Mutations in *KCND3* cause spinocerebellar ataxia type 22

Yi-Chung Lee, MD. PhD,[1,2,3 *] Alexandra Durr, MD, PhD,[4,5,6,7*] Karen Majczenko, MD,[8,9*] Yen-Hua Huang,MD, PhD,[10,11] Yu-Chao Liu, BS,[12] Cheng-Chang Lien, MD, PhD,[2,12] Yaeko Ichikawa, MD, PhD,[13] Jun Goto, MD, PhD,[13] Marie-Lorraine Monin, MD,[4,5,6] Jun Z. Li, PhD,[14,15] Ming-Yi Chung, PhD,[16,17] Emeline Mundwiller, BS,[4,5,6] James Dell' Orco, BS[9], Vikram Shakkottai, MD, PhD,[9] Tze-Tze Liu, PhD,[18] Christelle Tesson, MS,[4,5,6,19] Yi-Chun Lu, BS,[3] Pei-Chien Tsai, PhD,[2] Alexis Brice, MD,[4,5,6,7] Shoji Tsuji, MD, PhD,[13] Margit Burmeister, PhD,[8,14,15,20#] Giovanni Stevanin, PhD,[4,5,6,7,19 #] Bing-Wen Soong, MD, PhD[1,2,3 ,12#]

[1]Department of Neurology, National Yang-Ming University School of Medicine, Taipei, Taiwan
[2]Brain Research Center, National Yang-Ming University, Taipei, Taiwan
[3]Department of Neurology, Taipei Veterans General Hospital, Taipei, Taiwan
[4]INSERM, U975, Paris, France
[5]CNRS, UMR7225, Paris, France
[6]UPMC University Paris, Centre de Recherche du Cerveau et de la Moelle épinière, Hopital Pitie-Salpetriere, Paris, France
[7]APHP, Department of Genetics, Hopital Pitie-Salpetriere, Paris, France
[8]Molecular & Behavioral Neuroscience Institute, University of Michigan, Ann Arbor, MI, USA
[9]Department of Neurology, University of Michigan, Ann Arbor, MI, USA
[10]Department of Biochemistry, Faculty of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan
[11]Center for Systems and Synthetic Biology, National Yang-Ming University, Taipei, Taiwan
[12]Institute of Neuroscience, National Yang-Ming University, Taipei, Taiwan
[13]Department of Neurology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan
[14]Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA
[15]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA
[16]Faculty of Life Sciences and Institute of Genomic Sciences, National Yang-Ming University, Taipei, Taiwan
[17]Department of Medical Research and Education, Taipei Veterans General Hospital, Taipei, Taiwan

[18]Genome Research Center, National Yang-Ming University, Taipei, Taiwan
[19]Ecole Pratique des Hautes Etudes, Paris, France
[20]Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA

Corresponding author: Bing-Wen Soong, MD, PhD
Department of Neurology
National Yang-Ming University School of Medicine
Taipei Veterans General Hospital
#201, Sec 2, Shipai Road, Peitou District,
Taipei Taiwan 11217

Giovanni Stevanin, PhD,
UPMC University Paris, Centre de Recherche du Cerveau et de la Moelle épinière, Hopital Pitie-Salpetriere, Paris, France

Margit Burmeister, PhD,
Molecular & Behavioral Neuroscience Institute, University of Michigan, Ann Arbor, MI, USA

* These authors contributed equally.
# Corresponding authors

Supporting Information

Subjects and Methods

Exome Sequencing

In Families A and B, about 213,000 exonic regions from the genomic DNA of two family members (Fig 1A, IV-1 and III-10; Fig 1B, II-7 and III-6) were captured and enriched using the Agilent SureSelect Human All Exon 50 Mb kit according to the manufacturer's protocol (Agilent, CA). The DNA sequencing libraries were prepared using Illumina's paired end DNA sample preparation kit according to the manufacturer's instruction. The enriched samples were sequenced on the Illumina HiSeq2000 (Illumina, CA) platform using paired-end 2x100 bp protocol. All sequenced reads were mapped to the human reference genome (GRCh37 patch 2, obtained from ftp://ftp.ensembl.org/pub/release-61/) using the Burrows-Wheeler Aligner (BWA).[1] PCR duplicates were removed from the alignments using SAMtools, and then the variant calling was performed using SAMtools/BCFTools (http://samtools.sourceforge.net/mpileup.shtml).[2] To further ensure that the collected variants were of high confidence, we preserved only the variants with Phred-like confidence scores at least 30.0, as was suggested by GATK (http://www.broadinstitute.org/gsa/wiki/index.php/Best_Practice_Variant_Detection_with_the_GATK_v3) for processing high-coverage (>10x) datasets.

In Family C, exome capture was performed on the two most distantly related affected individuals (IV-6 and IV-7, see Fig 1 C) of the pedigree and was followed by next generation sequencing using an Illumina Hiseq2000 Instrument. Sequence reads were aligned to the human hg19 genome using bwa[1] and duplicated reads were removed using samtools.[2]

Because SCA22 is inherited in an autosomal-dominant manner, the causative mutation shared by the affected members should be heterozygous. Thus, variants that were homozygous or not harbored in the genomes of affected patients (and/or present in the unaffected parent of family B) were not taken for further analyses. Next, based on the hypothesis that the mutation of this rare SCA is not present in the general population, variants that had been identified in the Single Nucleotide Polymorphism Database (dbSNP, Human Build 132, http://www.ncbi.nlm.nih.gov/projects/SNP/), the 1000 Genomes Project,[3] or previously exome-sequenced samples of 10 individuals without SCA from Chinese population or 32 non-SCA individuals from France were excluded. Then, we removed variants not influencing coding sequences, including synonymous and most non-coding variants. In family A, significantly linked to 1p21-q23, variants not present in this region were also omitted. In families B and C, the analysis of exome sequencing was restricted to linked regions. Finally, Sanger sequencing was performed to determine whether and which of the remaining variants segregated with disease phenotype in other family members.

In silico analyses

The residue changes identified were further evaluated by *in silico* analyses with the following programs: SIFT, PMut and Polyphen 2.

URLs. NHLBI Exome Sequencing Project http://snp.gs.washington.edu/EVS/; Novocraft (http://www.novocraft.com); Ensembl http://uswest.ensembl.org; dbSNP132 http://www.ncbi.nlm.nih.gov/projects/SNP/; 1000 Genome Project http://www.1000genomes.org/; SIFT: http://sift.jcvi.org/; PMut: http://mmb.pcb.ub.es/PMut/PMut.jsp; PolyPhen 2:

http://genetics.bwh.harvard.edu/pph2/

Results

Exome Sequencing Identifies KCND3 Mutations in Families with Dominant Ataxia

In Family A, 53.9 and 50.9 million sequencing reads were produced for the two samples, comprising 5.39 and 5.09 billion bases, respectively. About 93% and 93.7% of the reads for both samples, respectively, could be aligned to the human reference genome. About 0.4% of the targeted exon regions in IV-1 and about 1% of these regions in III-10 were un-sequenced. The average sequencing depths for the targeted exonic regions were approximately 105x and 96x, respectively, for these two samples. After analysis and filtering (Supplementary Table 1A), 489 heterozygous coding variants were shared by both patients and were not present in the dbSNP, the 1000 Genomes Project or the exome data from 10 non-SCA ethnic controls. Among these variants, 11 were located in the 1p21-q23 region, to which SCA19/22 is linked. Sanger sequencing of these variants in other 23 family members, including 11 patients, 6 unaffected individuals and 6 married-in spouses, identified only one variant, c.679_681delTTC in *KCND3* gene, completely segregating with the disease phenotype. (Fig 2A) This mutation putatively results in a deletion of phenylalanine at position 227 (p.F227del) of Kv4.3. The deleted phenylalanine of Kv4.3 is highly conserved across a wide variety of species, from zebrafish, frog, platypus and mouse to humans. (Fig 2E) This mutation was not found in 500 normal Chinese controls. Mutations were also not found in *KCND3* in 105 unrelated probands with dominant ataxia or 55 patients with sporadic cerebellar ataxia in Taiwan by sequence analysis.

In Family B, whole genome scan analysis identified 11 regions cosegregating with the disease with LOD scores from -1 to +2.8. The analysis of 34 additional

microsatellites allowed to restrict some of these loci that represented a total of approximately 59 Mb. Following capture and next generation sequencing in subject II-7 and III-6, 71373389 to 91598083 non-duplicated reads were obtained and mapped on the reference genome hg19. The mean coverage ranged from 78 to 82% at 40x and from 89 to 90% at 20x. Of the 1648 nucleotide variants present in III-6 but absent in the unaffected father (II-7), 1240 were heterozygous, and 146 variants remained following removal of synonymous variants. Further filtering by excluding validated rs variants in dbSNP and 1000 genome (March 2012 version) and intergenic variants left 13 variants, of which 5 resided in the linked regions, 2 of which were not rare (>1% in Exome Variant Server). Therefore, three variants remained (Supplementary Table 2): a) c.1180G>T p.Asp394Tyr in *NDUFS1*, a mitochondrial protein involved in a different disease in the homozygous state: Mitochondrial Complex I Deficiency. b) c.430C>T p.Arg144Cys in *CXCR1*, receptor to interleukin-8, which is a powerful chemotactic factor to neutrophils. Mice homozygous for a gene disruption display normal morphology, clinical chemistry, hematology, and behavior. This variant was reported at 0.01% in the Exome Variant Server and was reported in a heterozygous case also in dbSNP (1/2273 subjects). c) c.678_680delCTT p.Phe227del, affecting a strongly conserved amino-acid in *KCND3* and that cosegregated with the disease. This variant was absent in 5,363 exomes at the Exome Variant Server and in 31 local exomes done for other non-SCA diseases. It was also absent from 152 French control chromosomes.

Family C presented with incompletely penetrant mid-life onset ataxia. Linkage analysis resulted in several linkage peaks on multiple chromosomes, reaching maximally a LOD score of 1.97 in a large region on chromosome 1. In addition, using high density SNP chips and PLINK,[4] shared haplotypes between the two most distant

relatives were identified (data not shown). Only two large shared haplotypes matched the linkage peaks, one 62 Mb long haplotype on chromosome 1, and a 33 Mb haplotype on chromosome 15. The shared haplotype with the highest LOD score on chromosome 1 overlapped the SCA19/SCA22 region. Next generation exome sequencing identified a large number of shared variants between the two individuals. These were filtered by linkage, haplotype as well as the SCA19/22 interval, absence on the Washington University Exome Server (Exome Variant Server, NHLBI Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [7/2011, rechecked 4/2012]) and predicted damage in function. Only one variant passed all these filters: G345V in *KCND3*. Conventional sequencing confirmed complete segregation of G345V in the family.

Supplementary Table 1. Bioinformatic analysis of exome sequencing in Family A

	IV-1	III-10
Total bases sequenced	10,773,734 Kb	10,177,251 Kb
Raw heterozygous variants (number)	165,125	156,440
Common heterozygous variants	35,953	
Variants not in dbSNP or 1000 genomes database	8,102	
Variants not found in other 10 non-SCA patients' exome data	5,715	
Variants resulting in amino acid sequence changes	489	
Variants present in the linked region (1p21-q23)	11	
Variants completely segregating with disease phenotype in the pedigree	1	

Supplementary Table 2. Bioinformatic analysis of exome sequencing in Family B.

	III-6	II-7
Variants	43364	39777
Variants present in the affected but absent in the unaffected father	1648	
Heterozygous variants	1240	
Variants resulting in amino acid sequence changes	146	
Variants not in dbSNP or 1000 genomes database	13	
Variants present in the 11 linked regions	5	
Variants absent in Exome variant server and local 31 exomes	3	
Variants segregating with disease phenotype in the pedigree	3 (including one on chromosome 1)	

Supplementary Table 3. Bioinformatic analysis of exome sequencing in Family C

	IV-6	IV-7
Variants	23475	28549
Variants present in both affected individuals as heterozygotes	5514	
Variants present in at least one individual that fall in the 11 potentially linked regions	1769	
Variants resulting in amino acid sequence or splice site changes and absent in dbSNP and 1000 genomes	31	
Variants also present in the two shared large haplotypes	13	
Variants (of these) predicted to be damaging, not present in the exome server, and segregating with the family	1	

Supplementary Table 4. Deleterious consequences from residue changes predicted by *in silico* analysis

Change in		Exon	Motif	Prediction Program			
Nucleotide	Amino acid			PMut	SIFT	PolyPhen-2	
						HumDiv	HumVar
c.1013T>C	p.V338E	2	S5	Pathogenic	Damaging	Probably damaging	Probably damaging
c.1034G>T	p.G345V	2	Re-entrant loop	Pathogenic	Damaging	Benign	Possibly damaging
c.1130C>T	p.T377M	3	Re-entrant loop	Pathogenic	Damaging	Probably damaging	Probably damaging

Supplementary Table 5. Clinical features of Family A

Subject	II-3	II-6	III-2	III-8	III-10	IV-1
Age at onset (yrs)	46	35	32	17	20	15
Age at exam (yrs)	81	78	57	48	59	32
Signs at onset	Ataxia	Ataxia	Ataxia	Ataxia	Ataxia	Ataxia
Oculomotor abnormalities*	Moderate	Moderate	Mild	Mild	Mild	Mild
Dysarthria/dysphagia	Moderate	Moderate	Mild	Mild	Mild	No
DTRs	Decreased	Decreased	Decreased	Decreased	Increased	Decreased
Extrapyramidal signs	No	No	No	No	No	No
Cognitive impairment	No	No	No	No	No	No
Babinski's sign	Absent	Absent	Absent	Absent	Absent	Absent
EKG/echocardiogram	NA	NA	Normal	Normal	NA	NA
EMG	CLR	NA	CLR	Mild SMP	NA	NA
VEP/AEP/SSEP	Unremarkable	Unremarkable	Unremarkable	Unremarkable	NA	NA
Brain MRI	Mild CA	NA	Mild CA	Mild CA	NA	NA

* Nystagmus, saccadic pursuits;

AEP: auditory evoked potentials; CA: cerebellar atrophy; CLR: chronic lumbar radiculopathies; DTR: deep tendon reflexes; EKG: electrocardiogram; EMG: electromyogram; MRI: magnetic resonance image; NA: not available; SMP: sensorimotor polyneuropathy; SSEP: somatosensory evoked potentials; VEP: visual evoked potentials.

Supplementary Table 6. Clinical features of Family B

Subject	II-6		II-8	II-9 (index)		III-1	III-4		III-6	III-8		III-9	
Age at onset (yrs)	35		50	51		40	30		31	30		24	
Age at exam (yrs)	61	77	61	54	61	44	38	45	32	31	40	28	44
Signs at onset	Ataxia		Ataxia	Ataxia		Ataxia	Ataxia		NA	Intermittent diplopia		Ataxia	
Disability (7 max worse)	4	6	4	3	4	1	2	2	1	1	2	1	5
Cerebellar gait ataxia	Moderate	Severe	Moderate	Moderate	Moderate	Mild	Mild	Moderate	Mild	Mild	Mild	Mild	Severe
Cerebellar dysarthria	Yes	Yes	NA	Yes	Yes	No	Yes	Yes	No	intermittent	NA	Yes	Yes
Oculomotor abnormalities	Nystagmus, vertical ophthalmoplegia	Nystagmus, vertical ophthalmoplegia	Normal	Normal	Nystagmus, saccadic pursuit	Normal	Normal	Diplopia, saccadic pursuit	Saccadic pursuit	Normal	Saccadic pursuit	Normal	Nystagmus, saccadic pursuit
DTRs	NA	No	Increased	No	Decreased	No	No	No	Increased	Increased	NA	No	NA
Babinski sign	NA	No	No	No	indifferent	No	No	No	No	No	NA	No	NA
Vibration sense	Decreased	Abolished	Normal	Normal	Decreased	Normal	Normal	Decreased	Normal	Normal	Normal	Normal	Normal
Dysphagia	NA	NA	No	No	No	No	No	No	No	NA	Yes	No	Yes
Extrapyramidal signs	NA	Rigidity	No	No	Rigidity	No	No	No	Tremor	Tremor	NA	No	No
Urinary dysfunction	NA	Incontinence	No	NA	Urgencies	No	No	No	Urgencies	No	NA	Urgencies	Urgencies
Cognitive impairment	NA	NA	No	No	No	No	No	No	No	No	NA	No	No
EMG	NA	Normal	NA	NA	Sensory polyneuropathy	NA	NA	Normal	Normal	Normal	NA	NA	NA

Brain MRI	NA	CA	CA	NA	CA	NA	NA	CA	NA	NA	CA	CA	NA

CA: cerebellar atrophy; DTR: deep tendon reflexes; EMG: electromyogram; MRI: magnetic resonance image; NA: not available;

Supplementary Table 7. Clinical features of Family C

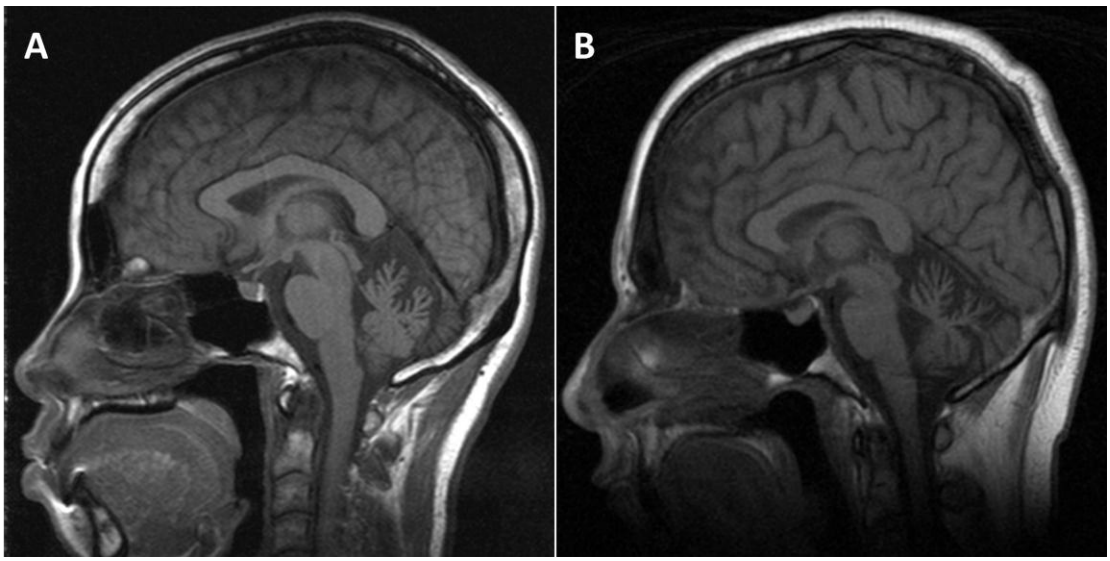
Subject	III-3	IV-6	IV-7
Age at onset (yrs)	<55	<35	39
Age at exam (yrs)	65	39	54
Signs at onset	Ataxia	Ataxia	Ataxia
Gait Ataxia	Mild	Very mild	Mild
Dysarthria	Present	NA	Present
ataxia upper limbs	Absent	NA	Present
ataxia lower limbs	present	NA	Present
Extrapyramidal signs	No	No	No
Muscle strength	Normal	NA	Normal
Oculomotor signs	Saccadic pursuit	NA	Normal
Cognitive impairment	No	No	No
DTRs	Normal	NA	Increased
Babinski's sign	Absent	NA	Present
Brain MRI	Mild CA	NA	NA

CA: cerebellar atrophy; DTR: deep tendon reflexes; MRI: magnetic resonance image; NA: not available;

Supplementary Table 8. Clinical features of Family D, E, F

Family	D			E			F			
Subject	I-2	II-1	II-4	I-1	II-2	II-3	I-2	II-1	III-1	III-2
Age at onset (yrs)	NA	35	NA	early 90s	<58	51	NA	40s	late teens	Early teens
Age at exam (yrs)	75	49	35	died at 94	63	61	NA	65	32	26
Signs at onset	NA	Delusion of persecution	NA	unstable gait	Ataxia	Ataxia	Ataxia	NA	Ataxia	Ataxia
Gait Ataxia	NA	Mild	NA	NA	Present	Present	Present	NA	Present	Present
Dysarthria	NA	Present	NA	NA	Present	NA	NA	NA	Present	NA
ataxia upper limbs	NA	NA	NA	NA	Normal	NA	NA	NA	Present	NA
ataxia lower limbs	NA	NA	NA	NA	Mild	NA	NA	NA	Present	NA
Extrapyramidal signs	NA	NA	NA	NA	Absent	NA	NA	NA	NA	NA
Muscle strength	NA	NA	NA	NA	Normal	NA	NA	NA	NA	NA
Oculomotor signs	NA	NA	NA	NA	Normal	NA	NA	NA	Saccadic pursuit	NA
Cognitive impairment	NA	Suspected	NA	NA	No	NA	NA	NA	No	NA
DTRs	NA	NA	NA	NA	Normal	NA	NA	NA	Increased	NA
Babinski's sign	NA	NA	NA	NA	Absent	NA	NA	NA	Absent	NA
Involuntary movement	NA	Grimacing face	NA	NA	No	NA	NA	NA	NA	NA
Brain MRI	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

DTR: deep tendon reflexes; MRI: magnetic resonance image; NA: not available



Supplementary FIGURE. MR imaging in SCA19/22. Sagittal T1-weighted scan of patient (A) III-2 (47 years of age) of family A and (B) III-9 (36 years of age) of family B, showing cerebellar volume loss and normal brainstem after 15 and 12 years, respectively, of disease duration, correlating with mild functional impairment.

References

1. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754-1760
2. Li H, Handsaker B, Wysoker A et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078-2079
3. Durbin RM, Altshuler D, Abecasis GR et al. A map of human genome variation from population-scale sequencing. Nature. 2010;467:1061-1073
4. Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559-575