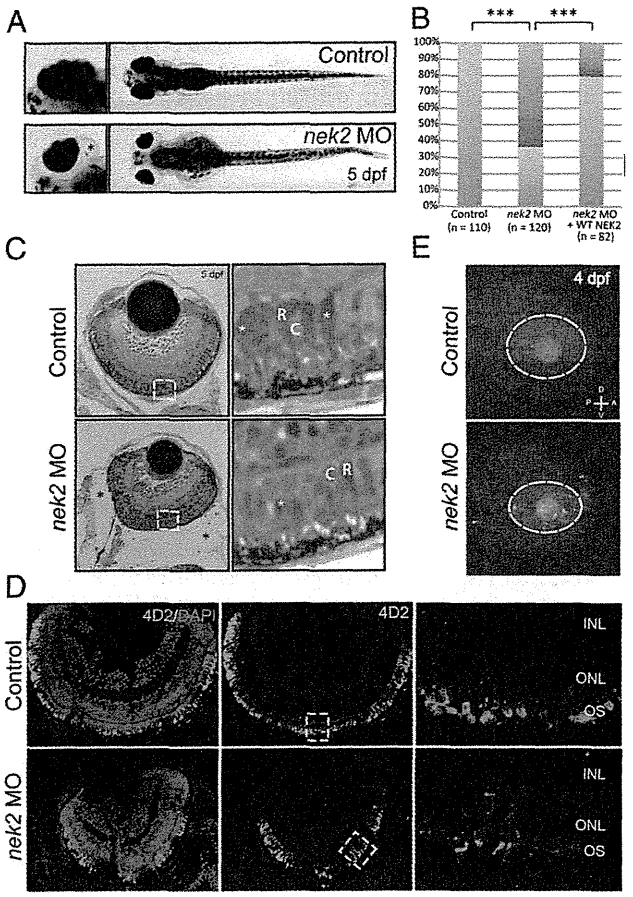

Fig. 4. In vivo functional evaluation of *nek2* loss in zebrafish. (*A*) Bright-field representation of 5-dpf control and *nek2* morphant zebrafish embryos. Magnified *Insets* highlight ocular phenotypes including microphthalmia and enlarged eye sockets (marked by the black asterisk). (*B*) Ocular phenotypes 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

choice when detection of structural variants or mutations in noncoding regions represents an important element of investigation. In the long term, considering that costs associated with massively parallel sequencing technology is expected to fall further and that analysis pipelines continue to evolve, it is probable that WGS would be just as workable economically and physically as WES. Limitations of WGS include the requirement of high-quality DNA to explore the full leverage of the mate-pair mapping and the lack of reliable pipelines to detect SVs ranging in size from 50 to a few hundred base pairs. Unexpectedly, the difficulty accompanied by handling the large amount of data produced by WGS was not a significant obstacle, given the power of desktop computers presently available on the market. Whereas samples with suitable quality could be obtained through careful preparation of fresh DNA samples, under detection of SVs may be a more problematic issue to solve. This occurs because current mapping is based on two steps: mapping of the short reads aimed at detecting variations between 1 and 50 bp and mate-pair mapping for detection of SVs larger than a few hundred bases; to our knowledge, a solution that could fill the gap between these two mapping approaches remains to be found.

In conclusion, in this study we identified clear-cut causative mutations among the overwhelming number of DNA variants present in the human genome, in single patients from genetically diverse populations. This happened without ambiguities in a highly heterogeneous disease, ARRP, and in more than 50% of the individuals analyzed. Furthermore, two cases presented mutations involving noncoding parts of the genome. Considering that the majority of patients referred for molecular genetics diagnosis are isolated individuals, our results are relevant not only to basic research, but also to future clinical genetic testing.

Methods

Our research protocol involving humans and animals was approved by the institutional review boards of our respective universities and organizations. Written informed consent for providing medical information and blood samples was obtained from each patient. Experimental procedures are described in detail in *SI Appendix, Methods*.

ACKNOWLEDGMENTS. We thank Anna M. Siemiatkowska and Frans P. M. Cremers for sharing material from a person with RP, Adriana Ransijn for technical help, as well as Andrea Superti-Furga, and Luisa Bonafé for fruitful suggestions. Data storage was ensured by the Vital-IT Center for high-performance computing of the Swiss Institute of Bioinformatics. This work was supported by the Swiss National Science Foundation (Grant 310030_138346) and the Gebert Rüf Foundation, Switzerland (Rare Diseases-New Technologies Grant) (both to C.R.); a Center Grant from the Foundation Fighting Blindness (to E.L.B.); National Institutes of Health Grants DK072301 and MH-084018 (to N.K.); Ministry of Health, Labor and Welfare (MHLW) of Japan [Grant 23300101 (to S.I. and N. Matsumoto) and Grant 23300201 (to S.I.)]; MHLW, the Japan Science and Technology Agency, and the Strategic Research Program for Brain Sciences (N. Matsumoto); and a Grant-in-Aid for Scientific Research on Innovative Areas (transcription cycle) from the Ministry of Education, Culture, Sports, Science and Technology of Japan and the Takeda Science Foundation (to N. Matsumoto).

1. Maguire AM, et al. (2008) Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med* 358(21):2240–2248.
2. Bainbridge JW, et al. (2008) Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med* 358(21):2231–2239.
3. Cideciyan AV, et al. (2008) Human gene therapy for RPE65 isomerase deficiency activates the retinoid cycle of vision but with slow rod kinetics. *Proc Natl Acad Sci USA* 105(39):15112–15117.
4. Berson EL (1993) Retinitis pigmentosa. The Friedenwald Lecture. *Invest Ophthalmol Vis Sci* 34(5):1659–1676.
5. Berson EL, Rosner B, Sandberg MA, Weigel-DiFranco C, Willett WC (2012) ω-3 intake and visual acuity in patients with retinitis pigmentosa receiving vitamin A. *Arch Ophthalmol* 130(6):707–711.
6. Berson EL, et al. (1993) A randomized trial of vitamin A and vitamin E supplementation for retinitis pigmentosa. *Arch Ophthalmol* 111(6):761–772.
7. Hartong DT, Berson EL, Dryja TP (2006) Retinitis pigmentosa. *Lancet* 368(9549): 1795–1809.
8. Jin ZB, et al. (2008) Identifying pathogenic genetic background of simplex or multiplex retinitis pigmentosa patients: A large scale mutation screening study. *J Med Genet* 45(7):465–472.
9. Tucker T, Marra M, Friedman JM (2009) Massively parallel sequencing: The next big thing in genetic medicine. *Am J Hum Genet* 85(2):142–154.
10. Tuson M, Marfany G, Gonzàlez-Duarte R (2004) Mutation of CERKL, a novel human ceramide kinase gene, causes autosomal recessive retinitis pigmentosa (RP26). *Am J Hum Genet* 74(1):128–138.
11. Aldahmesh MA, et al. (2009) Molecular characterization of retinitis pigmentosa in Saudi Arabia. *Mol Vis* 15:2464–2469.
12. Vervoort R, et al. (2000) Mutational hot spot within a new RPGR exon in X-linked retinitis pigmentosa. *Nat Genet* 25(4):462–466.
13. Sanuki R, et al. (2011) miR-124a is required for hippocampal axogenesis and retinal cone survival through Lhx2 suppression. *Nat Neurosci* 14(9):1125–1134.
14. Cottet S, Schorderet DF (2009) Mechanisms of apoptosis in retinitis pigmentosa. *Curr Mol Med* 9(3):375–383.
15. Rabbani B, Mahdieh N, Hosomichi K, Nakaoka H, Inoue I (2012) Next-generation sequencing: Impact of exome sequencing in characterizing Mendelian disorders. *J Hum Genet* 57(10):621–632.
16. Hosono K, et al. (2012) Two novel mutations in the EYS gene are possible major causes of autosomal recessive retinitis pigmentosa in the Japanese population. *PLoS ONE* 7(2):e31036.
17. Pieras JI, et al. (2011) Copy-number variations in EYS: A significant event in the appearance of arRP. *Invest Ophthalmol Vis Sci* 52(8):5625–5631.
18. van Wijk E, et al. (2006) The DFNB31 gene product whirlin connects to the Usher protein network in the cochlea and retina by direct association with USH2A and VLGR1. *Hum Mol Genet* 15(5):751–765.
19. Rivolta C, Sweklo EA, Berson EL, Dryja TP (2000) Missense mutation in the USH2A gene: Association with recessive retinitis pigmentosa without hearing loss. *Am J Hum Genet* 66(6):1975–1978.
20. Ebermann I, et al. (2007) A novel gene for Usher syndrome type 2: Mutations in the long isoform of whirlin are associated with retinitis pigmentosa and sensorineural hearing loss. *Hum Genet* 121(2):203–211.
21. Yang J, et al. (2010) Ablation of whirlin long isoform disrupts the USH2 protein complex and causes vision and hearing loss. *PLoS Genet* 6(5):e1000955.
22. McLaughlin ME, Sandberg MA, Berson EL, Dryja TP (1993) Recessive mutations in the gene encoding the beta-subunit of rod phosphodiesterase in patients with retinitis pigmentosa. *Nat Genet* 4(2):130–134.
23. Rivolta C, Sharon D, DeAngelis MM, Dryja TP (2002) Retinitis pigmentosa and allied diseases: Numerous diseases, genes, and inheritance patterns. *Hum Mol Genet* 11(10): 1219–1227.
24. Nishiguchi KM, Rivolta C (2012) Genes associated with retinitis pigmentosa and allied diseases are frequently mutated in the general population. *PLoS ONE* 7(7):e41902.
25. Fry AM, Meraldi P, Nigg EA (1998) A centrosomal function for the human Nek2 protein kinase, a member of the NIMA family of cell cycle regulators. *EMBO J* 17(2): 470–481.
26. Quarmby LM, Mahjoub MR (2005) Caught Nek-ing: Cilia and centrioles. *J Cell Sci* 118(Pt 22):5161–5169.
27. Fakhro KA, et al. (2011) Rare copy number variations in congenital heart disease patients identify unique genes in left-right patterning. *Proc Natl Acad Sci USA* 108(7): 2915–2920.
28. Bahe S, Stierhof YD, Wilkinson CJ, Leiss F, Nigg EA (2005) Rootletin forms centriole-associated filaments and functions in centrosome cohesion. *J Cell Biol* 171(1):27–33.
29. Yang J, et al. (2002) Rootletin, a novel coiled-coil protein, is a structural component of the ciliary rootlet. *J Cell Biol* 159(3):431–440.
30. Hollingsworth TJ, Gross AK (2012) Defective trafficking of rhodopsin and its role in retinal degenerations. *Int Rev Cell Mol Biol* 293:1–44.
31. Luo N, Lu J, Sun Y (2012) Evidence of a role of inositol polyphosphate 5-phosphatase INPP5E in cilia formation in zebrafish. *Vision Res* 75:98–107.
32. Patil SB, Hurd TW, Ghosh AK, Murga-Zamalloa CA, Khanna H (2011) Functional analysis of retinitis pigmentosa 2 (RP2) protein reveals variable pathogenic potential of disease-associated missense variants. *PLoS ONE* 6(6):e21379.
33. Chang GQ, Hao Y, Wong F (1993) Apoptosis: Final common pathway of photoreceptor death in rd, rds, and rhodopsin mutant mice. *Neuron* 11(4):595–605.
34. Davis EE, Katsanis N (2012) The ciliopathies: A transitional model into systems biology of human genetic disease. *Curr Opin Genet Dev* 22(3):290–303.
35. Fahim AT, et al. (2011) Allelic heterogeneity and genetic modifier loci contribute to clinical variation in males with X-linked retinitis pigmentosa due to RPGR mutations. *PLoS ONE* 6(8):e23021.

ARTICLE

Mutations in *KLHL40* Are a Frequent Cause of Severe Autosomal-Recessive Nemaline Myopathy

Gianina Ravenscroft,[1,28] Satoko Miyatake,[2,28] Vilma-Lotta Lehtokari,[3,4] Emily J. Todd,[1] Pauliina Vornanen,[3,4] Kyle S. Yau,[1] Yukiko K. Hayashi,[5] Noriko Miyake,[2] Yoshinori Tsurusaki,[2] Hiroshi Doi,[2] Hirotomo Saitsu,[2] Hitoshi Osaka,[6] Sumimasa Yamashita,[6] Takashi Ohya,[7] Yuko Sakamoto,[7] Eriko Koshimizu,[2] Shintaro Imamura,[8] Michiaki Yamashita,[8] Kazuhiro Ogata,[9] Masaaki Shiina,[9] Robert J. Bryson-Richardson,[10] Raquel Vaz,[10] Ozge Ceyhan,[11] Catherine A. Brownstein,[11] Lindsay C. Swanson,[11] Sophie Monnot,[12] Norma B. Romero,[13] Helge Amthor,[12] Nina Kresoje,[14] Padma Sivadorai,[15] Cathy Kiraly-Borri,[16] Goknur Haliloglu,[17] Beril Talim,[17] Diclehan Orhan,[18] Gulsev Kale,[18] Adrian K. Charles,[19] Victoria A. Fabian,[15] Mark R. Davis,[15] Martin Lammens,[20] Caroline A. Sewry,[21] Adnan Manzur,[21] Francesco Muntoni,[21] Nigel F. Clarke,[22] Kathryn N. North,[23] Enrico Bertini,[24] Yoram Nevo,[25] Ekkhard Willichowski,[26] Inger E. Silberg,[27] Haluk Topaloglu,[17] Alan H. Beggs,[11] Richard J.N. Allcock,[14] Ichizo Nishino,[5] Carina Wallgren-Pettersson,[3,4] Naomichi Matsumoto,[2,29,*] and Nigel G. Laing[1,29,*]

Nemaline myopathy (NEM) is a common congenital myopathy. At the very severe end of the NEM clinical spectrum are genetically unresolved cases of autosomal-recessive fetal akinesia sequence. We studied a multinational cohort of 143 severe-NEM-affected families lacking genetic diagnosis. We performed whole-exome sequencing of six families and targeted gene sequencing of additional families. We identified 19 mutations in *KLHL40* (kelch-like family member 40) in 28 apparently unrelated NEM kindreds of various ethnicities. Accounting for up to 28% of the tested individuals in the Japanese cohort, *KLHL40* mutations were found to be the most common cause of this severe form of NEM. Clinical features of affected individuals were severe and distinctive and included fetal akinesia or hypokinesia and contractures, fractures, respiratory failure, and swallowing difficulties at birth. Molecular modeling suggested that the missense substitutions would destabilize the protein. Protein studies showed that KLHL40 is a striated-muscle-specific protein that is absent in *KLHL40*-associated NEM skeletal muscle. In zebrafish, *klhl40a* and *klhl40b* expression is largely confined to the myotome and skeletal muscle, and knockdown of these isoforms results in disruption of muscle structure and loss of movement. We identified *KLHL40* mutations as a frequent cause of severe autosomal-recessive NEM and showed that it plays a key role in muscle development and function. Screening of *KLHL40* should be a priority in individuals who are affected by autosomal-recessive NEM and who present with prenatal symptoms and/or contractures and in all Japanese individuals with severe NEM.

[1]Western Australian Institute for Medical Research and the Centre for Medical Research, University of Western Australia, Nedlands, Western Australia 6009, Australia; [2]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama 236-0004, Japan; [3]The Folkhälsan Institute of Genetics, Samfundet Folkhälsan, Biomedicum Helsinki, PB 63 (Haartmaninkatu 8), University of Helsinki, Helsinki 00014, Finland; [4]Department of Medical Genetics, Haartman Institute, University of Helsinki, Helsinki 00014, Finland; [5]Department of Neuromuscular Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo 187-8502, Japan; [6]Division of Neurology, Clinical Research Institute, Kanagawa Children's Medical Center, Yokohama 232-8555, Japan; [7]Department of Pediatrics, Odawara Municipal Hospital, Odawara 232-8558, Japan; [8]National Research Institute of Fisheries Science, Yokohama 236-8648, Japan; [9]Department of Biochemistry, Yokohama City University Graduate School of Medicine, Yokohama 236-0004, Japan; [10]School of Biological Sciences, Monash University, Victoria 3800, Australia; [11]The Manton Center for Orphan Disease Research, Genetics and Genomics, Boston Children's Hospital and Harvard Medical School, Boston MA, 02115, USA; [12]Service Génétique Médicale, Hôpital Necker-Enfants Malades, Université Paris Descartes, Paris 75015, France; [13]Unité de Morphologie Neuromusculaire, Institut de Myologie, Institut National de la Santé et de la Recherche Médicale, Paris 75651, France; [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]Genetic Services of Western Australia, Princess Margaret Hospital for Children and King Edward Memorial Hospital for Women, Subiaco, Western Australia 6008, Australia; [17]Department of Pediatric Neurology, Hacettepe University Children's Hospital, Ankara 06100, Turkey; [18]Department of Pediatric Pathology, Hacettepe University Children's Hospital, Ankara 06100, Turkey; [19]School of Women's and Infants' Health, University of Western Australia, Crawley, Western Australia 6009, Australia; [20]Department of Pathology, University Hospital Antwerp, Antwerp 2650, Belgium; [21]Dubowitz Neuromuscular Centre, UCL Institute of Child Health and Great Ormond Street Hospital for Children, London WC1N 1EH, UK; [22]Institute for Neuroscience and Muscle Research, Children's Hospital at Westmead, Sydney 2145, Australia; [23]Murdoch Childrens Research Institute, The Royal Children's Hospital, Parkville, Victoria 3052, Australia; [24]Unit of Neuromuscular Disorders, Bambino Gesù Children's Hospital, Rome, Lazio 00165, Italy; [25]Neuropediatric Unit, Hadassah Medical Center, Hebrew University of Jerusalem, Jerusalem 91240, Israel; [26]Department of Pediatrics and Pediatric Neurology, University Medicine Göttingen, Göttingen 37075, Germany; [27]Neonatal Intensive Care Unit, Department of Pediatric Research, Women and Children's Division, Oslo University Hospital Rikshospitalet, Oslo 0424, Norway
[28]These authors contributed equally to this work
[29]These authors contributed equally to this work
*Correspondence: nigel.laing@uwa.edu.au (N.G.L.), naomat@yokohama-cu.ac.jp (N.M.)
http://dx.doi.org/10.1016/j.ajhg.2013.05.004. ©2013 by The American Society of Human Genetics. All rights reserved.

CrossMark

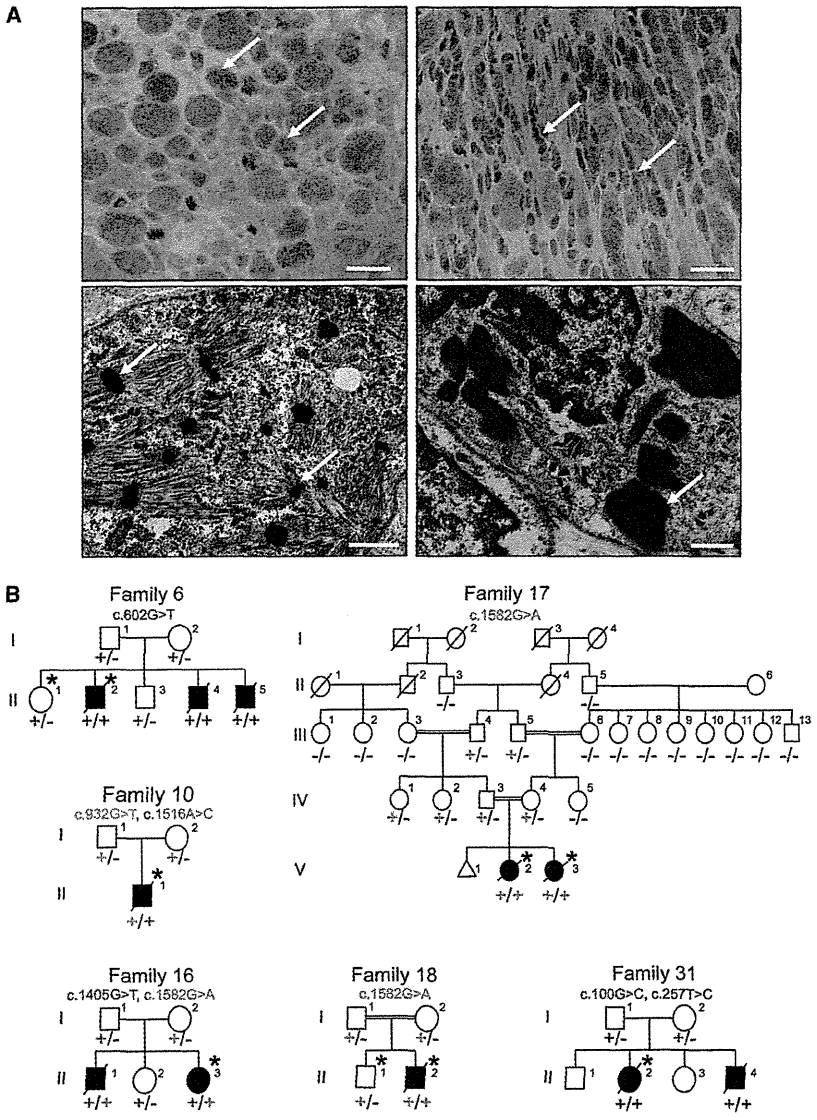

Figure 1. Family Pedigrees and Light and Electron Microscopy of Muscle Biopsies

(A) Modified Gomori trichrome (upper) and electron microscopy (lower) of muscle biopsies from affected individuals of families 15 (right) and 20 (left). Abnormal variation in fiber size, together with many small myofibers and sometimes increased connective tissue, and the presence of numerous red- or purple-stained nemaline bodies (arrows) can be seen (upper panels). Numerous nemaline bodies with varying sizes and shapes and a lack of normal myofibrils are visible by electron microscopy (arrows). Scale bars represent 20 μm for modified Gomori trichrome and 1 μm for electron microscopy.

(B) Pedigrees for the families in which exome sequencing and analysis were performed on the probands. Asterisks indicate the individuals whose DNA was analyzed by exome sequencing. Segregation of the mutations identified in each pedigree is shown.

Introduction

Nemaline myopathy (NEM) is a common form of nondystrophic congenital myopathy and is defined clinically by skeletal-muscle dysfunction and pathologically by the presence of nemaline bodies within myofibers.[1,2] Typical clinical symptoms include hypotonia, muscle weakness of proximal dominance, respiratory insufficiency, and feeding problems. Congenital onset is usual, but a wide variation in age of onset and disease severity is recognized. Mutations in seven genes are known to cause NEM (NEM1–NEM7).[1,2] Six of these encode sarcomere-thin-filament proteins or associated proteins: *ACTA1* (MIM 102610),[3] *CFL2* (MIM 601443),[4] *NEB* (MIM 161650),[5] *TNNT1* (MIM 191041),[6] *TPM2* (MIM 190990),[7] and *TPM3* (MIM 191030);[8] the seventh, *KBTBD13* (kelch-repeat- and BTB-[POZ]-domain-containing 13 [MIM 613727])[9] is involved in the ubiquitin proteasome pathway.[10] Nevertheless, some forms of NEM remain genetically unsolved.

One such subtype, which has long been recognized,[11,12] has apparent autosomal-recessive inheritance and is characterized by severe weakness, in utero presentation of fetal akinesia or hypokinesia and associated abnormalities, and muscle biopsy often showing numerous small nemaline bodies, sometimes only visible by electron microscopy and frequently with virtually no normal myofibrils remaining ("miliary NEM" Figure 1A and Figure S1, available online). We aimed to identify genetic causes of these severe NEM cases by using a combination of linkage analysis, or homozygosity mapping, SNP array, and whole-exome sequencing (WES) in selected families. We have identified loss-of-function mutations in *KLHL40* as a frequent cause of severe NEM and have shown through functional studies that KLHL40 is crucial for myogenesis and skeletal-muscle maintenance.

Subjects and Methods

Subject Details and Ethics

We recruited 143 genetically unresolved severe-NEM-affected families from large congenital-myopathy cohorts in major centers around the world (Boston, Helsinki, Perth, and Tokyo). All individuals within the cohorts were diagnosed with NEM on the basis of muscle-biopsy findings.

Written informed consent was obtained for participation in this study, which was approved by the Human Research Ethics

Committee of the University of Western Australia (UWA), the ethics committee of the Children's Hospital of the University of Helsinki, Yokohama City University School of Medicine, and the Boston Children's Hospital institutional review board. The UWA Animal Ethics Committee approved animal studies.

Microscopy

Light microscopy and electron microscopy of biopsies was performed as previously described.[13]

Whole-Genome SNP Genotyping, Linkage Analysis, and WES

Genotyping was performed for families 6 and 18 with the use of the HumanOmniExpress BeadChip Kit (Illumina) and Infinium II Assay Workflow (Illumina) at the Institute for Molecular Medicine Finland (FIMM). Data were analyzed with PLINK v.1.07. Multiple large homozygous regions were identified, but none included known myopathy-associated genes. WES was performed on one healthy and one affected sibling from family 6 and the proband from family 18 with the SeqCap EZ Human Exome Library v.2.0 exome system (Nimblegen, Roche Diagnostics). Coverage depths were 31- to 62-fold. Variant quantification was performed with the FIMM Variant Calling Pipeline v.1.0 and the Integrative Genomics Viewer (IGV, Broad Institute of MIT and Harvard). All known and heterozygous SNPs were excluded. Healthy siblings' genotypes were used for the exclusion of shared homozygous variants.

Five individuals from family 16 were genotyped with the Human Mapping 10K XbaI 142 2.0 array (Affymetrix) and GeneChip Genotyping Analysis Software (Gtypev4.1). Parametric linkage analysis was performed with Allegro v.2 with a fully penetrant autosomal-recessive model. WES was performed on the proband with the use of the SureSelect Human All Exon 50 Mb Kit (Agilent Technologies) and sequenced in one lane on a GAIIx platform (Illumina) with 108 bp paired-end reads. Reads were aligned to the UCSC Genome Browser (GRCh37/hg19) with Novoalign (Novocraft Technologies). Mean coverage depth was 59-fold. Single-nucleotide variants and small indels were identified with GATK UnifiedGenotyper and filtered according to the Broad Institute's Best Practices guidelines v.3. Variants registered in dbSNP132 were filtered. The filter-passed variants were annotated with ANNOVAR. Only genes with homozygous variants or more than two variants located in the candidate linkage regions were included.

Family 17 was genotyped with the HumanCytoSNP-12 BeadChip (Illumina). MERLIN was used for performing linkage analysis on a subset of 14,514 SNPs.[14] WES was performed for the proband from family 10 and for both siblings from family 17 as described.[15] Coverage depth was 61- to 97-fold. Variants were called with LifeScope 2.5 (Life Technologies) and filtered with ANNOVAR[16] against ENCODE GENCODE v.11 (October 2011 freeze, GRCh37).[17] Two custom variant-filtering steps were used: (1) one against the 1000 Genomes database (February 2012 release) (variants with a minor allele frequency > 0.5% were excluded) and (2) one against the dbSNP135 common database.

Family 31 (BOS74) was one in a cohort of 59 NEM-affected families who underwent WES by the Intellectual and Developmental Disabilities Research Center Core Next-Gen Sequencing Facility of Boston Children's Hospital and Harvard Medical School in collaboration with Axeq Technologies, Complete Genomics, Integrated Genetics (LabCorp), and the Boston Children's Hospital Gene Partnership. Exome sequencing was performed with the Illumina HiSeq 2000 platform. Reads were mapped with the Burrows-Wheeler Aligner (v.0.5.8). SNPs and indels were called with SAMtools (v.0.1.7). Data analysis and variant calling were performed with the Broad GATK Best Practices for identification of SNPs and small indels. Annotated variants were filtered against dbSNP135, the 1000 Genomes Project database (October 2011 edition), and the National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project Exome Variant Server (EVS).

Sequencing

Bidirectional Sanger sequencing of *KLHL40* (RefSeq accession number NM_152393.2) was performed on biobanked DNA from additional probands with severe NEM and their family members in Boston, Helsinki, Perth, Yokohama, and Tokyo. Identified variants were then screened in all available family members. Primer sequences and conditions are available upon request. For detection of the c.1582G>A (p.Glu528Lys) mutation in normal Japanese controls, high-resolution melting (HRM) analysis with and without the spike-in method[18] was performed on LightCycler 480 System II (Roche Diagnostics). If samples showed any aberrant melting patterns, Sanger sequencing was performed for confirmation of the mutation.

LOD Scores

Where possible, MERLIN was used for calculating LOD scores for individual families.[14]

Expression Analysis on Human cDNAs

TaqMan quantitative real-time PCR analyses were performed with cDNAs of human adult (Human MTCPanel I, #636742, Clontech Laboratories) and fetal (Human Fetal MTC Panel, #636747, Clontech Laboratories) tissues.[19] Predesigned TaqMan probe sets for human *KLHL40* (*KBTBD5*, Hs00328078_m1, Applied Biosystems) and human β-actin (*ACTB*, 4326315E, Applied Biosystems) were used. PCR was performed on a Rotor-Gene Q (QIAGEN) (conditions are available upon request) and analyzed with the Rotor-Gene Q Series Software by the $2^{-\Delta\Delta Ct}$ method. Relative concentrations of cDNA were normalized to concentrations obtained from the hearts.

Calculations of the Free-Energy Change upon Amino Acid Substitutions

Molecular structures were drawn with PyMOL. FoldX v.3.0 beta[20] was used through a graphics interface as a plugin for the YASARA molecular viewer.[21] Crystal structures of the kelch domain of human KLHL40 (Protein Data Bank [PDB] code 4ASC) and the BTB (bric-a-brac, tram-track, broad-complex)-BACK (BTB and C-terminal kelch) domain of human KHLH11 (PDB code 3I3N) were energy-minimized with the RepairPDB command implemented in FoldX and subsequently with the BuildModel command for mutagenesis. Protein stabilities were calculated by the Stability command, and the free-energy changes were estimated by subtraction of the free-energy value of the wild-type protein from those of the altered proteins. The procedure was repeated three times for each substitution, and the resultant data were presented as an average value with SDs.

Immunoblotting and Immunohistochemistry

SDS-PAGE and immunoblotting were performed as described.[22,23] For protein studies, C2C12 myoblasts and myotubes were grown and prepared for immunoblotting and immunofluorescence as

described.[23] For KLHL40 immunoblots, the Human Protein Atlas (HPA) rabbit polyclonal KLHL40 (KBTBD5) antibody from Sigma was used (HPA024463 [1:2,500 dilution]). Immunostaining of human and mouse muscle samples was performed as described[13,23] with a KLHL40 antibody (KBTBD5; HPA024463 [1:100 dilution]).

Zebrafish Studies

In Situ Hybridization

Digoxigenin probes for *klhl40a* and *klhl40b* were generated by cDNA amplification of 1,340 and 694 bp sequences, respectively (Table S1). In situ hybridizations were performed as described previously.[24]

Morpholino Microinjection

Antisense translation-blocking morpholinos (Table S1) for *klhl40a* (*klhl40a*-MO) and *klhl40b* (*klhl40b*-MO and *klhl40b*-MO2) were coinjected into 1- to 2-cell-stage embryos at a final concentration of 0.25 or 0.5 mM. Morpholino efficacies were tested by immunoblotting for Klhl40.

Zebrafish Immunohistochemistry

Immunohistochemistry of zebrafish embryos was performed as described[24,25] with myosin heavy chain (MHC) antibody (F59 [1:20 dilution] or A4.1025 [1:10 dilution]; Developmental Studies Hybridoma Bank) and α-actinin (1:100 dilution; Sigma) and filamin C (1:100 dilution; Sigma) antibodies, and Alexa-Fluor-488-conjugated phalloidin (1:100 dilution; Molecular Probes) was used for labeling F-actin. Immunoreactivity was detected with an Alexa-Fluor-594-conjugated anti-mouse secondary antibody diluted in blocking buffer (1:200).

Statistical Analyses

Statistical analyses of clinical features were carried out with SPSS Statistics 19 (IBM) software. Individuals for whom information for a clinical feature was not available were excluded from the analysis of that feature. Either Chi-square tests or Fisher's exact tests were applied for comparing each phenotypic variable between different genotypes. $p < 0.05$ was considered statistically significant.

Results

WES identified homozygous or compound-heterozygous mutations in *KLHL40* (kelch-like family member 40; also known as *KBTBD5* [kelch-repeat- and BTB-(POZ)-domain-containing 5] and *SYRP* [sarcosynapsin]) in six NEM-affected families (families 6, 10, 16–18, and 31; Figure 1B and Table 1). Subsequent screening of *KLHL40* by Sanger sequencing in additional probands with severe NEM resulted in the identification of a total of 19 variants (4 frameshifts, 12 missense mutations, 2 nonsense mutations, and 1 splice site) in 28 (19.6%) apparently unrelated families (Table 1) from the cohort of 143 families affected by severe NEM. In addition, 129 probands with milder NEM were screened, but no *KLHL40* mutations were identified in this cohort, confirming that *KLHL40* mutations are most likely exclusive to cases of severe NEM.

In all cases where it was possible to test unaffected parents, siblings, and extended family, the mutations cosegregated with disease in an autosomal-recessive fashion (Figure 1B), giving a combined LOD score of 5.66 (Table 1). All mutations were either absent from the NHLBI EVS and the 1000 Genomes database[26] or present at low frequencies in the heterozygous state (Table 1). In five additional NEM-affected families, only single *KLHL40* variants were identified (Table S2); the significance of these variants in these individuals remains unclear.

In Japanese persons, *KLHL40* mutations are the most common cause of this severe form of NEM (13/47 [~28%]) as a result of a founder effect with the c.1582G>A mutation. Given that this mutation was present in Turkish, Kurdish, and Japanese families, we completed a haplotype analysis of Japanese and Turkish families (families 16 and 17) but did not identify a common haplotype between them (Figure S2). HRM with confirmatory Sanger sequencing of 510 normal Japanese individuals revealed a heterozygous c.1582G>A mutation in one individual. Therefore, the mutant-allele frequency in the Japanese population was estimated to be 0.0098. According to the equation described by Kimura and Ota[27] and under the assumption of 25 years per generation, the age of this mutation is calculated to be 4,900 years old.

The identified *KLHL40* mutations were scattered throughout all exons (Table 1 and Figure 2A) encoding mostly conserved residues (Figure S3). To investigate disease mechanisms, all substitutions except p.Arg311Leu were mapped to the crystal structures of the kelch domain of human KLHL40 and the BTB-BACK domain of human kelch-like protein 11 (KLHL11; Figures 2B and 2C and Figure S4). p.Arg311Leu (c.932G>T) was predicted to be in the structurally flexible region, a linker of nonconserved amino acids connecting the BACK and kelch domains (Figure S7D), and was therefore excluded from structural consideration. All the modeled substituted residues are involved in intramolecular interactions, and thus the substitutions would most likely destabilize the hydrophobic cores of the BTB-BACK domain (p.Leu86Pro [c.257T>C], p.Val194Glu [c.581T>A], and p.Trp201Leu [c.602G>A]), the kelch domain (p.Pro397Leu [c.1190C>T], p.His455Arg [c.1364A>G], and p.Gly469Cys [c.1405G>T]), the β sheet (p.Thr506Pro [c.1516A>C] and p.Ala538Pro [c.1612G>C]), or the hydrogen bonds between the main chain and side chain (p.Asp34His [c.100G>C] and p.Glu528Lys [c.1582G>A]) or between side chains (p.Glu588Lys [c.1762G>A]) (Figures S5–S7). The p.Pro397Leu and p.Glu588Lys substitutions appear to be conservative for the hydrophobic core and hydrogen bonding, respectively. The former substitution is predicted to affect the polyproline II helix conformation (residues 396–399; Figure S6A). The calculated free-energy change for most substitutions was estimated to be over 2.0 kcal/mol (Figure 2D), which is typically associated with destabilization of domain folds.[28] These analyses suggested that most *KLHL40* missense mutations impair protein stability.

To investigate *KLHL40* expression and KLHL40 abundance, we performed quantitative RT-PCR and immunoblotting of human and mouse tissues. *KLHL40* transcripts

Table 1. *KLHL40* Mutations by Family, Individual LOD Scores, Ethnicity, and Population-wide Incidence

Family	Exon(s)	Mutation		LOD Score	Ethnicity	Incidence from EVS (1^{st}; 2^{nd})	Incidence from 1000 Genomes (1^{st}; 2^{nd})
		Nucleotide Change	Amino Acid Change				
Family 31[a]	1	c.[100G>C];[257T>C]	p.[Asp34His];[Leu86Pro]	0.6	Vietnamese	ND; ND	ND; ND
Family 2	1	c.[134delC];[134delC]	p.[Pro45Argfs*19]; [Pro45Argfs*19]	NA	Italian	NA	ND
Family 3	1	c.[270C>G];[270C>G]	p.[Tyr90*];[Tyr90*]	NA	Turkish	ND	ND
Family 5	1	c.[581T>A];[581T>A]	p.[Val194Glu];[Val194Glu]	0.6	Israeli	ND	ND
Family 6[a]	1	c.[602G>T];[602G>T]	p.[Trp201Leu];[Trp201Leu]	1.454	Turkish	ND	ND
Family 7	1	c.[602G>A];[602G>A]	p.[Trp201*];[Trp201*]	NA	Norwegian	ND	ND
Family 9	1	c.[790delC];[790delC]	p.[Arg264Alafs*59]; [Arg264Alafs*59]	0.25	Turkish	NA	ND
Family 10[a]	1 and 4	c.[932G>T];[1516A>C]	p.[Arg311Leu];[Thr506Pro]	NA	Chinese	ND; ND	ND; ND
Family 34	2 and 6	c.[1190C>T];[1762G>A]	p.[Pro397Leu];[Glu588Lys]	NA	Turkish	ND; ND	ND; A = 2 and G = 2,184
Family 12	2 and 4	c.[1270_1272delinsAGATC AAGGT];[1582G>A]	p.[Asp424Argfs*23]; [Glu528Lys]	NA	Japanese	NA; ND	ND; ND
Family 13	2 and 4	c.[1281_1294delCTGCCTGG ACTCGG];[1582G>A]	p.[Cys428Hisfs*12]; [Glu528Lys]	NA	Korean	NA; ND	ND; ND
Family 14	3	c.[1364A>G];[1364A>G]	p.[His455Arg];[His455Arg]	NA	Turkish	ND	ND
Family 15	3	c.[1405G>T];[1405G>T]	p.[Gly469Cys];[Gly469Cys]	NA	Japanese	ND	ND
Family 16[a]	3 and 4	c.[1405G>T];[1582G>A]	p.[Gly469Cys];[Glu528Lys]	0.727	Japanese	ND; ND	ND; ND
Family 17[a]	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	1.654	Turkish	ND	ND
Family 18[a]	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	0.125	Kurdish	ND	ND
Family 19	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	0.25	Kurdish	ND	ND
Family 20	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	NA	Japanese	ND	ND
Family 21	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	NA	Japanese	ND	ND
Family 22	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	NA	Japanese	ND	ND
Family 23	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	NA	Japanese	ND	ND
Family 24	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	NA	Japanese	ND	ND
Family 25	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	NA	Japanese	ND	ND
Family 26	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	NA	Japanese	ND	ND
Family 27	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	NA	Japanese	ND	ND
Family 28	4	c.[1582G>A];[1582G>A]	p.[Glu528Lys];[Glu528Lys]	NA	Japanese	ND	ND
Family 29	4/5	c.[1608–1G>A];[1608–1G>A]	NA	NA	Turkish	ND	ND
Family 30	5	c.[1612G>C];[1612G>C]	p.[Ala538Pro];[Ala538Pro]	NA	Turkish	ND	ND

The individual pedigree LOD scores are given where possible. This table also shows the incidence of the mutations reported within the NHLBI EVS and the 1000 Genomes browser. Abbreviations are as follows: NA, not available; and ND, not detected.
[a]Families for whom WES was performed.

and their encoded proteins were exclusive to developing and adult skeletal muscle (Figures 3A–3C) and more abundant in fetal muscle than in postnatal muscle (Figure 3C). Confocal microscopy suggested that KLHL40 might localize to the sarcomeric A-band (Figure 3D and Figure S8), a region not previously linked to NEM. Immunoblotting showed that KLHL40 is absent or of low abundance in *KLHL40*-associated NEM muscle (Figure 3E), even for persons harboring two missense mutations (F10 and F17). Immunohistochemistry confirmed that KLHL40 was absent or very scarce in *KLHL40*-associated NEM myofibers (Figure 3F).

We further investigated Klhl40 function in zebrafish. The zebrafish genome contains two orthologs of *KLHL40*: *klhl40a* and *klhl40b*, which have 57% (*klhl40a*) and 55.7% (*klhl40b*) amino acid similarity to human KLHL40. RT-PCR demonstrated expression of both *klhl40* genes at 24 and 48 hr postfertilization (hpf) (Figure S9A). In adult

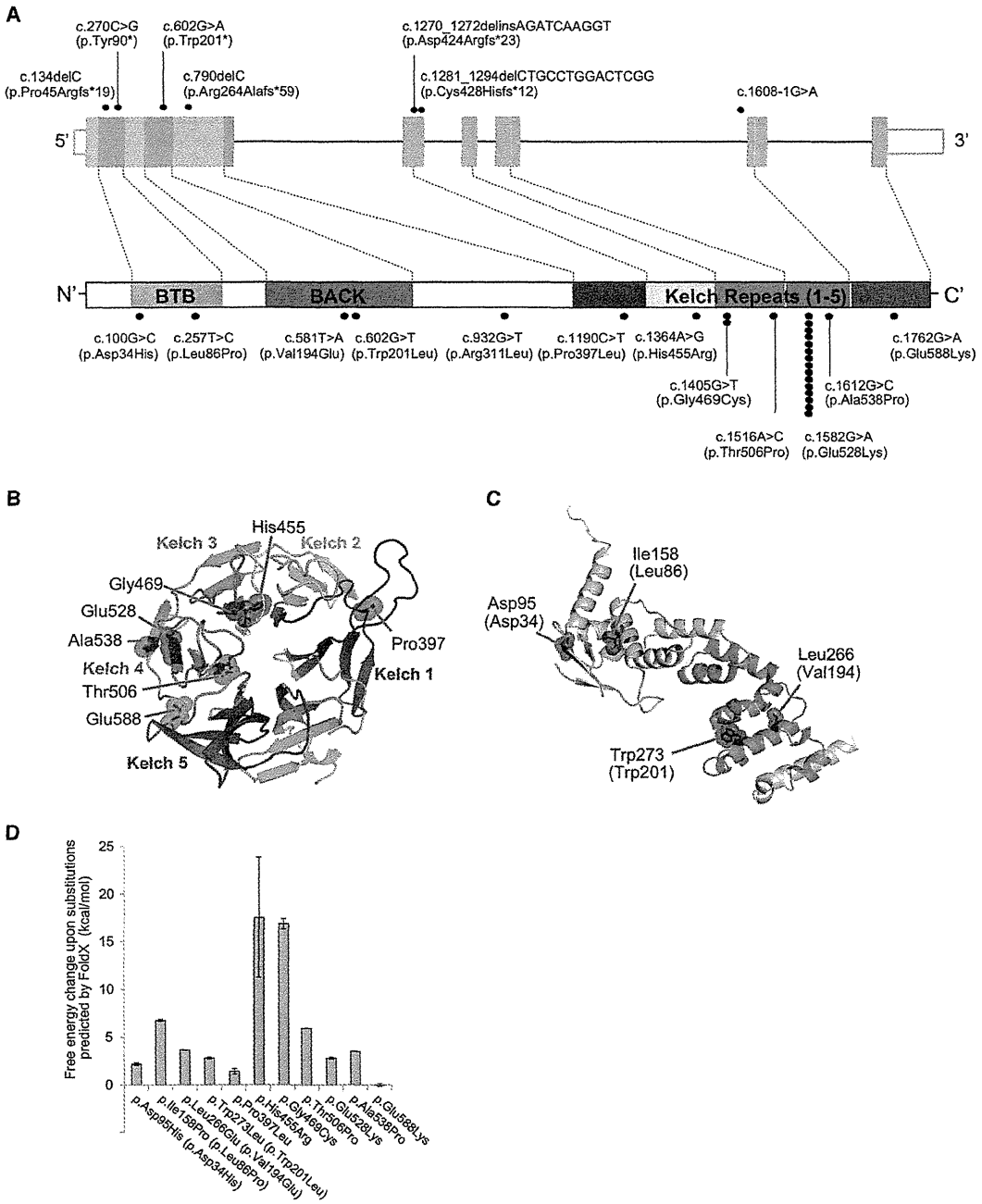

Figure 2. Mutations Identified in Our Cohort and the Structural Modeling of the Missense KLHL40 Substitutions

(A) Schematic presentation of the genomic structure of *KLHL40* (upper) and its encoded protein, KLHL40, with the BTB-BACK domain and kelch repeats (lower). The localization of mutations and substitutions identified is depicted with dots, and the number of dots for each mutation or substitution indicates the number of times it was found. Most substitutions occurred at conserved amino acids. The dots above *KLHL40* indicate truncating mutations, and those below *KLHL40* indicate missense mutations.

(B and C) Structural modeling of the missense KLHL40 substitutions. The crystal structures of the (B) kelch domain of KLHL40 and the (C) BTB-BACK domain of KLHL11 and the location of the substitutions are shown. p.Pro397Leu, p.His455Arg, p.Glu469Cys, p.Thr506Pro, p.Glu528Lys, p.Ala538Pro, and p.Glu588Lys map to the kelch repeats (B), p.Asp34His and p.Leu86Pro map to the BTB domain, and p.Val194Lys and p.Trp201Leu map to the BACK domain (C). The side chains of the mutated residues are shown as sticks with space-filling spheres in red. α helices, β sheets, and loops are drawn as ribbons, arrows, and threads, respectively. Each kelch repeat (B) is color coded in the kelch domain, and the BTB and BACK domains (C) are colored pink and green, respectively. Molecular structures were drawn with PyMOL.

(D) The calculated free-energy changes resulting from the missense substitutions in the kelch domain of human KLHL40 and the BTB-BACK domain of human KLHL11 were predicted by FoldX. Data are presented as the mean ± SD. Residue numbers used in (C) and (D) refer to human KLHL11, and those corresponding to human KLHL40 are in parentheses.

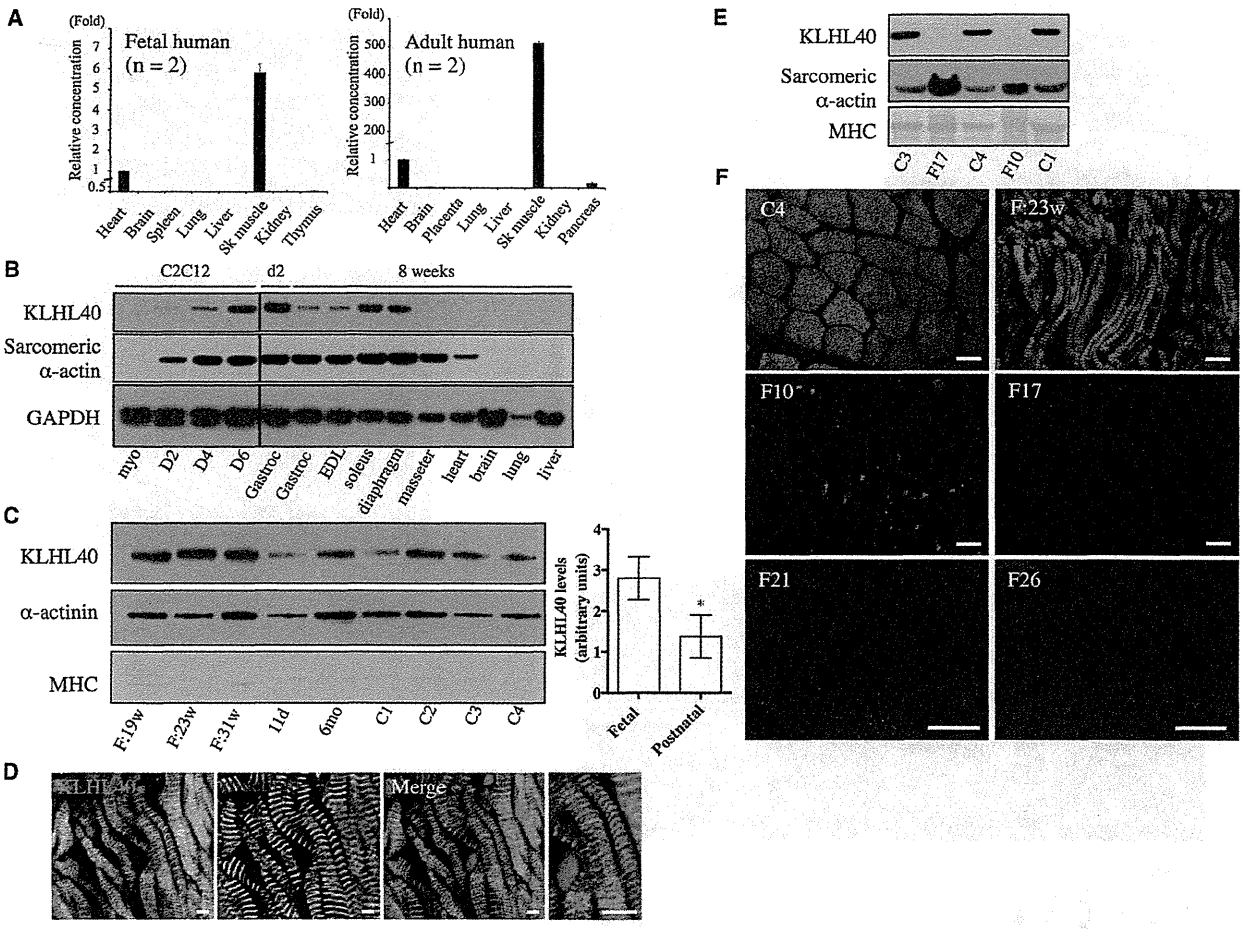

Figure 3. ***KLHL40*** **Expression in Human and Mouse Tissues**

(A) Taqman quantitative real-time PCR analysis of cDNA from adult or fetal human tissues. Error bars represent the SD. The following abbreviation is used: Sk, skeletal.

(B) KLHL40 levels in C2C12 cells and mouse tissues (HPA, top panel) and immunoblotting for sarcomeric α-actin (clone 5C5, middle panel) and GAPDH (lower panel). Lanes are as follows: myo, C212 myoblasts; D2, myotubes on day 2 of differentiation; D4, myotubes on day 4 of differentiation; D6, myotubes on day 6 of differentiation; Gastroc (left), C57BL/6 postnatal day 2 (d2) gastrocnemius; Gastroc (right), C57BL/6 8-week-old gastrocnemius; and EDL (extensor digitorum longus) to liver, C57BL/6 8-week-old tissues. For all mouse tissue lysates, samples were pooled from three different mice.

(C) On the left is KLHL40 expression in human skeletal muscle (HPA, top panel), immunoblotting for α-actinin (clone EA-53, middle panel), and Coomassie staining of MHC band (bottom panel). Lanes are as follows: F:19w, 19-week-old fetus; F:23w, 23-week-old fetus; F:31w, 31-week-old fetus; 11d, 11-day-old neonate; 6mo, 6-month-old baby; and C1–C4, healthy adult controls of 19–42 years of age. On the right, KLHL40 intensity normalized to MHC for fetal muscle is 3.34 ± 0.92 (n = 3) versus 1.37 ± 0.21 (n = 6) for postnatal skeletal muscle. *p = 0.023, unpaired two-tailed t test. Error bars represent the SEM.

(D) Single Z-plane confocal microscopy showing localization of KLHL40 (green) and α-actinin (red) in a longitudinal section of skeletal muscle from a 31-week-old fetus; costaining with Hoechst (blue) is also shown (Merge). Scale bars represent 5 μm.

(E) Immunoblotting shows that KLHL40 is absent in *KLHL40*-associated NEM muscle (II-1 from family 10 [F10] and V-2 from family 17 [F17]) compared with healthy control muscle (C1, C3, and C4). Coomassie staining of the MHC band (bottom panel) and immunoblotting for sarcomeric α-actin (clone 5C5, middle panel) indicate similar or greater loading for the *KLHL40*-associated NEM samples compared with control samples.

(F) Immunofluorescence for KLHL40 in a human 23-week-old fetal skeletal muscle sample (F:23w), an adult healthy control (C4), and *KLHL40*-associated NEM muscle biopsies (II-1 from family 10 [F10], V-2 from family 17 [F17], family 21 [F21], and family 26 [F26]). Scale bars represent 50 μm.

zebrafish, *klhl40a* was most abundant in the skeletal muscle and heart and *klhl40b* was most abundant in the skeletal muscle (Figure S9A). At the 16 and 24 hpf stages, expression of both genes was restricted to the muscle precursor cells in the somites (Figure 4A). We knocked down zebrafish *klhl40a* and *klhl40b* with antisense morpholino oligonucleotides (*klhl40a*-MO, *klhl40b*-MO, and *klhl40b*-MO2) (Figures S9B and S10A). Embryos injected with *klhl40a*-MO, *klhl40b*-MO, and *klhl40a*-MO/*40b*-MO (double morpholinos) showed a curved trunk and small head at 48 hpf (Figures 4B and 4C). A normal phenotype resulted from 5 bp mismatched morpholinos (5mis-MOs).

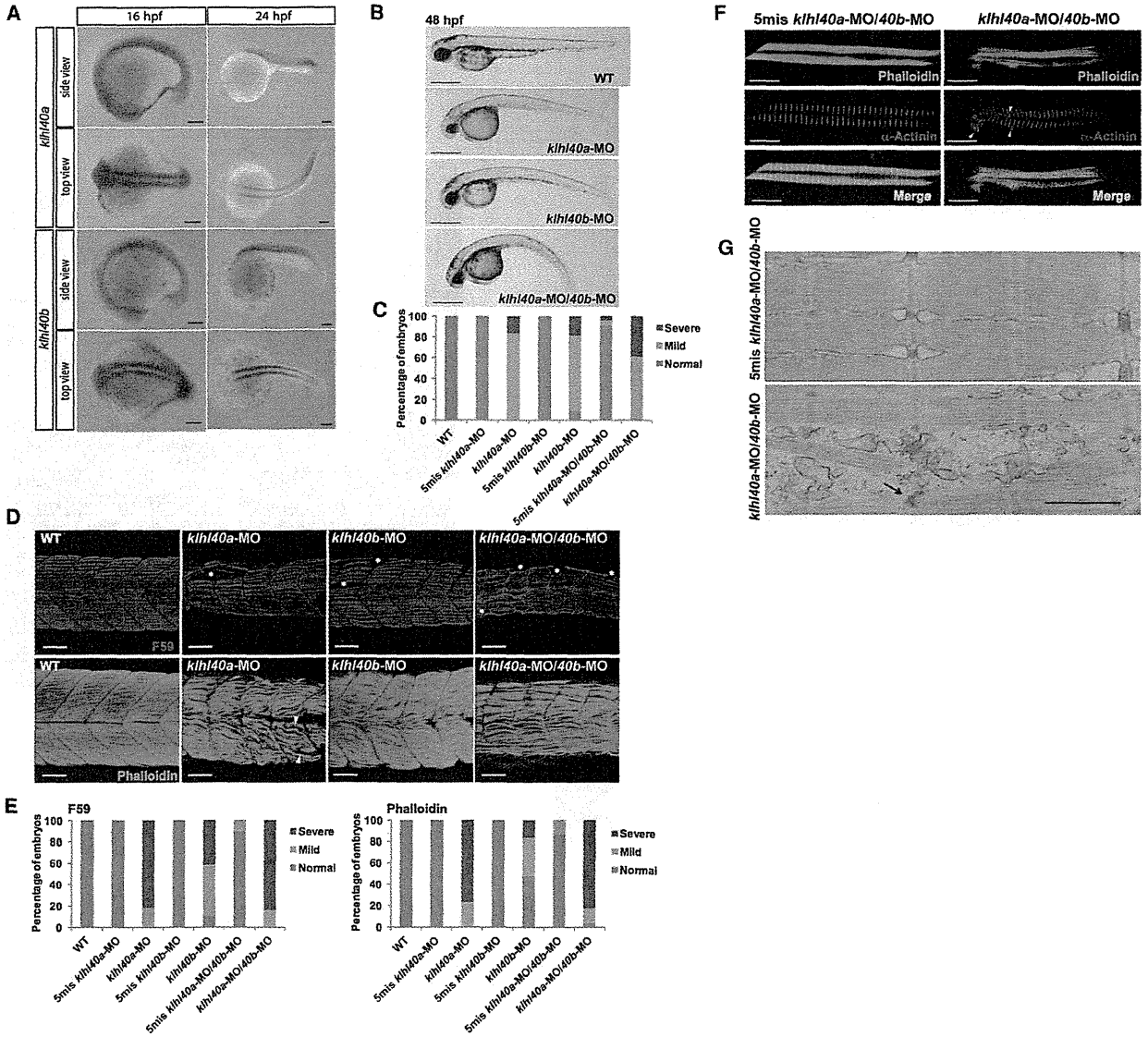

Figure 4. Expression and Function of *klhl40* in Zebrafish

(A) In situ hybridization demonstrates that expression of both *klhl40a* and *klhl40b* is restricted to the skeletal muscle at 16 and 24 hpf.

(B) Gross morphology of uninjected embryos (WT) and embryos injected with *klhl40a*-MO, *klhl40b*-MO, and *klhl40a*-MO/*40b*-MO. Lateral views of MO-injected embryos (4 ng) at 48 hpf are shown. Scale bars represent 500 μm.

(C) Percentage of embryos categorized in phenotypic classes after injection with the 5mis-MO control, *klhl40a*-MO, *klhl40b*-MO, or *klhl40a*-MO/*40b*-MO. We categorized the phenotypes at 48 hpf into normal (normal appearance), mild (curved trunk), and severe (tail defect and severe development delay) (n = 111–130).

(D) Knockdown of *klhl40a*, *klhl40b*, or both resulted in severe disruption of the skeletal muscle: fibers appeared wavy, and there were extensive gaps between fibers in contrast to the densely packed and aligned fibers of the controls. Maximum-intensity projection images from a confocal image series followed immunolabeling with a myosin antibody (F59, upper panels) at 36 hpf and F-actin (lower panels) at 72 hpf.

(E) Embryos injected with 5mis-MO, *klhl40a*-MO, *klhl40b*-MO, or *klhl40a*-MO/*40b*-MO were categorized phenotypically on the basis of the presence of myofiber detachment affecting one to two somites (mild) or multiple (three or more) somites (severe) (n = 25–44).

(F) Double-labeled immunofluorescence was performed on isolated myofibers from 72 hpf embryos with the use of phalloidin (green) and α-actinin (red). Frequent areas of aberrant α-actinin accumulation were detected in *klhl40a*-MO/*40b*-MO myofibers (arrowheads).

(G) Electron microscopy of 72 hpf myofibers. A 5mis-MO-injected embryo shows correctly aligned sarcomeres and T-tubules (upper panel). A *klhl40a*-MO/*40b*-MO-injected embryo (lower panel) shows disarranged myofibrils with widened Z-disks (arrow), but thin filament lengths are unchanged. The scale bar represents 0.7 μm.

We analyzed slow myofibers in more detail by immunostaining slow myosin heavy chains (Figure 4D, upper panels). *klhl40* morphants showed disruption of muscle patterning with an irregular, wavy appearance of the striated myofibers and extensive gaps between the myofibers (Figures 4D and 4E and Figure S10B) and a greatly

Table 2. Summary of Clinical Features of NEM Individuals with *KLHL40* Mutations

	Individuals with *KLHL40* Mutations (n = 32 Cases from 28 Families)	
	Total	**Percentage**
Family history	17/28	60.7%
Consanguinity	10/28	35.7%
Prenatal Period		
Prenatal symptoms	24/29	82.8%
Fetal akinesia or hypokinesia	16/21	76.2%
Polyhydramnios	14/29	48.3%
Neonatal Period		
Respiratory function		
respiratory failure	28/29	96.6%
requiring ventilation	11/29	37.9%
Facial involvement	26/26	100%
weakness	23/23	100%
ophthalmoparesis	4/23	17.4%
mild dysmorphology	15/15	100%
Dysphagia	23/24	95.8%
with tube feeding or gastrostomy	13/24	54.2%
Muscle weakness	29/29	100.0%
with no spontaneous antigravity movements	13/29	44.8%
Contracture(s)	24/27	88.9%
Pathological fracture(s)	10/19	52.6%
Average age at death	5 months (n = 14)	
Average gestation age at birth	37 weeks (n = 27)	
Average birth weight	2,558 g (n = 26)	

Total numbers were calculated as the number of individuals with the clinical features over the total number of individuals whose medical records were available for each category.

diminished birefringence (Figure S10C). Isolated myofibers from *klhl40a*-MO/*40b*-MO fish, coimmunostained with phalloidin and an α-actinin antibody (Z-disk), showed disorganized and irregular patterns with small aggregates of α-actinin, suggesting nemaline bodies (Figure 4F). Aggregation of Z-disk material was also confirmed by immunostaining for filamin C in *klhl40a*-MO/*40b*-MO fish (Figure S11). Electron-microscopic analysis revealed disarranged myofibrils with widened Z-disks (Figure 4G). Fish injected with *klhl40a*-MO, *klhl40b*-MO, *klhl40b*-MO2, or *klhl40a*-MO/*40b*-MO2 (double morpholinos) exhibited sporadic muscle tremors, and coordinated swimming behavior was not observed (Movies S1 and S2). These results suggest that Klhl40a and Klhl40b are required for muscle development and function and that loss of either isoform in the early embryo is sufficient to impair normal mobility.

Detailed clinical records were collected and analyzed for 32 affected individuals from the 28 unrelated kindreds afflicted with *KLHL40* mutations. These individuals were from various ethnicities, such as European, Middle and Near Eastern, or Asian. Clinical features of individuals with *KLHL40* mutations were severe and distinctive (Table 2 and Table S3). Eighty-three percent of affected individuals showed prenatal symptoms, and 76% displayed fetal akinesia or hypokinesia. Most persons had severe respiratory compromise (97%), and approximately a third required ventilatory support (38%). Almost all affected individuals (96%) also had swallowing problems, and half required tube feeding or gastrostomy. Muscle weakness was severe. Forty-five percent of individuals had no spontaneous antigravity movement. Seventeen percent of affected individuals were also noted to have ophthalmoparesis, a relatively rare symptom in NEM. Multiple joint contractures and pathological bone fracture were other common features. Dysmorphic facial features and deformities of the chest, spine, fingers, and feet were also frequent. The average age of death was 5 months. Many families, including a previously described family (family 30 herein, cases 2–6 in Lammens et al.),[11] were consanguineous.

We further evaluated whether there are any genotype-phenotype correlations in *KLHL40*-associated NEM. We compared the clinical features of individuals according to the type of mutation they had (either two truncating mutations, one truncating mutation and one missense mutation, or two missense mutations) and the pattern of mutations (homozygous or compound heterozygous). No significant differences in frequencies of these clinical features were observed (data not shown). We also compared the clinical features of persons with the recurrent c.1582G>A genotype (either with this mutation [genotype G/A or A/A as group A] or without [genotype G/G as group G]). Prenatal symptoms, including fetal akinesia or hypokinesia, were frequently observed (73.3% in group A versus 92.9% in group G). Respiratory failure was common in both groups (100% in group A versus 92.9% in group G), but there were significantly fewer individuals requiring ventilation in group A than in group G (20.0% in group A versus 57.1% in group G; $p = 0.047$). Dysphagia was also common in both groups (100% in group A versus 90.0% in group G), but there were fewer persons requiring tube feeding or gastrostomy in group A than in group G, although the difference was not significant (42.9% in group A versus 70.0% in group G; $p = 0.127$). Facial weakness was observed in all affected individuals in both groups, but fewer individuals in group A had ophthalmoparesis (7.7% in group A versus 30.0% in group G; $p = 0.281$). All persons also had muscle weakness, but significantly fewer individuals in group A had the most severe form of muscle weakness with no antigravity movements (20.0% in group A versus 71.4% in group G; $p = 0.018$). Significantly fewer affected individuals in group A were deceased at the time of study than in group G (23.5% in group A versus 71.4% in group G; $p = 0.012$;

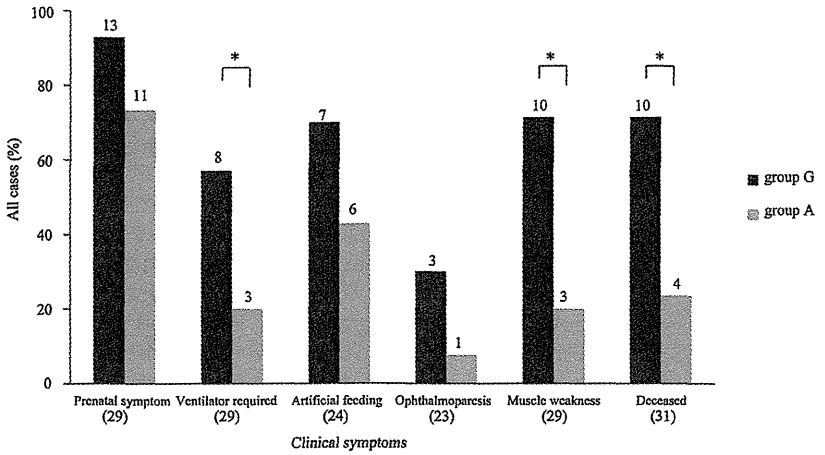

Figure 5. Correlation between the c.1582G>A (p.Glu528Lys) Mutation and Clinical Features

The clinical characteristics of NEM are shown for the two groups of affected individuals (32 total), either with the c.1582G>A (p.Glu528Lys) mutation (as group A) or without it (as group G). The numbers of total affected individuals with clinical records regarding either the presence or the absence of each characteristic are indicated below the bars, and the numbers of affected individuals in each group are indicated above the respective bars. Labels on the x axis are as follows: prenatal symptoms, individuals demonstrating either fetal akinesia or hypokinesia, polyhydramnios, or fetal edema or effusion; ventilator required, individuals with respiratory failure requiring ventilation; artificial feeding, dysphagia-affected persons requiring tube feeding or gastrostomy; ophthalmoparesis, individuals with ophthalmoparesis along with facial weakness; muscle weakness, individuals with the most severe form of muscle weakness and demonstrating no antigravitory movement; and deceased, individuals who were deceased at the time of study. Asterisks indicate that statistical significance was observed.

odds ratio = 8.125; 95% confidence interval = 1.62–40.75) (Figure 5). We further compared the clinical features of individuals of different ethnicities (either European or Asian descent) according to the c.1582G>A genotype, and similar tendencies were demonstrated (data not shown). There was, however, great variation in severity for individuals with or without the c.1582G>A genotype.

Discussion

We have described the identification of recessive *KLHL40* mutations in individuals with severe NEM from 28 unrelated families of various ethnicities. The c.1582G>A mutation was the most frequently detected mutation and was found in Japanese, Kurdish, and Turkish persons. However, comparison of haplotypes between a Japanese family and a Turkish family suggested that the mutation arose independently in these ethnic groups. We have shown several lines of evidence of the pathogenicity of the *KLHL40* mutations. The missense mutations occurred mostly in conserved functional domains within KLHL40, and they were predicted to destabilize the intramolecular interactions and thus impair protein stability. This was corroborated by the absence of KLHL40 even in the skeletal muscle of individuals harboring two missense mutations. We have established a locus-specific database for *KLHL40* mutations at the Leiden Muscular Dystrophy Pages.

Expression of KLHL40 in fetal and adult skeletal muscle indicates that KLHL40 plays a role in both myogenesis and mature muscle. KLHL40 appears to be more abundant in fetal skeletal muscle than in postnatal skeletal muscle and most likely accounts for the prevalence of in utero presentations in this NEM cohort. Perhaps KLHL40 is more important for myogenesis than for muscle maintenance; this could account for the fact that the disease ranges so much in severity, from some individuals' dying within hours of being born to others' surviving into adolescence. Our zebrafish studies have demonstrated that Klhl40a and Klhl40b are not required for the specification of muscle cells but rather for muscle patterning and function and that loss of either isoform in the early embryo is sufficient to impair normal mobility, supporting the involvement of KLHL40 in NEM-associated fetal akinesia. It has previously been suggested that KLHL40 is also important for muscle maintenance through the process of degeneration and regeneration.[29,30] *Klhl40* is upregulated in myogenic precursors after cardiotoxin injury of mouse skeletal muscle, supporting a role for *Klhl40* in the response to muscle damage.[29] Studies of cattle muscle have shown increased *Klhl40* expression in another catabolic process, undernutrition, further suggesting a role for *KLHL40* in the stress response.[30]

KLHL40 belongs to the superfamily of kelch-repeat-containing proteins that form characteristic β-propeller structures,[31] which bind substrate proteins and are involved in a wide variety of functions. In humans, 71 kelch-repeat-containing proteins have been identified.[31] The majority contain an N-terminal BTB domain (also known as the POZ [poxvirus and zinc finger] domain) and a BACK motif. Proteins containing both a BTB domain and a kelch repeat have previously been implicated in neuromuscular disease. A dominant *KLHL9* mutation causes an early-onset distal myopathy (distal myopathy 1 [MIM 160500]),[32] and dominant *KBTBD13* mutations cause nemaline myopathy with cores (MIM 609273).[9] We now show that *KLHL40*, encoding KLHL40, which contains both a BTB domain and a kelch repeat, is associated with autosomal-recessive neuromuscular disease. BTB domains function as substrate-specific adaptors for cullin 3 (Cul3),[33,34] a component of the E3-ubiquitin-ligase complex. Both KLHL9 and KBTBD13 bind Cul3.[10,32] MuRF1,

an E3-ubiquitin ligase, is known to be recruited to M-line titin and is thought to modulate myofibrillar turnover and the trophic state of muscle.[35] KLHL40 appears to be present at the A-band and might be similarly involved through the ubiquitin-proteasome pathway.

We have characterized the severe and distinctive features of this disease as fetal akinesia or hypokinesia during the prenatal period, respiratory failure and swallowing difficulty at birth, contractures and fractures along with dysmorphic features, and in most cases, early death. We have also shown that persons with the recurrent c.1582G>A mutation tend to have relatively milder symptoms compared to those of individuals without c.1582G>A. However, the severity of the disease in persons with or without the c.1582G>A genotype varied greatly (for example, from death at 20 days to still being alive at 11 years for persons homozygous for the c.1582G>A genotype), suggesting modifying factors.

Fetal akinesias are clinically and genetically heterogeneous, and the majority of cases still remain genetically unsolved.[36] Primary muscle diseases account for up to 50% of such syndromes.[37] On the basis of our study, *KLHL40* mutations cause a significant proportion of severe NEM cases of fetal akinesia sequence and the disease shows worldwide prevalence. *KLHL40* should be considered when a clinician encounters an individual presenting with prenatal symptoms, such as fetal akinesia or hypokinesia, or clinical features and/or pathology of severe NEM at birth (especially miliary NEM, which was present in at least 20% of our *KLHL40*-mutation cases), along with an autosomal-recessive pattern of family history. Fractures are a relatively frequent presentation within this cohort, unlike other NEM cohorts, and should also be used for directing genetic screening of *KLHL40*. We show that KLHL40 immunohistochemistry, immunoblotting, or genetic screening will identify the disease and thus allow genetic counseling for the affected individual's family.

In conclusion, this study associates loss-of-function *KLHL40* mutations with severe, often in utero, NEM. Many probands who do not harbor *KLHL40* mutations present with NEM in utero, suggesting further genetic heterogeneity. Clarification of KLHL40 function and interactions might lead to a greater understanding of the pathogenesis of disease, the identification of other candidates for this severe form of NEM, and the investigation of possible therapies.

Supplemental Data

Supplemental Data include 11 figures, three tables, and two movies and can be found with this article online at http://www.cell.com/AJHG.

Acknowledgments

This research was supported by the National Health and Medical Research Council of Australia (fellowships APP1035955 to G.R. and APP1002147 to N.G.L. and grant APP1022707) and Association Francaise contre les Myopathies (AFM; AFM15734). E.T. and K.S.Y. are supported by university postgraduate awards. This work received grants from the Ministry of Health, Labour, and Welfare (N. Miyake, H.S., and N. Matsumoto), Japan Science and Technology Agency (N. Matsumoto), Strategic Research Program for Brain Sciences (E.K. and N. Matsumoto), and Takeda Science Foundation (N. Miyake and N. Matsumoto) and Grants-in-Aid for Scientific Research on Innovative Areas (Transcription Cycle) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (N. Miyake and N. Matsumoto) and for Scientific Research from the Japan Society for the Promotion of Science (N. Miyake, H.S., and N. Matsumoto). The A.H.B. laboratory was supported by the National Institutes of Health (R01-AR044345) and the Muscular Dystrophy Association (MDA201302). O.C. is a Dubai Harvard Foundation for Medical Research Fellow and a grantee of the Schlumberger Foundation Faculty for the Future Program. E.B. is supported by grants from Telethon (GUP08005) and the Ministry of Health on Congenital Myopathies. F.M. is supported by the Great Ormond Street Children's Charity and National Specialist Commissioning Group. P.V. and V.-L.L. were supported by grants to C.W.-P. by the AFM, Sigrid Jusélius Foundation, Academy of Finland, Finska Läkaresällskapet, and Medicinska Understödsföreningen Liv och Hälsa r.f. R.V. is supported by a Monash Graduate Research Scholarship and a Faculty of Science Dean's International Postgraduate Research Scholarship.

Received: March 15, 2013
Revised: April 25, 2013
Accepted: May 3, 2013
Published: June 6, 2013

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes Project, http://www.1000genomes.org/
dbSNP, http://www.ncbi.nlm.nih.gov/projects/SNP/
Leiden Open Variation Database, www.LOVD.nl/KLHL40
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, http://evs.gs.washington.edu/EVS/
Online Mendelian Inheritance in Man (OMIM), http://www.omim.org
PyMOL, http://www.pymol.org
RefSeq, http://www.ncbi.nlm.nih.gov/RefSeq

References

1. Nance, J.R., Dowling, J.J., Gibbs, E.M., and Bönnemann, C.G. (2012). Congenital myopathies: an update. Curr. Neurol. Neurosci. Rep. *12*, 165–174.
2. Nowak, K.J., Davis, M.R., Wallgren-Pettersson, C., Lamont, P.J., and Laing, N.G. (2012). Clinical utility gene card for: nemaline myopathy. Eur. J. Hum. Genet. *20*. Published online April 18, 2012. http://dx.doi.org/10.1038/ejhg.2012.70.
3. 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.
4. Agrawal, P.B., Greenleaf, R.S., Tomczak, K.K., Lehtokari, V.L., Wallgren-Pettersson, C., Wallefeld, W., Laing, N.G., Darras,