Appendix

for

Shashi *et al*. "Loss of tubulin deglutamylase CCP1 causes infantile-onset neurodegeneration"

Appendix Note S1 - Members of the Undiagnosed Diseases Network.

Appendix Figure S1 - Partial deletion of the *CCP1* gene (case J1).

Appendix Figure S2 - Sanger validation and segregation analysis of *CCP1* variants.

Appendix Figure S3 - CCP1 levels in transfected cells.

Appendix Figure S4 - Axon numbers in the saphenous nerve.

Appendix Table S1 - WES and data output.

Appendix Table S2 - Frequencies of disease-related CCP1 variants in control data sets.

Appendix Table S3 - Bioinformatic results for CCP1 amino acid substitutions.

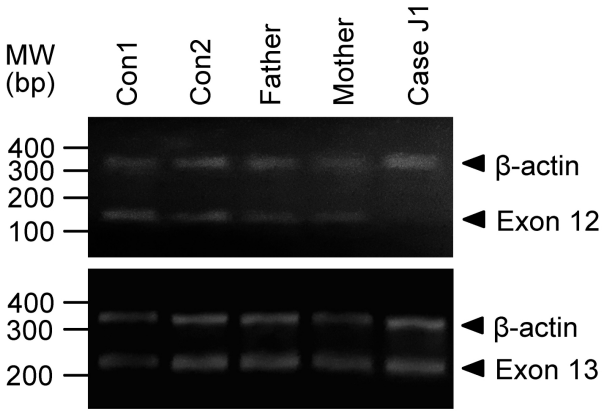
Appendix Table S4 - Clinical findings in 13 individuals with biallelic *CCP1* variants.

Appendix References

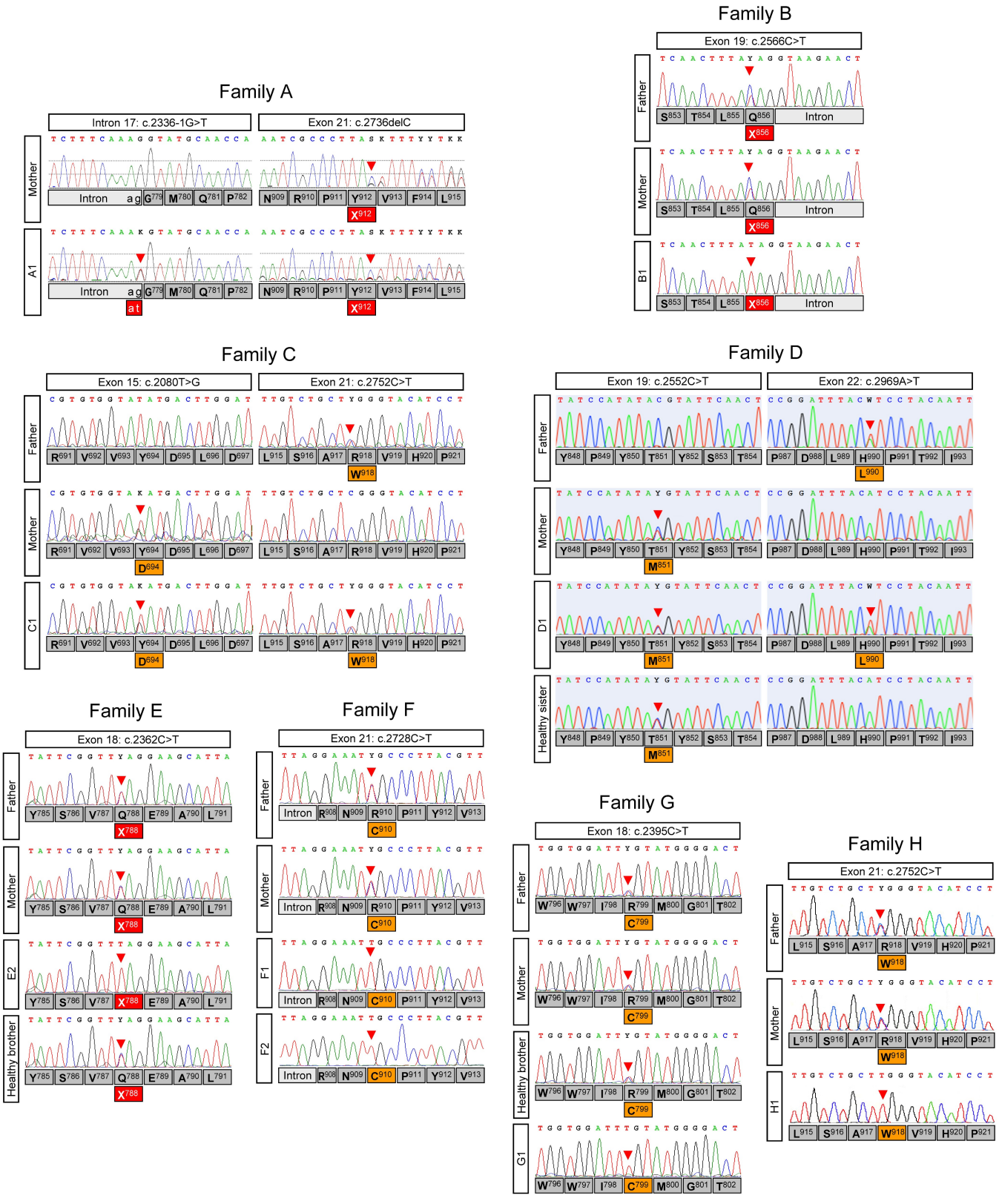
Appendix Note S1 - Members of the Undiagnosed Diseases Network.

David R. Adams, Mercedes E. Alejandro, Patrick Allard, Euan A. Ashley, Mahshid S. Azamian, Carlos A. Bacino, Ashok Balasubramanyam, Hayk Barseghyan, Gabriel F. Batzli, Alan H. Beggs, Babak Behnam, Hugo J. Bellen, Jonathan A. Bernstein, Anna Bican, David P. Bick, Camille L. Birch, Devon Bonner, Braden E. Boone, Bret L. Bostwick, Lauren C. Briere, Donna M. Brown, Matthew Brush, Elizabeth A. Burke, Lindsay C. Burrage, Manish J. Butte, Shan Chen, Gary D. Clark, Terra R. Coakley, Joy D. Cogan, Cynthia M. Cooper, Heidi Cope, William J. Craigen, Precilla D'Souza, Mariska Davids, Jean M. Davidson, Jyoti G. Dayal, Esteban C. Dell'Angelica, Shweta U. Dhar, Katrina M. Dipple, Laurel A. Donnell-Fink, Naghmeh Dorrani, Daniel C. Dorset, Emilie D. Douine, David D. Draper, Annika M. Dries, David J. Eckstein. Lisa T. Emrick, Christine M. Eng, Gregory M. Enns, Ascia Eskin, Cecilia Esteves, Tyra Estwick, Liliana Fernandez, Carlos Ferreira, Paul G. Fisher, Brent L. Fogel, Noah D. Friedman, William A. Gahl, Emily Glanton, Rena A. Godfrey, David B. Goldstein, Sarah E. Gould, Jean-Philippe F. Gourdine, Catherine A. Groden, Andrea L. Gropman, Melissa Haendel, Rizwan Hamid, Neil A. Hanchard, Lori H. Handley, Matthew R. Herzog, Ingrid A. Holm, Jason Hom, Ellen M. Howerton, Yong Huang, Howard J. Jacob, Mahim Jain, Yong-hui Jiang, Jean M. Johnston, Angela L. Jones, David M. Koeller, Isaac S. Kohane, Jennefer N. Kohler, Donna M. Krasnewich, Elizabeth L. Krieg, Joel B. Krier, Jennifer E. Kyle, Seema R. Lalani, C. Christopher Lau, Jozef Lazar, Kimberly LeBlanc, Brendan H. Lee, Hane Lee, Shawn E. Levy, Richard A. Lewis, Sharyn A. Lincoln, Sandra K. Loo, Joseph Loscalzo, Richard L. Maas, Ellen F. Macnamara, Calum A. MacRae, Valerie V. Maduro, Marta M. Majcherska, May Christine V. Malicdan, Laura A. Mamounas, Teri A. Manolio, Thomas C. Markello, Ronit Marom, Martin G. Martin, Julian A. Martínez-Agosto, Shruti Marwaha, Thomas May, Allyn McConkie-Rosell, Colleen E. McCormack, Alexa T. McCray, Jason D. Merker, Thomas O. Metz, Matthew Might,

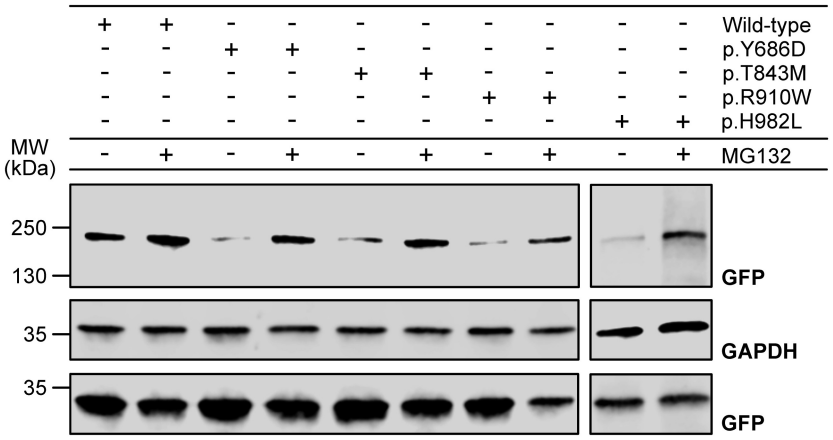
Paolo M. Moretti, Marie Morimoto, John J. Mulvihill, Jennifer L. Murphy, Donna M. Muzny, Michele E. Nehrebecky, Stan F. Nelson, J. Scott Newberry, John H. Newman, Sarah K. Nicholas. Donna Novacic, Jordan S. Orange, J. Carl Pallais, Christina GS. Palmer, Jeanette C. Papp, Neil H. Parker, Loren DM. Pena, John A. Phillips III, Jennifer E. Posey, John H. Postlethwait, Lorraine Potocki, Barbara N. Pusey, Chloe M. Reuter, Amy K. Robertson, Lance H. Rodan, Jill A. Rosenfeld, Jacinda B. Sampson, Susan L. Samson, Kelly Schoch, Molly C. Schroeder, Daryl A. Scott, Prashant Sharma, Vandana Shashi, Edwin K. Silverman, Janet S. Sinsheimer, Kevin S. Smith, Rebecca C. Spillmann, Joan M. Stoler, Nicholas Stong, Jennifer A. Sullivan, David A. Sweetser, Queenie K.-G. Tan, Cynthia J. Tifft, Camilo Toro, Alyssa A. Tran, Tiina K. Urv, Zaheer M. Valivullah, Eric Vilain, Tiphanie P. Vogel, Daryl M. Waggott, Colleen E. Wahl, Nicole M. Walley, Chris A. Walsh, Jijun Wan, Michael F. Wangler, Patricia A. Ward, Katrina M. Waters, Bobbie-Jo M. Webb-Robertson, Monte Westerfield, Matthew T. Wheeler, Anastasia L. Wise, Lynne A. Wolfe, Elizabeth A. Worthey, Shinya Yamamoto, Yaping Yang, Amanda J. Yoon, Guoyun Yu, Diane B. Zastrow, Chunli Zhao, Allison Zheng



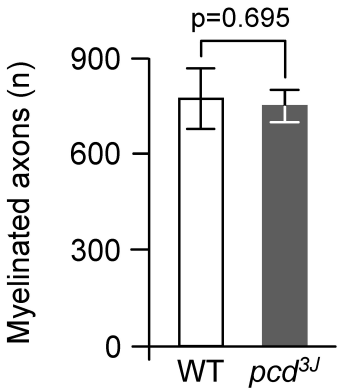
Appendix Figure S1 - Partial deletion of the *CCP1* gene (case J1).
Narrowing of the intragenic breakpoint by PCR amplification of fragments corresponding to exonic sequences predicted to be inside (exons 12) and outside (exon 13) the deletion region. In agreement with array CGH results, amplification of the exon 12-related fragment failed in patient's DNA while the exon 13-related product was detectable. Similar results were obtained for exons 10, 11 (inside the deletion region, no products) and 14 and 15 (outside the deletion region, products of the expected sizes) (data not shown). A β-actin primer set was used to ensure appropriate PCR set-up and sufficient template DNA quality and quantity. Con1-2: DNA samples from healthy, unrelated probands.



Appendix Figure S2 - Sanger validation and segregation analysis of *CCP1* variants. Variants segregate with the disease consistent with autosomal recessive inheritance in the families. All affected probands carried biallelic variants.



Appendix Figure S3 - CCP1 levels in transfected cells.
HEK293 cells were co-transfected with expression vectors encoding mouse wild-type and mutant CCP1-YFP and GFP (to control for transfection efficiency). After 7 hours, cells were treated with 5 μM MG132 proteasome inhibitor or vehicle (DMSO) for additional 17 hours. Protein extracts were subjected to western blotting and proteins were detected using primary antibodies as indicated. CCP1 mutants showed lower baseline protein levels which were normalized upon treatment with MG132. Residues Y686, T843, R910 and H982 of mouse CCP1 align to residues Y694, T851, R918 and H990 of the human ortholog.



Appendix Figure S4 - Axon numbers in the saphenous nerve.
Quantification of myelinated axons in the pure sensory femoral saphenous nerves of WT and *pcd* mice. Graphs: Mean ± SD, n = 3. P-value: Student's unpaired two-tailed t-test.

Appendix Table S1 - WES and data output.

	Family A	Family B	Family C	Family D	Family E	Family F	Family G	Family H	Family I
Enrichment	Agilent Clinical Research Exome	Agilent SureSelect50Mbv5	Agilent SureSelectXTv4	Illumina TruSeq Exome Enrichment	Illumina Rapid Capture Enrichment	Agilent SureSelect50Mbv5	Nextera Rapid Capture Exome v1.2	Illumina TruSeq Exome Enrichment	Roche VCRome_v2.1
Sequencing type	Mother & affected proband WES	Father & mother of affected proband WES[3)]	Affected proband WES	Father, mother, affected proband & healthy sibling WES	Affected proband WES	Affected proband & affected sibling WES	Affected proband WES	Affected proband WES	Father, mother & affected proband WES
Platform	Illumina HiSeq2000	Illumina HiSeq2000	Illumina HiSeq2000	Illumina HiSeq2000	Illumina HiSeq4000	Illumina HiSeq2000	Illumina HiSeq4000	Illumina HiSeq4000	Illumina HiSeq2000
Read alignment	BWA-Mem (v0.7.8)	BWA (v0.5.8)	BWA (v0.5.8)	NovoAlign	BWA (v0.5.8)	BWA (v0.5.8)	BWA (v0.5.8)	BWA (v0.5.8)	Edico DRAGEN BioIT platform
Variant calling	SAMtools (v0.1.18) & GATK Haplotype Caller (v3.1.7)	SAMtools (v0.1.7)	GATK (v3.3)	MPG	GATK (v3.6)	GATK (v3.3)	GATK (v3.6)	GATK (v3.3)	Edico Dragen haplotype-based variant calling
Criteria for variant exclusion	Variants with MAF >.01 in 1000G, >.05 in ExAC, >.005 in >80,000 internal exomes, genes with homozygous variants in >3 controls, or SNV quality <20	HapMap SNPs with average hetero-zygosity ≥.02 (dbSNP135), variants with MAF >.0007 in >11,000 internal exomes, or SNV quality <30	Variants with MAF >.05 in dbSNP137, >.01 in >2,000 internal exomes, or missense variants with a PhyloP60way score <3.5	Variants with MAF ≥.005 in ExAC, ESP & 587 internal exomes, CADD score ≤20, or coverage ≤10	Variants with MAF ≥.001 in ExAC & >5,000 internal exomes, or coverage ≤10	Variants with MAF ≥.001 in ExAC & >5,000 internal exomes, or coverage ≤10	Variants with MAF ≥.001 in ExAC & >5,000 internal exomes, or coverage ≤10	Variants with MAF ≥.01 in ExAC, ESP 3,000 internal exomes, or coverage ≤10	Variants with MAF >.01 in ExAC
Read depth[1)]									
Mean	158x	130x/117x	76x	58x	124x	84x	110x	100x	150x
≥10x	97%	99%/99%	98%	95%	98%	95%	97%	96%	99%
≥20x	95%	98%/98%	93%	79%	96%	88%	95%	99%	99%
Number of identi-fied variants[2)]									
Putative biallelic	1	40[3)]	2	2	7	2	6	4	10
Putative *de novo*	Not available	Not available	Not available	1	Not available	Not available	Not available	Not available	2
Putative X-linked	Not considered	Not considered	Not considered	1	Not considered	Not considered	Not considered	Not considered	1

[1)]Figures refer to proband's sample except for family B where values are given in the format paternal sample/maternal sample.

[2)]Frameshift, nonsense, splice site, missense, in-frame indel.

[3)]Quality of the DNA sample from the deceased proband B1 was not suitable for WES. Analysis of DNA samples from the parents yielded 40 genes in which they both carried identical heterozygous variants. Eight variants were considered unlikely based on frequency in databases, low scores for conservation and bioinformatic predictions of pathogenicity, different or no phenotype in knockout mouse models, or known association to another, non-overlapping human disease. For the remaining 32 variants, we preformed Sanger sequencing using the proband's DNA sample and found homozygosity for 10 variants while the remaining 22 variants were either heterozygous or absent. We did not consider genes in which the parents carried different heterozygous variants because we expected a homozygous disease-causing variant in the child due to known parental consanguinity.

Appendix Table S2 - Frequencies of disease-related CCP1 variants in control data sets.

	gnomAD[1)]	ESP[2)]	1000G[3)]	HZM[4)]
Number of samples	138,632	6,503	2,504	11,325
c.988C>T;p.R330*	1 (3.23e-5)	0	0	0
c.2080T>G;p.Y694D	0	0	0	0
c.2336-1G>T;p.M780fs	0	0	0	0
c.2362C>T;p.Q788*	0	0	0	0
c.2395C>T;p.R799C	0	0	0	0
c.2552C>T;p.T851M	0	0	0	1 (4.42e-5)
c.2566C>T;p.Q856*	0	0	0	0
c.2728C>T;p.R910C	0	0	0	0
c.2736delC;p.Y912*	1 (4.12e-6)	0	0	0
c.2752C>T;p.R918W	0	0	0	0
c.2969A>T;p.H990L	0	0	0	0

[1)]GnomAD: 123,136 exomes and 15,496 genomes from unrelated individuals sequenced as part of various disease-specific and population genetic studies contained in the Genome Aggregation Database (gnomAD) (Lek *et al.*, 2016)
[2)]ESP: Exomes from 6,503 individuals with heart, lung and blood disorders included in the NHLBI GO Exome Sequencing Project (ESP) (Fu *et al.*, 2013)
[3)]1000G: Genomes from 2,504 individuals who declared themselves to be healthy at the time the samples were collected (Sudmant *et al.*, 2015)
[4)]HZM: Exomes from 11,325 individuals with various genetic and acquired diseases contained in the WES repository of the Helmholtz Zentrum München (HZM)

Appendix Table S3 - Bioinformatic results for CCP1 amino acid substitutions.

	c.2080T>G; p.Y694D	c.2395C>T; p.R799C	c.2552C>T; p.T851M	c.2728C>T; p.R910C	c.2752C>T; p.R918W	c.2969A>T; p.H990L
PROVEAN prediction (score)[1)]	Deleterious (-7.84)	Deleterious (-7.53)	Deleterious (-5.43)	Deleterious (-6.63)	Deleterious (-7.14)	Deleterious (-7.16)
SIFT prediction (score)[2)]	Damaging (0)	Damaging (0)	Damaging (0)	Damaging (0)	Damaging (0)	Damaging (0)
PolyPhen-2 prediction (score)[3)]	Probably damaging (.99)	Probably damaging (.99)	Probably damaging (1)	Probably damaging (.99)	Probably damaging (.99)	Probably damaging (.95)
CADD prediction (phred-like score)[4)]	Damaging (29)	Damaging (35)	Damaging (33)	Damaging (35)	Damaging (35)	Damaging (33)
LRT prediction (LRT_{new} score)[5)]	Deleterious (1.00)	Deleterious (1.00)	Deleterious (1.00)	Deleterious (1.00)	Deleterious (1.00)	Deleterious (1.00)
SNAP2 prediction (score)[6)]	Effect (85)	Effect (52)	Effect (65)	Effect (64)	Effect (95)	Effect (31)
MutationAssessor prediction (score)[7)]	Medium (2.79)	High (3.91)	High (4.30)	High (3.7)	High (3.84)	Medium (3.25)
VEST3 score[8)]	.95	.98	.95	.95	.97	.98
MutationTaster2 prediction (score)[9)]	Disease causing (.99)	Disease causing (.99)	Disease causing (.99)	Disease causing (.99)	Disease causing (.99)	Disease causing (.99)
PMut prediction (score)[10)]	Disease (.79)	Disease (.89)	Disease (.89)	Disease (.87)	Disease (.91)	Disease (.87)
GERP++ score[11)]	5.54	5.50	4.95	5.65	5.65	5.69
PhastCons100way score[12)]	1.00	1.00	1.00	1.00	1.00	1.00
PhyloP100way score[12)]	8.80	4.90	9.87	7.77	5.12	8.04

[1)]PROVEAN (Choi *et al.*, 2012) scores equal or below -2.50 are considered "deleterious".
[2)]SIFT (Kumar *et al.*, 2009) scores $\leq$.05 are assigned the prediction "damaging".
[3)]PolyPhen-2 (Adzhubei *et al.*, 2010) scores near 1 are most strongly predicting a "damaging" effect of an amino substitution.
[4)]CADD (Kircher *et al.*, 2014) phred-like rank scores above 15 (for a more conservative estimate: above 20) are considered "damaging".
[5)]Values for the LRT_{new} (Chun & Fay, 2009) score range from 0 to 1 with higher values indicating a variant is more likely to be "deleterious".
[6)]SNAP2 (Hecht *et al.*, 2015) scores range from -100 strong neutral prediction to +100 strong effect prediction.
[7)]MutationAssessor (Reva *et al.*, 2011) scores range from -5.14 to 6.49 with higher scores indicating increasing likelihood of functional impact of a variant. Score cutoff between "neutral", "low", "medium" and "high" predictions are 0.8, 1.94 and 3.50.
[8)]VEST3 (Carter *et al.*, 2013) score ranges from 0 to 1. The larger the score the more likely the variant may cause functional change.

[9)]The probability value given by MutationTaster2 (Schwarz *et al.*, 2010) is the probability of the prediction, i.e. a value close to 1 indicates a high "security" of the prediction.
[10)]The PMut (López-Ferrando *et al.*, 2017) classifier outputs a prediction score between 0 and 1; mutations scoring from 0 to 0.5 are classified as "neutral", and those scoring from 0.5 to 1 are classified as "disease".
[11)]GERP++ (Davydov *et al.*, 2010) estimates evolutionary constraint of specific positions in 36 mammalian species. Scores range from -12.36 to 6.18 with higher scores indicating more conserved sites.
[12)]PhastCons and PhyloP (Pollard *et al.*, 2010) conservation scores are based on multiple alignments of 100 vertebrate genomes. Scores range from 0 to 1 for PhastCons and from -20 to 9.87 for PhyloP with higher scores suggesting stronger conservation of the site.

Appendix Table S4 - Clinical findings in 13 individuals with biallelic *CCP1* variants.

	Family A	Family B	Family C	Family D	Family E			Family F		Family G	Family H	Family I	Family J
Consanguinity / ethnicity	No / Caucasian	Yes / Persian	No / Caucasian	No / Caucasian	Yes / Egyptian			Yes / Egyptian		Yes / Egyptian	Yes / Saudi Arab	No / African American, Caucasian	Yes / Caucasian
Identified variants	c.2336-1G>T, p.Y912*	p.Q856* (hom.)	p.Y694D, p.R918W	p.T851M, p.H990L	p.Q788* (hom.)			p.R910C (hom.)		p.R799C (hom.)	p.R918W (hom.)	p.R330*, p.Y912*	Del exons 1-12 (hom.)
Individual	A1	B1	C1	D1	E1	E2	E3	F1	F2	G1	H1	I1	J1
Sex / age	F / 2 y	F / † 16 m	F / 5 y	M / 14 y	M / † 8 m	F / † 5 m	M / † 4 m	M / 8 y	F / 5 y	M / 8 y	M / † 12 m	F / 20 m	F / † 7 m
Pregnancy	Normal	Normal	Normal	Normal	Normal	Normal	Fetal MRI: Abnormal corpus callosum	Normal	Normal	Bleeding in early pregnancy	Normal	Normal	Normal
Birth	Normal	C-section	Normal	C-section	Normal	C-section	C-section	Normal	C-section	Normal	C-section	C-section	Normal
Age of onset / first symptom	2 m / hypotonia	4 m / abnormal eye movements	3 m / hypotonia	Birth / hypotonia	2 m / GDD, FTT	2 m / GDD, FTT	2 m / GDD, FTT	20 m / hypotonia	4 m / hypotonia	3 m / hypotonia	3 m / hypotonia, tongue fasciculations, weak cry	6 m / eye movements abnormal, GDD	3 m / upper limb tremor, weak suck
Progressive disease course	Worsening of head control	Loss of head control	Loss of rolling over	Worsening of balance	Yes	Yes	Yes	Loss of crawling	Yes	Worsening of dystonia	Yes	Yes	Yes
Microcephaly	Yes	No	Yes	No	Yes	Yes	Yes (-2.5 SD)	No	No	Yes (-2.8 SD)	Yes	No	Yes (-2.2 SD)
Motor delay	Profound	Profound	Profound	Mild	Profound	Profound	Profound	Severe	Severe	Severe	Profound	Severe	Profound
Best motor ability / age when reached	Knee flexion / 6 m	Neck flexion / 4 m	Rolling over / n.a.	Walking / 14 m	No motor development	No motor development	No motor development	Crawling / 12 m	Neck flexion / 4 m	Standing with support	No motor development	Rolling over / 8 m	No motor development
Cognitive delay, mental retardation	No speech	No speech	No speech	Autism, learning disability	No visual recognition	No visual recognition	No visual recognition	Severe	Severe	No speech	No visual recognition	No speech	Profound
Eye movement abnormalities	Strabismus, slow saccades	Strabismus	Strabismus	Upgaze palsy, slow saccades	Strabismus	No	Strabismus	No	Upgaze palsy	Strabismus, nystagmus	Strabismus	Strabismus, nystagmus	Nystagmus
Hypotonia	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Muscle weakness	Tetraplegia	Tetraparesis	Tetraparesis	Lower limb	Yes	Yes	Yes	Yes	Yes	Yes	Tetraplegia	No	Tetraparesis
Muscle wasting	Yes	Yes	No	No	No	No	No	Yes	Yes	Yes	No	No	Yes
Fasciculations	Tongue	Tongue	No	No	No	No	No	No	No	No	Tongue	No	Tongue
Tendon reflexes	0	0	3+	3+	0	0	0	3+	3+	3+	0	2+	0

(continued on next page)

Appendix Table S4 - Clinical findings in 13 individuals with biallelic *CCP1* variants. (continued from previous page).

	Family A	Family B	Family C	Family D	Family E			Family F		Family G	Family H	Family I	Family J
Individual	A1	B1	C1	D1	E1	E2	E3	F1	F2	G1	H1	I1	J1
Ataxia	n.e.	n.e.	n.e.	Yes	n.e.	n.e.	n.e.	Yes	Yes	Yes	n.e.	Yes	n.e.
Dystonia	No	No	Yes	No	No	No	No	Yes	Yes	Yes	No	No	No
Spasticity	No	No	Lower limb	Lower limb	No	No	No	Yes	Yes	Yes	No	Yes	No
Respiratory insufficiency	Tracheostomy, nightly BIPAP	Yes	No	No	No	No	No	No	No	No	Yes	No	Yes
Feeding difficulties	Gastrostomy	Yes	Gastrostomy	No	FTT	FTT	FTT	No	No	No	No	Yes	Yes
Retina degeneration	No	No	n.d.	No	No	No	No	No	No	No	n.d.	No	n.d.
Brain MRI	Cerebellar atrophy, dysplastic corpus callosum	Cerebellar atrophy, dysplastic corpus callosum	Cerebellar atrophy	Cerebellar atrophy	Cerebellar atrophy, dysplastic corpus callosum	Cerebellar atrophy, dysplastic corpus callosum	Cerebellar atrophy, dysplastic corpus callosum	Cerebellar atrophy	Cerebellar atrophy	Cerebellar atrophy	Cerebellar atrophy, global brain atrophy	Cerebellar atrophy	Cerebellar atrophy, hypoplastic corpus callosum, small pons
Nerve conduction studies	Axonal motor neuropathy	Axonal motor neuropathy	Normal	Axonal motor predominant neuropathy	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	Motor neuropathy	n.d.	Axonal motor neuropathy
Electromyography	Denervation	Denervation, marked neurogenic reorganization	Normal	Mild neurogenic pattern	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.a.	n.d.	Denervation
Muscle ultrasound	Hyperechoic, atrophy	n.d.	n.d.	Hyperechoic, atrophy	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Muscle biopsy	n.d.	Denervation atrophy, type 1 fibre predominance	Mild type 1 fibre predominance	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	Denervation atrophy	n.d.	n.d.

† = Deceased; BIPAP = Biphasic positive airway pressure ventilation; F = Female; FTT = Failure to thrive; GDD = Global developmental delay; hom. = Homozygous; M = Male; m = Months; MRI = Magnetic resonance imaging; n.a. = Not available; n.d. = Not done; n.e.= Not examinable (due to severe weakness); SD = Standard deviation; y = Years; Deep tendon reflexes: 0 = Absent, 1+ = Diminished, 2+ = Normal, 3+ = Brisk, 4+ = Hyperactive and clonus

Appendix References

Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, Gabriel S, Rieder MJ, Altshuler D, Shendure J, Nickerson DA, Bamshad MJ, Akey JM (2013) Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* 493: 216-20

Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, Zhang Y, Ye K, Jun G, Hsi-Yang Fritz M, Konkel MK, Malhotra A, Stutz AM, Shi X, Paolo Casale F, Chen J, Hormozdiari F, Dayama G, Chen K, Malig M et al. (2015) An integrated map of structural variation in 2,504 human genomes. *Nature* 526: 75-81