Loss of Cajal Bodies in Motor Neurons from patients with novel mutations in VRK1

Lara El-Bazzal, Khalil Rihan, Nathalie Bernard-Marissal, Christel Castro, Eliane Chouery-Khoury, Jean-Pierre Desvignes, Alexandre Atkinson, Karine Bertaux, Salam Koussa, Nicolas Lévy, et al.

Loss of Cajal Bodies in Motor Neurons from patients with novel mutations in *VRK1*

Lara El-Bazzal[1], Khalil Rihan[1], Nathalie Bernard-Marissal[1], Christel Castro[1], Eliane Chouery-Khoury[2], Jean-Pierre Desvignes[1], Alexandre Atkinson[1], Karine Bertaux[3], Salam Koussa[4], Nicolas Lévy[1,5], Marc Bartoli[1], André Mégarbané[6,7] , Rosette Jabbour[8], Valérie Delague[1,*]

1. Aix Marseille Univ, Inserm, MMG, U 1251, Marseille, France

2. Unité de génétique médicale, Université Saint Joseph, Campus des Sciences Médicales, Beyrouth, Liban

3. Medical Genetics, Biological Resource Center – Tissue, DNA, Cells, CRB TAC, La Timone Children's Hospital, Marseille, France

4. Department of Neurology, Lebanese University Hospital-Geitaoui, Beirut, Lebanon

5. Department of Medical Genetics, Children's Hospital La Timone, Marseille, France

6. Centre médical et psychopédagogique, Beyrouth, Liban

7. Institut Jérôme Lejeune, Paris, France

8. Neurology Division, Department of Internal Medicine, St George Hospital University Medical Center, University of Balamand, Beirut, Lebanon

*__Correspondence to__:
Valérie Delague
U 1251, Marseille Medical Genetics
Faculté de Médecine de la Timone
27 bd Jean Moulin
13385 Marseille cedex 05
Phone: +334 91 32 49 02
Fax :+33491804319
e-mail: valerie.delague@univ-amu.fr

Abstract

Distal Hereditary Motor Neuropathies are a heterogeneous group of diseases, resembling Charcot-Marie-Tooth syndromes, but characterized by an exclusive involvement of the motor part of the peripheral nervous system.

Here, we describe two new compound heterozygous mutations in *VRK1*, the vaccinia-related kinase 1 gene, in two sibs from a Lebanese family, affected with distal Hereditary Motor Neuropathy associated with upper motor neurons signs. The mutations lead to severely reduced levels of VRK1 by impairing its stability, and to a shift of nuclear VRK1 to cytoplasm. Depletion of VRK1 from the nucleus alters the dynamics of coilin, a phosphorylation target of VRK1, by reducing its stability through increased proteasomal degradation. In human induced Pluripotent Stem Cell-derived Motor Neurons from patients, we demonstrate that this drop in VRK1 levels leads to Cajal Bodies disassembly and to defects in neurite outgrowth and branching.

Mutations in *VRK1* have been previously reported in several neurological diseases affecting lower or both upper and lower motor neurons. Here, we describe a new phenotype linked to *VRK1* mutations, presenting as a classical slowly progressive motor neuropathy, beginning in the second decade of life, with associated upper motor neuron signs.

We provide, for the first time, evidence for a role of VRK1 in regulating Cajal Body assembly **in Motor Neurons**. The observed Motor Neuron defects are consistent with a length dependent axonopathy affecting lower and upper motor neurons, and we propose that diseases due to mutations in *VRK1* should be grouped under a unique entity named “VRK1-related motor neuron disease”.

Introduction

Hereditary Motor and Sensory Neuropathy (HMSN), or Charcot-Marie-Tooth disease (CMT), are the most common group of Inherited peripheral neuropathies (IPN), with an overall prevalence of 1/2500 (1). These diseases are characterized by length-dependent progressive degeneration of the Peripheral Nervous System (PNS) and extensive phenotypic and genetic heterogeneity, with alterations in over 80 genes reported to date (2, 3). Clinically most patients present with distal weakness and muscle amyotrophy predominant at the lower limbs, resulting in gait difficulties. The onset usually occurs within the first two decades, and the disease then progresses slowly. Progression to the upper limbs is frequent but occurs later in the disease course. Skeletal abnormalities such as pes cavus, reduced or absent deep tendon reflexes are also present, while sensory loss is also often, but not always described. Patients presenting with exclusive motor neuropathy, are affected with distal Hereditary Motor Neuropathy (dHMN), also known as neuronopathy, distal Spinal Muscular Atrophy or spinal CMT. It is not always clear if dHMN must be considered as a separate clinical entity: indeed, many forms of dHMN have minor sensory abnormalities or begin as pure motor neuropathies and progress toward an axonal CMT (CMT2) with time. As a consequence, there is a genetic overlap between CMT2 and dHMN, and mutations in the same gene, even the same mutation in a gene, may cause both phenotypes (4). The same kind of overlap also exists between CMT/dHMN and upper motor neuron diseases, such as hereditary spastic paraplegias (5). Similarly, minor sensory involvement might also be present in close clinical conditions affecting the motor neurons, such as Amyotrophic Lateral Sclerosis (ALS) and spinal muscular atrophy (SMA) (4, 6).

Here, we describe two novel compound heterozygous missense mutations in the Vaccinia-related Kinase 1 gene (*VRK1)*, in two siblings affected with a distal Hereditary

Motor Neuropathy (dHMN), associated with upper motor neuron signs. Since the first description of *VRK1* as the culprit gene in Spinal Muscular Atrophy with Pontocerebellar Hypoplasia (SMA-PCH/PCH1A, MIM 607596) (7), mutations in this gene have been reported in a spectrum of neurological diseases affecting motor neurons (either lower only or lower and upper) (8-11), or motor and sensory nerves (12).

VRK1 encodes the Vaccinia Related kinase 1, a ubiquitously expressed, mainly nuclear, serine/threonine kinase, playing a crucial role in regulating cell cycle (13, 14). Several substrates are known to date for VRK1, including VRK1 (15), the transcription factors p53 (15), c-jun (16), ATF2 (17) and CREB (18), proteins involved in DNA replication and repair, such as histone H2B (15), histone H3 (19) and NBS1 (20), coilin, the main component of Cajal Bodies (CBs) (21), myelin basic protein MBP (15), a component of central myelin and peripheral myelin, or BAF, a protein required for nuclear envelope assembly (22, 23). *VRK1* encodes a 396 amino acid (aa) protein containing a N-terminal serine/threonine kinase domain (residues 37-275) (24) and a C-terminal Nuclear Localization Signal (NLS) (residues 356–360). VRK1 is mainly a nuclear protein, although the presence of a small fraction in the cytoplasm and membrane compartments has been described (24, 25).

In this study, we describe bi-allelic mutations in *VRK1* in a new form of distal Hereditary Motor Neuropathy associated with upper motor neuron signs, leading to decreased VRK1 expression. We show that loss of VRK1 likely contributes to altered coilin dynamics in fibroblasts. In patient's human induced Pluripotent Stem Cells (hiPSC)-derived motor neurons (MNs), we describe altered neurite length and branching and depletion of Cajal Bodies (CBs). The loss of CBs, center for the assembly of ribonucleoproteins required for splicing, suggests that VRK1 has a major role in regulating splicing activity in motor neurons.

Results

Subjects

Two siblings (patients II.2 and II.3), aged 47 and 38 years respectively, from a Lebanese non consanguineous family presented with distal Hereditary Motor Neuropathy (dCMT/dHMN) associated with upper motor neuron signs (Fig. 1A). Clinically, both patients share slowly progressive distal muscular weakness and atrophy of the four limbs, which started around age 10 for both patients. At age 38, patient II.3 is slightly more severely affected than his sister (II.2) and uses sometimes a walking stick. Other features include hyperreflexia and Babinski bilateral signs. Physical examination showed no sensory component. Brain and spinal cord MRI performed at age 34 in patient II.3 showed no abnormalities (Fig. 1B). Cranial Nerves were normal. Nerve conduction studies revealed some drop in the amplitudes of the compound muscle action potentials (CMAPs) of all motor nerves in the upper and lower extremities with mild slowing in their conduction velocities (Table 1). The electromyography (EMG) revealed mild neurogenic motor units action potentials with preserved recruitment after insertion of a bipolar needle electrode (Table 1). These findings are suggestive of mild peripheral motor neuropathy. Genealogical data indicate an autosomal recessive mode of inheritance (Fig. 1A).

Identification of compound heterozygous mutations in VRK1

Although we searched in priority for apparently compound heterozygous variants shared by Patient II.2 and Patient II.3, a search for homozygous variants was performed, which showed the absence of shared homozygous variants between the two sibs (Supp. Table 1). In contrast, WES sequencing led to the identification of 4265 possibly compound heterozygous variants, in the coding sequences of genes shared by the two sibs, of which 134 had a frequency above

1% in the gnomAD dataset (http://gnomad.broadinstitute.org/). Additional filtering narrowed down the number of candidates to 10 variants in 4 genes (*LAMA1, TTN, EVPL and VRK1*). Details of the filtering steps are available in Supplementary Table 1. *EVPL* and *TTN* variants have been eliminated by segregation studies and *LAMA1* variants were also excluded because one of the two variants was described 3 times at the homozygous state in the gnomAD dataset. We focused on two compound heterozygous variants in the *VRK1* gene, because of their complete absence from the gnomAD dataset and a prediction for pathogenicity by several prediction tools. Moreover, at that time, *VRK1* was described as the causative gene for Spinal Muscular Atrophy with Pontocerebellar Hypoplasia (PCH1A, MIM 607596) (7), a neurological disease, sharing clinical signs with the disease affecting our patients. *VRK1* (NM_003384) comprises 13 exons and encodes a 396 amino acid (aa) protein (NP_003375), containing a N-terminal serine/threonine kinase domain (residues 37-275) (24) and a C-terminal Nuclear Localization Signal (NLS) (residues 356–360). Both variants are novel missense heterozygous transversions: c.656G>T (p.Arg219Ile) and c.761G>T (p.Trp254Leu) in *VRK1* exons 8 and 9 respectively. Segregation studies confirmed that the mutations are, indeed, compound heterozygous, and segregate with the disease in the family (Fig. 1A and Fig.1C). Both the arginine at aa 219 and the Tryptophan at aa 254 are highly conserved in vertebrates (Fig. 1E). RT-PCR performed between exons 7 and 9 of *VRK1* in both patients showed that both mutations are true missense variations with no effect on splicing (data not shown). Finally, we showed, in mouse, that *vrk1* is expressed in the tissues affected in our patients' disease: the central (spinal cord) and peripheral (sciatic nerve) mouse nervous system (Fig. 1D). In consequence, we propose that the two novel missense variations c.656G>T (p.Arg219Ile) and c.761G>T (p.Trp254Leu) in VRK1 (Clinvar Accession number SCV000882440 and SCV000882740 respectively) are the molecular defects underlying this

particular form of distal hereditary motor neuropathy associated with upper motor neuron signs, bringing to twelve the number of mutations described to date in *VRK1* (Fig. 1F).

Decrease of VRK1 levels in patients' cells due to post-translational defects

In order to further demonstrate the pathogenicity of the mutations, we studied the expression of VRK1 in patients' cells by western blot and qRT-PCR. We demonstrated a decrease in VRK1 protein levels in patients as compared to controls in both lymphocytes and fibroblasts: 59% and 68% decrease respectively (Fig. 2A-B). To test whether this decrease is a result of an impairment of the transcriptional process, VRK1 mRNA levels were measured by qRT-PCR: no significant changes in VRK1 transcript levels were observed between controls and patients (Fig. 2C). These results evidenced that the decrease in VRK1 levels observed in patients cells is due to a post-translational defect. Interestingly, blockade of the proteasome by MG132 treatment increased VRK1 protein levels by 44% in patients' fibroblasts, showing that the decrease in VRK1 levels in patients' cells is due to its degradation by the proteasome (Fig. 2D).

Mislocalization of VRK1 in patients' fibroblasts

Although small fractions of VRK1 can be present in the cytoplasm and other cytoplasmic compartments, VRK1 is mainly a nuclear protein. Protein subcellular localization was evaluated in patients and controls' fibroblasts: three possible VRK1 locations were observed (Fig. 3). Interestingly, VRK1 was mislocalized in more than 70% of analyzed patients' cells. In the remaining 30%, VRK1 was either both in the nucleus and the cytoplasm or completely absent. However, VRK1 was never strictly nuclear, unlike in 50% of controls' cells. Nuclear localization was restored in at least 97% of patients' fibroblasts after MG132 treatment

suggesting that the missense mutations lead to the export of VRK1 to the cytoplasm for degradation by the proteasome (Fig. 3).

VRK1 depletion in patients' fibroblasts leads to reduced coilin levels by facilitating its proteasomal degradation

We analyzed the effect of the absence of VRK1 from the nucleus of patients' fibroblasts on coilin expression. We observed a significant decrease of coilin levels in patients' cells as compared to controls' (Fig. 4A-B). Having demonstrated above that MG132 blockade of the proteasome is able to restore VRK1 levels and nuclear localization; we checked whether it would also restore coilin levels. Indeed, MG132 treatment of patients' fibroblasts resulted in the restoration of normal coilin levels (Fig. 4). Therefore, we conclude that in the absence of VRK1, coilin is likely degraded in the proteasome.

VRK1 is necessary for CBs assembly in hiPSC-derived MNs

Given the upper and lower MN signs described in our patients, we sought to investigate the effects of VRK1 and coilin abnormalities in a more relevant cellular model, and therefore, we produced spinal MNs by differentiation of hiPSCs from patients II.2 and one control hiPSC-derived MNs. Differentiation efficiency was assessed, in each hiPSC cell line, by immunostaining of HB9 and ISLET1, two transcription factors of spinal MNs. As shown in Supplementary Fig. 1, we observed no differences in the differentiation efficiency between patient II.2 and the control. Also, there were no differences in the iPS proliferation rates between patient II.2 and control (data not shown). In hiPSC-derived MNs, immunostaining and western blot experiments confirmed the decrease of VRK1 levels in patient versus control (Fig. 5A). Quantification by western blot evidenced a 65% decrease of VRK1 levels in patient II.2 as compared to control (Fig. 5B-C). Interestingly, we showed that this decrease

is accompanied by disassembly of CBs in patient's MNs: indeed, while in controls, more than 80% of MNs have nicely assembled CBs, in patient II.2, CBs are present in only 20% of MNs. Moreover, when present, CBs have a smaller size than control's (Fig. 5D-E).

Impaired neurite length and branching in patient's hiPSC-derived MNs

In order to further investigate the effects of reduced VRK1 levels, and the resulting CB loss and disruption, on the function of MNs, we studied neurite length and branching in hiPSC-derived MNs from patient II.2 as compared to a control. To this end, patient and control's hiPSC-derived MNs were infected by low levels of AAV 2/6-CMV-GFP virus. Indeed the use of low virus titers led to weak infection and GFP expression by few MNs, thereby allowing tracing of GFP-expressing neurites in isolated MNs. Interestingly, patient's MNs exhibited a significant decrease in: i) total neurite length, ii) number of branching points, and, iii) number of branches per cell, as compared to control (Fig .5F-I).

Discussion

In this work, we identified two novel compound heterozygous missense mutations in *VRK1*, the vaccinia related kinase 1 gene, in two patients from a non-consanguineous Lebanese family, affected with a slowly-progressive distal form of CMT (dHMN) associated with upper motor neuron signs. The mutations, c.656G>T (p.Arg219Ile) and c.761G>T (p.Trp254Leu), located in the kinase domain of VRK1, segregate in the pedigree, following an autosomal recessive mode, are absent from reference datasets, are well conserved among species, and are considered to be damaging by pathogenicity prediction algorithms (Fig. 1).

Mutations in *VRK1* were first described in spinal muscular atrophy with pontocerebellar hypoplasia (SMA-PCH) (7), but then proved to be responsible for several other neurological diseases, affecting motor neurons (lower or lower and upper) (8-11), with or without sensory abnormalities (12). In total, including ours, 12 mutations are now reported in *VRK1* in 14 patients from 10 families (Fig. 1), described under 5 different clinical entities (Table 2). There is no clear phenotype-genotype correlation, although the p.Arg358* mutation is often associated with more severe phenotypes, in particular when it is found at the homozygous state. The presence of pontocerebellar hypoplasia (PCH), first described as a cardinal clinical sign of the disease, should however not be seen as a distinctive clinical manifestation, because most patients do not have PCH. The same is true for microcephaly, which has been described in only 5 patients from 3 families (Table 2), of whom 3 are homozygous for the p.Arg358* mutation. In contrast, the involvement of upper motor neuron is almost constant, as manifested by brisk tendon reflexes, even in patients described as affected with distal SMA (Table 2).

Although the phenotype that we describe here is not new, this is the first time that patients with VRK1 mutations present with a classical slowly progressive motor neuropathy, beginning in the second decade of life and associated to upper motor neuron involvement. The description of this more classical "dHMN" phenotype, further expands the range of diseases related to defects in *VRK1*, and underlines the heterogeneity of phenotypes resulting from VRK1 mutations, as emphasized by the diversity of clinical terms used to describe the disease and its progression: spinal muscular atrophy (SMA), distal SMA, Amyotrophic Lateral Sclerosis (ALS), juvenile ALS, Hereditary Motor and Sensory Neuropathy (HMSN), pure motor neuropathy (distal HMN). Considering the possible overlap among all these diseases, and the fact that the feature, common to all patients, is the involvement of lower motor neurons, we propose that patients with mutations in *VRK1* should be grouped under a

unique entity named “VRK1-related motor neuron disease”, and that the diagnosis might be evoked by slowly progressive distal motor neuropathy associated or not with upper motor neuron signs.

VRK1 encodes an ubiquitously expressed, mainly nuclear, serine/threonine kinase, from the vaccinia-related kinase (VRK) family, playing a crucial role in many cellular processes like cell division and cell cycle progression (13, 14). In nervous tissues, VRK1 has been described to be expressed in fetal and adult brain (26). Additionally, we show that *vrk1* is expressed in mouse tissues from the central (brain, cerebellum and spinal cord) and peripheral nervous system (sciatic nerve) (Fig. 1C), the latter two being affected in patients with *VRK1* mutations. While the role of *VRK1* in cell cycle and cellular proliferation has been widely studied, its function in the nervous system has only been addressed in one study, where Vrk1 has been shown to be important for neuronal migration, through amyloid β precursor protein (APP)-dependent mechanisms (27). Here, we show that the two identified missense mutations lead to a significant decrease (more than 50%) of VRK1 levels in patients in all tested cell types (fibroblasts, immortalized lymphoblastoid cells and hiPSC-derived MNs,thereafter termed MNs) due to post-translational degradation of the mutated proteins (Fig. 2A and Fig. 5B). Additionally to decreased protein levels, we also demonstrate that the remaining mutated VRK1 proteins are mislocalized, and shift from the nucleus to the cytoplasm (Fig. 3), although this phenomenon is less striking in hiPSC-derived MNs (Fig. 5A). Both protein levels and nuclear localization, are restored to normal after MG132 treatment (Fig. 2C to E), thereby suggesting that VRK1 is likely transported to the cytoplasm for ubiquitination and degradation in the proteasome. Knowing that VRK1 is a nuclear kinase, the absence, or reduced levels, of VRK1 from the nucleus in patients’ cells strongly suggest abnormal phosphorylation of VRK1 nuclear targets. Several substrates are known to date for VRK1, including VRK1 (15), the transcription factors p53 (15), c-jun (16), ATF2

(17) and CREB (19), proteins involved in DNA replication and repair, such as histone H2B (15), histone H3 (19) and NBS1 (20), myelin basic protein MBP (15), a component of central myelin and peripheral myelin, BAF, a protein required for nuclear envelope assembly (22) and coilin (21). Among all VRK1 phosphorylation targets, we focused on coilin, the main component of Cajal Bodies (CBs), due to the fact that defects in CBs' size and quantity have been described in motor neurons from patients with Spinal Muscular Atrophy (SMA, MIM 253300) (28). CBs were first discovered by Ramón y Cajal in 1903 in pyramidal neurons from the cerebral cortex. They are dynamic, membrane-free, nuclear organelles enriched in several nuclear proteins and RNA-protein complexes (29), and playing an important role in RNA processing, particularly splicing, by guiding the modification of the snRNA moiety of snRNP proteins (small nuclear ribonucleoproteins) from the spliceosome (30, 31). CBs also participate in the biogenesis and delivery of telomerases to telomeres (32). CBs assemble on coilin, which acts as a scaffold protein, but they do not exist in all cells. Indeed, CBs are most prominent in cells that are transcriptionally active (33), such as post-mitotic neurons. In dividing cells, dynamic assembly/disassembly of CBs is regulated, during cell cycle, by post-translational modifications, including phosphorylation (34, 35).

VRK1 being a nuclear kinase, also regulated during cell cycle progression, it is a good candidate to regulate CBs dynamics (14). Indeed, recent work has suggested that VRK1 regulates CBs dynamics during cell cycle, by phosphorylating coilin and protecting coilin from ubiquitination and degradation in the proteasome (34). In the same study, the authors showed that depletion of VRK1 by knockdown in three cell lines (Hela, MCF7 and SH-SY5Y) resulted in a drop of nuclear coilin and subsequently in the loss of CBs; and that coilin's proteasomal degradation is prevented by VRK1 (34). Here, we confirm that VRK1 is crucial for coilin expression and stability. Indeed, in dividing cells (fibroblasts) of patients with bi-allelic mutations in *VRK1*, we show that the resulting depletion of VRK1 from the

nucleus leads to reduced coilin levels by facilitating its proteasomal degradation (Fig.4). However, in these cells, as expected in primary cells (36), coilin is not assembled in CBs, even in normal conditions. In order to investigate CBs organization and function in the disease context, independently of cell-cycle stage and ploidy, we developed an *in vitro* model of patient's MNs, by producing hiPSCs-derived spinal MNs mimicking post-mitotic MNs from the spinal cord. Most importantly, we show that, while CBs are well assembled in control MNs, decreased VRK1 levels lead to a disintegration of CBs in patients MNs. Moreover when CBs are present, they have a smaller size (Fig. 5) suggestive of their abnormal assembly. Several studies have shown that CBs number correlates with neuronal size and global transcriptional activity (33). For the first time, we bring evidence that VRK1 is necessary for CBs assembly in human Motor Neurons. Although we provide no experimental evidence that coilin phosphorylation is altered in patients, our results strongly suggest that the remaining mutant VRK1 proteins in our patient's MNs are not sufficient to maintain the phosphorylation levels of coilin required for CBs assembly.

Considering the fact that CBs number positively correlates with global transcriptional activity and the cellular demand for pre-mRNA and pre-rRNA processing (29, 33), our results strongly suggest that altered RNA metabolism due to CBs depletion is an essential component of the pathophysiological mechanisms of this VRK1-related disease. In post-mitotic neurons, RNA metabolism is essential to maintain metabolic and electrical activity, and indeed, we demonstrate that the patient's hiPSC-derived MNs have shorter neurite length and altered branching, which are consistent with a length dependent axonopathy affecting both lower and upper motor neurons.

There are few studies about the contribution of CBs to the pathophysiology of neurological disorders, however, the depletion of CBs and the presence of smaller CBs, that we describe here in VRK1 patients' MNs, are also hallmarks of Spinal Muscular Atrophy (SMA, MIM

253300) (28), another motor neuron disease (lower MN) caused by loss or mutations in the Survival Motor Neuron (SMN) protein. SMN is part of the SMN complex, which localizes to the nucleus, where it accumulates in Cajal Bodies (CB). The SMN protein shuttles between the cytoplasm and the nucleus, and its subcellular localization is regulated by complex regulatory cues mediated by phosphorylation of serine/threonine and tyrosine residues (37). Knowing the function of SMN in pre-mRNA splicing, it is therefore tempting to speculate that VRK1 might be involved in the regulation of SMN function through its nuclear phosphorylation and that the defects observed in our patient's MNs are due to impaired RNA metabolism.

In conclusion, we describe new VRK1 mutations in two sibs affected with distal Hereditary Motor Neuropathy associated with upper motor neuron signs. We provide, for the first time, evidence for a role of VRK1 in regulating CBs assembly **in Motor Neurons**. Many questions remain to be addressed regarding the interactions between components of the CBs, such as coilin and SMN, but our study provides strong evidence that VRK1 is a crucial protein in maintaining CBs integrity in motor neurons and might be a key factor in pathophysiology of lower and/or upper motor neuron diseases. Taken together, our results strongly suggest that defective RNA metabolism might be a unifying theme in diseases affecting motor neurons, such as SMA and the one that we describe here.

Targeting VRK1 in motor neuron diseases might therefore open up new research areas for therapy of these diseases.

Material and Methods

Genetic analyses

Samples

After informed consent was obtained from all individuals included in this study, EDTA blood samples were collected, and genomic DNA was extracted from lymphocytes with the use of standard methods. All protocols performed in this study complied with the ethics guidelines of the institutions involved.

Whole Exome Sequencing (WES)

Whole exome sequencing was carried out on both patients II.2 and II.3 (Fig. 1A). Library preparation, capture and sequencing were performed by the French National Genotyping Center (CNG, Evry, France), using the *in solution* Agilent SureSelect Human All Exon kit v3.0 (38) and then sequenced on an Illumina HiSeq2000, using a paired-end 100-bp read sequencing protocol. Image analysis and base calling were performed using the Illumina Data Analysis Pipeline Software 1.5 with default parameters. Raw data were mapped to the built of the human genome (hg19) by using BWA 0.7.5 (39). Variant calling was subsequently performed using GATK (40) and annotation was done with ANNOVAR (41). All subsequent steps were performed using our *in-house* software for variant annotation and segregation VarAFT (42). WES data from both patients were analyzed simultaneously and segregated using this tool. Variants were filtered based on an autosomal recessive inheritance pattern: considering the absence of consanguinity, we searched in priority for apparently compound heterozygous mutations shared by the two patients, although a search for shared homozygous variants was also performed. The analysis was performed in a two-phased approach: (i) analysis of a list of 88 genes known to be implicated in IPNs; (ii) analysis of the WES data. In order to refine the obtained lists of candidates, filtering was performed by removing all variants with a frequency above 1% on the Genome Aggregation Database (gnomAD,

http://gnomad.broadinstitute.org/). Additional filtering was then performed using several frequency datasets: the Greater Middle East (GME) Variome (http://igm.ucsd.edu/gme/), and an *in-house* exome database. Finally, manual filtering was performed by removing variants with frequencies above 1% in gnomAD subpopulations, or variants present at the homozygous state in gnomAD.

In order to predict the deleterious effect of the identified sequence variations, different bioinformatics tools were applied: MutationTaster (43), SIFT (44), PolyPhen-2 (45), UMD predictor (46) and CADD (47).

Segregation Analysis by capillary Sanger Sequencing

Candidate variants identified by WES sequencing in patients were tested by Sanger sequencing of PCR amplified fragments in family members, for whom DNA was available: the mother II.1, the two patients II.2 and II.3, and the two unaffected sibs II.1 and II.4 (Fig.1). Genomic and cDNA sequences of the candidate gene (*VRK1*, Genbank accession number: NM_003384) were obtained from the UCSC Genomic Browser, February 2009 human reference sequence (GRCh37 https://genome.ucsc.edu/). Primers used for PCR amplification were designed using Primer3 software (http://frodo.wi.mit.edu) to amplify the region surrounding the candidate DNA variations (available upon request). PCR products were purified by mixing with an equal volume (10ul) of AMPure beads (Beckman Coulter, USA) according to the manufacturer's instructions. Both strands were sequenced as described in Jobling *et al*. (48). Electrophoregrams were aligned to the reference sequence using Sequencher v5.4.6 (Genecodes, USA).

Reverse Transcriptase PCR (RT-PCR) and quantitative Real-Time PCR (qRT-PCR)

Total RNA was extracted from cells using Pure Link RNA Minikit (Thermo Fisher Scientific, USA) and treated with DNAse I, on-column, according to the manufacturer's instructions. Complementary DNAs (cDNA) were synthesized from total RNA using the SuperScript III Reverse transcriptase and random hexamers (Thermo Fisher Scientific, USA) according to the manufacturer's protocol. For RT-PCR, specific primers were designed through the Primer3 software in order to detect the three possible isoforms of mouse *vrk1* (available upon request). qRT-PCR experiments were performed on total RNA from patients and controls' cells to measure *VRK1* mRNA levels. Experiments were performed in triplicate on a Lightcycler 480 (Roche, Basel, Switzerland) using the Luminaris Color HiGreen qPCR Master Mix (Thermo Fisher Scientific, USA). 50ng of RNAse H-treated cDNA were amplified in a 10µl reaction mixture containing primers at a final concentration of 0.3µM. For each sample, the expression levels of *VRK1* were normalized to three different endogenous genes: *GAPDH*, *GUS-B* and *B2*M. Relative expression of *VRK1* was calculated as fold-differences using the comparative C_T method ($\Delta\Delta C_T$). Two calibrator samples were used for each sample. Primer sequences are available upon request.

Cellular studies

Cell culture

Lymphoblastoid cells

Immortalized lymphoblastoid cell lines from patients and controls were prepared at the accredited Biological Resource Center (CRB TAC), Department of Medical Genetics, Timone Hospital of Marseille, from Acid Citrate Dextrose (ACD) blood samples transformed with Epstein-Barr virus using standard procedures. All used cell lines are stored at CRB TAC and belong to a biological sample collection declared to the French ministry of Health

(declaration number DC-2008-429) whose use for research purposes was authorized by the French ministry of Health (authorization number AC-2011-1312 and AC-2017-2986). Human lymphoblastoid cells were grown as described in Jobling *et al*. (48).

Fibroblasts

Control fibroblast cell lines were purchased from the Coriell Cell Repository (Coriell Institute for Medical Research, Philadelphia, USA): AG08471 (male, foreskin fibroblasts) and AG13091 (40 years old male, foreskin fibroblasts). Patients' fibroblasts II.2 and II.3 were prepared from a skin biopsy and stored by the CRB TAC, Department of Medical Genetics, Timone Hospital of Marseille according to the French regulation. Primary human fibroblasts were cultured as described in Jobling *et al*. (48).

hiPSCs maintenance and differentiation into Spinal Motor Neurons

Reprogrammation of skin fibroblasts from patient II.2 (Fig.1) into induced Pluripotent Stem Cells (iPSC) was performed by our Cell Reprogramming and Differentiation Facility at U 1251/Marseille Medical Genetics (Marseille, France), using retroviral transduction of Oct4, Sox2, Klf4, and c-myc protocols (49). The Control iPSC cell line was provided by the Cell Reprogramming and Differentiation Facility. All hiPSCs included in the study are declared in a collection authorized by the competent authorities in France (declaration number DC-2018-3207). For each cell line, we used classical quality control criteria, including expression of pluripotency markers, absence of expression of the reprogramming transgenes, phosphatase se alkaline staining and karyotype for checking the absence of chromosomal abnormalities. Expression of several pluripotency markers was verified by flow cytometry: keratin Sulfate antigens Tra1-60 and Tra1-81 and the glycolipid antigens SSEA3 and SSEA4 for the control hiPSC cell line, or intracellular pluripotency markers Oct3/4, Sox2 and Nanog for Patient II.2

hiPSC cell line. Pluripotency of hiPSC cell lines was also assessed by an *in vitro* embryoid body (EB) assay (50), where the formation of the three developmental germ layers (ectoderm, endoderm, and mesoderm) was tested by qRT-PCR analysis of germ layer-specific genes (*FOXA2, AFP, TBXT, NCAM, COL2A1, ALDhA1*) in the formed EBs at day 15.

The presence of the c.656G>T (p.Arg219Ile) and c.761G>T (p.Trp254Leu) mutations in *VRK1* has been checked, in the patient II.2 hiPSC clone, after reprogrammation, by fluorescent Sanger Sequencing.

iPSC clones were maintained and expanded on matrigel (BD biosciences)-coated dishes in mTeSR1 medium (STEM CELL Technologies 05851) under standard procedures.

For differentiation of iPSCs into Spinal MNs, we used an established 30-days differentiation protocol based on early activation of the Wnt signaling, coupled to activation of the Hedgehog pathway and inhibition of Notch signaling (51). Using this protocol, we obtained mature MNs, which express acetylcholine transferase and can fire Action Potentials (unpublished data). Differentiation efficiency was assessed by HB9 and Islet1 immunostaining as described in Maury *et al*. (51). Both iPS and differentiated MNs were tested for mycoplasma contamination every other week.

Protein extraction and Immunoblotting

Protein extraction and immunoblotting were performed from fibroblasts or lymphocytes pellets as mentioned in Jobling *et al.* (48). Primary antibodies were: rabbit polyclonal antibody to VRK1 (SIGMA-ALDRICH, #HPA017929) and GAPDH (Santa-Cruz Biotechnology, #sc-48167), diluted at 1:200 and 1:1000 respectively. GAPDH served as a loading control. Secondary antibodies were donkey anti-rabbit IRDye 800 and donkey anti-goat IRDye 680, (Li-Cor Biosciences), diluted at 1/10000.

Immunostaining and Microscopy

Conditions for immunostaining have been described in Jobling *et al.* (48). VRK1 was detected with a rabbit polyclonal antibody (1:200) (SIGMA-ALDRICH, #HPA017929), coilin with mouse Pdelta monoclonal antibody (1:50) (Santa-Cruz Biotechnology, #sc-56298), Neurofilament M with chicken polyclonal antibody (1:1000) (Covance, # PCK-593P). The following secondary antibodies were used: Goat Anti-Rabbit IgG H&L (Alexa Fluor®488) ab150077 (abcam, UK) at 1:400 dilution, donkey Anti-Mouse IgG H&L (DyLight® 550) ab96876 (abcam, UK) at 1:400 dilution, Donkey Anti-Rabbit IgG H&L (DyLight® 550) ab96892 (abcam, UK)) at 1:400 dilution, Goat Anti-Chicken IgY H&L (Alexa Fluor® 488) ab150169 (abcam, UK) at 1: 1000 dilution and Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647, A-31571 (Thermo Fisher Scientific, USA) at 1:1000 dilution.

Proteasome blockade

Fibroblasts from patients and controls were cultured with culture medium containing MG132 (474790, Merck Chemical LTD) at a final concentration of 1.5 μM, or the same volume of vehicle (DMSO, 0.025% v/v) for 1 h at 37°C and 5% CO2.

Constructs and production of AAV vectors

Adeno-associated-virus (AAV)2/6 vectors were produced by the Bertarelli platform for Gene Therapy at EPFL (Lausanne, Switzerland). Production and titration of AAV2/6-CMV-GFP was performed as described by Dirren et *al.* (52).

Quantification of neurite length and branching in GFP+ hiPSC-derived infected MNs

Quantification of neurite length and branching was performed in GFP positive-hiPSC-derived MNs after infection with AAV2/6-CMV-GFP (2 TU/cell). After fixation of the cells, GFP signal was amplified by immunohistochemistry using a goat anti-GFP antibody (Abcam, ab5449) following incubation with an anti-goat secondary antibody (Abcam, ab96931). Fluorescence was observed under a Zeiss ApoTome.2 fluorescence microscope with a 10x objective. Neuronal processes were then analyzed using the plugin NeuronJ on ImageJ software. We measured total neurite length and branching: i) neurite length corresponds to the sum of the length of all neuritic branches per cell; ii) a branch is defined as a neurite segment starting at a branching point or at the cell body, and ending at the next branching point or at the extremity of the neurite processes; iii) branching points were defined when a secondary or tertiary neurite branch diverges from the primary path.

Statistical analyses

Statistical analyses were performed using unpaired two-tailed Student's t-test, a one-way or a two-way ANOVA test for parametric data, and a Mann-Whitney test for non-parametric data, as appropriate. Details of the used test are described in the legends of the figures. Data are reported as mean ± SEM. Significance was accepted as the level of $P < 0.05$ (*$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$).

Acknowledgements

We would like to thank the patients and families for their kind cooperation and their participation in this study. We thank the Bertarelli platform for Gene Therapy at EPFL (Lausanne, Switzerland) for viral vector production of AAV-CMV-GFP. We thank Claire El Yazidi and Morgane Thomas, from the Cell Reprogramming and Differentiation Facility at U

1251/Marseille Medical Genetics (Marseille, France) for reprogrammation and characterization of patient's II.2 fibroblasts into iPSC cell line and for providing control iPSC lines. We thank Dr Antoine Salloum for performing skin biopsies.

This study was partly sponsored by the French Association against Myopathies "Association Française contre les Myopathies" (AFM), the "Agence Universitaire de la Francophonie" (AUF) and the A.MIDEX foundation. Lara El-Bazzal was supported by a grant from the AFM and by a fellowship from the HERMES Programme of the European Union's Seventh Framework Programme for research, technological development and demonstration.

Conflicts of interest statement

None

References

1 Skre, H. (1974) Genetic and clinical aspects of Charcot-Marie-Tooth's disease. *Clin. Genet.*, **6**, 98-118.

2 Baets, J., De Jonghe, P. and Timmerman, V. (2014) Recent advances in Charcot-Marie-Tooth disease. *Curr. Opin. Neurol.*, **27**, 532-540.

3 Jerath, N.U. and Shy, M.E. (2015) Hereditary motor and sensory neuropathies: Understanding molecular pathogenesis could lead to future treatment strategies. *Biochim. Biophys. Acta Mol. Cell. Res.*, **1852**, 667-678.

4 Irobi, J., Dierick, I., Jordanova, A., Claeys, K.G., De Jonghe, P. and Timmerman, V. (2006) Unraveling the genetics of distal hereditary motor neuronopathies. *Neuromol. Med.*, **8**, 131-146.

5 Timmerman, V., Clowes, V.E. and Reid, E. (2013) Overlapping molecular pathological themes link Charcot-Marie-Tooth neuropathies and hereditary spastic paraplegias. *Exp. Neurol.*, **246**, 14-25.

6 Rossor, A.M., Polke, J.M., Houlden, H. and Reilly, M.M. (2013) Clinical implications of genetic advances in Charcot-Marie-Tooth disease. *Nat. Rev. Neurol.*, **9**, 562-571.

7 Renbaum, P., Kellerman, E., Jaron, R., Geiger, D., Segel, R., Lee, M., King, M.C. and Levy-Lahad, E. (2009) Spinal muscular atrophy with pontocerebellar hypoplasia is caused by a mutation in the VRK1 gene. *Am. J. Hum. Genet.*, **85**, 281-289.

8 Nguyen, T.P., Biliciler, S., Wiszniewski, W. and Sheikh, K. (2015) Expanding Phenotype of VRK1 Mutations in Motor Neuron Disease. *J. Clin. Neuromuscul. Dis.*, **17**, 69-71.

9 Stoll, M., Teoh, H., Lee, J., Reddel, S., Zhu, Y., Buckley, M., Sampaio, H., Roscioli, T., Farrar, M. and Nicholson, G. (2016) Novel motor phenotypes in patients with VRK1 mutations without pontocerebellar hypoplasia. *Neurology*, **87**, 65-70.

10 Tomaselli, P.J., Rossor, A.M., Horga, A., Laura, M., Houlden, H. and Reilly, M.M. (2016) A Compound Heterozygous Mutation in the Vaccinia Related Kinase-1 Gene Is a Cause of Hereditary Motor Neuropathy with Upper Motor Neuron Signs. *J. Peripher. Nerv. Syst.*, **21**, 305-306.

11 Li, N., Wang, L., Sun, X., Lu, Z., Suo, X., Li, J., Peng, J. and Peng, R. (2019) A novel mutation in VRK1 associated with distal spinal muscular atrophy. *J. Hum. Genet.*, **64**, 215-219.

12 Gonzaga-Jauregui, C., Lotze, T., Jamal, L., Penney, S., Campbell, I.M., Pehlivan, D., Hunter, J.V., Woodbury, S.L., Raymond, G., Adesina, A.M. *et al.* (2013) Mutations in VRK1 associated with complex motor and sensory axonal neuropathy plus microcephaly. *JAMA Neurol.*, **70**, 1491-1498.

13 Valbuena, A., Sanz-García, M., López-Sánchez, I., Vega, F.M. and Lazo, P.A. (2011) Roles of VRK1 as a new player in the control of biological processes required for cell division. *Cell. Signal.*, **23**, 1267-1272.

14 Valbuena, A., López-Sánchez, I. and Lazo, P.A. (2008) Human VRK1 is an early response gene and its loss causes a block in cell cycle progression. *PLoS ONE*, **3**, e1642.

15 Lopez-Borges, S. and Lazo, P.A. (2000) The human vaccinia-related kinase 1 (VRK1) phosphorylates threonine-18 within the mdm-2 binding site of the p53 tumour suppressor protein. *Oncogene*, **19**, 3656-3664.

16 Sevilla, A., Santos, C.R., Barcia, R., Vega, F.M. and Lazo, P.A. (2004) c-Jun phosphorylation by the human vaccinia-related kinase 1 (VRK1) and its cooperation with the N-terminal kinase of c-Jun (JNK). *Oncogene*, **23**, 8950-8958.

17 Sevilla, A., Santos, C.R., Vega, F.M. and Lazo, P.A. (2004) Human vaccinia-related kinase 1 (VRK1) activates the ATF2 transcriptional activity by novel phosphorylation on Thr-73 and Ser-62 and cooperates with JNK. *J. Biol. Chem.*, **279**, 27458-27465.
18 Kang, T.-H., Park, D.-Y., Kim, W. and Kim, K.-T. (2008) VRK1 phosphorylates CREB and mediates CCND1 expression. *J. Cell Sci.*, **121**, 3035-3041.
19 Kang, T.-H., Park, D.-Y., Choi, Y.H., Kim, K.-J., Yoon, H.S. and Kim, K.-T. (2007) Mitotic histone H3 phosphorylation by vaccinia-related kinase 1 in mammalian cells. *Mol. Cell. Biol.*, **27**, 8533-8546.
20 Monsalve, D.M., Campillo-Marcos, I., Salzano, M., Sanz-García, M., Cantarero, L. and Lazo, P.A. (2016) VRK1 phosphorylates and protects NBS1 from ubiquitination and proteasomal degradation in response to DNA damage. *Biochim. Biophys. Acta. Mol. Cell. Res.*, **1863**, 760-769.
21 Sanz-García, M., Vázquez-Cedeira, M., Kellerman, E., Renbaum, P., Levy-Lahad, E. and Lazo, P.A. (2011) Substrate profiling of human vaccinia-related kinases identifies coilin, a Cajal body nuclear protein, as a phosphorylation target with neurological implications. *J. Proteomics*, **75**, 548-560.
22 Nichols, R.J., Wiebe, M.S. and Traktman, P. (2006) The vaccinia-related kinases phosphorylate the N′ terminus of BAF, regulating its interaction with DNA and its retention in the nucleus. *Mol. Biol. Cell*, **17**, 2451-2464.
23 Molitor, T.P. and Traktman, P. (2014) Depletion of the protein kinase VRK1 disrupts nuclear envelope morphology and leads to BAF retention on mitotic chromosomes. *Mol. Biol. Cell*, **25**, 891-903.
24 Nichols, R.J. and Traktman, P. (2004) Characterization of three paralogous members of the Mammalian vaccinia related kinase family. *J. Biol. Chem.*, **279**, 7934-7946.
25 Valbuena, A., López-Sánchez, I., Vega, F.M., Sevilla, A., Sanz-García, M., Blanco, S. and Lazo, P.A. (2007) Identification of a dominant epitope in human vaccinia-related kinase 1 (VRK1) and detection of different intracellular subpopulations. *Arch. Biochem. Biophys.*, **465**, 219-226.
26 Nezu, J., Oku, A., Jones, M.H. and Shimane, M. (1997) Identification of two novel human putative serine/threonine kinases, VRK1 and VRK2, with structural similarity to vaccinia virus B1R kinase. *Genomics*, **45**, 327-331.
27 Vinograd-Byk, H., Sapir, T., Cantarero, L., Lazo, P.A., Zeligson, S., Lev, D., Lerman-Sagie, T., Renbaum, P., Reiner, O. and Levy-Lahad, E. (2015) The Spinal Muscular Atrophy with Pontocerebellar Hypoplasia Gene VRK1 Regulates Neuronal Migration through an Amyloid-β Precursor Protein-Dependent Mechanism. *J. Neurosci.*, **35**, 936-942.
28 Tapia, O., Bengoechea, R., Palanca, A., Arteaga, R., Val-Bernal, J.F., Tizzano, E.F., Berciano, M.T. and Lafarga, M. (2012) Reorganization of Cajal bodies and nucleolar targeting of coilin in motor neurons of type I spinal muscular atrophy. *Histochem. Cell Biol.*, **137**, 657-667.
29 Cioce, M. and Lamond, A.I. (2005) Cajal bodies: a long history of discovery. *Annu. Rev. Cell Dev. Biol.*, **21**, 105-131.
30 Bohmann, K., Ferreira, J., Santama, N., Weis, K. and Lamond, A.I. (1995) Molecular analysis of the coiled body. *J Cell Sci*, **1995**, 107-113.
31 Matera, A.G. and Frey, M.R. (1998) Coiled bodies and gems: Janus or gemini? *Am. J. Hum. Genet.*, **63**, 317.
32 Cristofari, G., Adolf, E., Reichenbach, P., Sikora, K., Terns, R.M., Terns, M.P. and Lingner, J. (2007) Human telomerase RNA accumulation in Cajal bodies facilitates telomerase recruitment to telomeres and telomere elongation. *Mol. Cell*, **27**, 882-889.
33 Lafarga, M., Tapia, O., Romero, A.M. and Berciano, M.T. (2017) Cajal bodies in neurons. *RNA Biol.*, **14**, 712-725.

34 Cantarero, L., Sanz-García, M., Vinograd-Byk, H., Renbaum, P., Levy-Lahad, E. and Lazo, P.A. (2015) VRK1 regulates Cajal body dynamics and protects coilin from proteasomal degradation in cell cycle. *Sci. Rep.*, **5**, 10543.
35 Hebert, M.D. and Poole, A.R. (2017) Towards an understanding of regulating Cajal body activity by protein modification. *RNA Biol.*, **14**, 761-778.
36 Hearst, S.M., Gilder, A.S., Negi, S.S., Davis, M.D., George, E.M., Whittom, A.A., Toyota, C.G., Husedzinovic, A., Gruss, O.J. and Hebert, M.D. (2009) Cajal-body formation correlates with differential coilin phosphorylation in primary and transformed cell lines. *J. Cell Sci.*, **122**, 1872-1881.
37 Husedzinovic, A., Oppermann, F., Draeger-Meurer, S., Chari, A., Fischer, U., Daub, H. and Gruss, O.J. (2014) Phosphoregulation of the human SMN complex. *Eur. J. Cell Biol.*, **93**, 106-117.
38 Sulonen, A.-M., Ellonen, P., Almusa, H., Lepistö, M., Eldfors, S., Hannula, S., Miettinen, T., Tyynismaa, H., Salo, P., Heckman, C. *et al.* (2011) Comparison of solution-based exome capture methods for next generation sequencing. *Genome Biol.*, **12**, R94.
39 Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754-1760.
40 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. *et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.*, **20**, 1297-1303.
41 Wang, K., Li, M. and Hakonarson, H. (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.*, **38**, e164.
42 Desvignes, J.P., Bartoli, M., Delague, V., Krahn, M., Miltgen, M., Beroud, C. and Salgado, D. (2018) VarAFT: a variant annotation and filtration system for human next generation sequencing data. *Nucleic Acids Res.*, **46**, W545-W553.
43 Schwarz, J.M., Rödelsperger, C., Schuelke, M. and Seelow, D. (2010) MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods*, **7**, 575-576.
44 Kumar, P., Henikoff, S. and Ng, P.C. (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.*, **4**, 1073-1081.
45 Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S. and Sunyaev, S.R. (2010) A method and server for predicting damaging missense mutations. *Nat. Methods*, **7**, 248-249.
46 Salgado, D., Desvignes, J.-P., Rai, G., Blanchard, A., Miltgen, M., Pinard, A., Lévy, N., Collod-Béroud, G. and Béroud, C. (2016) UMD-Predictor: A High-Throughput Sequencing Compliant System for Pathogenicity Prediction of any Human cDNA Substitution. *Hum. Mutat.*, **37**, 439-446.
47 Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M. and Shendure, J. (2014) A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.*, **46**, 310-315.
48 Jobling, R.K., Assoum, M., Gakh, O., Blaser, S., Raiman, J.A., Mignot, C., Roze, E., Durr, A., Brice, A., Levy, N. *et al.* (2015) PMPCA mutations cause abnormal mitochondrial protein processing in patients with non-progressive cerebellar ataxia. *Brain*, **138**, 1505-1517.
49 Bohmann, K., Ferreira, J., Santama, N., Weis, K. and Lamond, A.I. (1995) Molecular analysis of the coiled body. *J. Cell Sci.*, **1995**, 107-113.
50 Sheridan, S.D., Surampudi, V. and Rao, R.R. (2012) Analysis of embryoid bodies derived from human induced pluripotent stem cells as a means to assess pluripotency. *Stem Cells Int*, **2012**, 738910.

51 Maury, Y., Come, J., Piskorowski, R.A., Salah-Mohellibi, N., Chevaleyre, V., Peschanski, M., Martinat, C. and Nedelec, S. (2015) Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nat Biotechnol*, **33**, 89-96.
52 Dirren, E., Towne, C.L., Setola, V., Redmond Jr, D.E., Schneider, B.L. and Aebischer, P. (2013) Intracerebroventricular Injection of Adeno-Associated Virus 6 and 9 Vectors for Cell Type–Specific Transgene Expression in the Spinal Cord. *Hum. Gene Ther.*, **25**, 109-120.
53 Najmabadi, H., Hu, H., Garshasbi, M., Zemojtel, T., Abedini, S.S., Chen, W., Hosseini, M., Behjati, F., Haas, S. and Jamali, P. (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature*, **478**, 57.

Legends to figures

Figure 1. Genealogical, clinical and molecular characteristics of patients with *VRK1* mutations.

(**A**) Pedigree from the studied family**.** Affected individuals are shaded in black. (**B**) Normal MRI images of patient II.3: Axial FLAIR (B1); Sagittal T1-weighted image (B2); Sagittal T2-weighted image (B3). (**C**) Chromatograms of patient and control showing the two compound heterozygous mutations identified in *VRK1* (NM_003384): c.656G>T (p.Arg219Ile) and c.761G>T (p.Trp254Leu) in exon 8 and 9 respectively. **(D)** RT–PCR amplification of *vrk1* transcripts from mouse neuromuscular tissues. Expected sizes of amplicons are 168 pb, 228 pb and 300pb for the three *vrk1* isoforms; NM_001029844, NM_001029843 and NM_011705 respectively. **(E)** Alignment of VRK1 amino acid sequences between different species using CLUSTAL 2.1. Arginine 219 and Tryptophan 254 highlighted in green and in red respectively (NP_003375) are highly conserved across vertebrates. **(F)** Schematic representation of the VRK1 gene and protein. The human VRK1 gene is covering a genomic region of 84, 268 bp at chromosome 14q32.2 and is composed of 13 exons. The transcript (NM_003384) is 1745 bp long, with a coding sequence of 1191 bp (exons 2-13), encoding a 396 AA protein (NP_003375). Untranslated regions (UTRs, 5'- and 3'-) are hashed. Kinase domain is highlited in purple and the putative nuclear localization signal (NLS) in blue. All mutations identified to date in *VRK1* are indicated.

(*)Our study, **(1)**Renbaum et al. 2009, **(2)**Najmabadi et al. 2011, **(3)**Gonzaga-Jauregui et al. 2013, **(4)**Nguyen et al. 2015, **(5)**Stoll et al. 2016, **(6)**Tomaselli et al. 2016, **(7)**Li et al. 2019.

Figure 2. Decrease of VRK1 levels in patients' cells due to post-translational defects.

(A) Western blotting of VRK1 in fibroblasts and immortalized lymphoblastoid cells' lysates from controls and patients shows a strong decrease of VRK1 protein levels in patients (both cell types). VRK1 is detected at 45Kda (lower band in fibroblasts, unique band in lymphoblastoid cells) and GAPDH at 37 KDa. **(B)** This decrease was quantified as VRK1/GAPDH ratio. Bars represent the mean of three independent experiments ± SEM. Statistical significance was evaluated by 2way ANOVA test, *p <0.05, **p <0.005. **(C)** qRT-PCR analyses did not show any significant differences in *VRK1* mRNA levels in patients' lymphocytes and fibroblasts versus controls. *VRK1* expression levels were normalized to three different reference genes: *GAPDH*, *GUS-B* and *B2M*. Relative expression of *VRK1* was calculated as fold-differences using the comparative $\Delta\Delta C_T$ Method. Two calibrator samples were used (Ctrl 1F and Ctrl 2F for fibroblasts and Ctrl 1L and Ctrl 2L for lymphocytes), and each experiment was repeated three times. Data are expressed as geometric mean ± SEM of $\Delta\Delta C_T$ values for each sample. Statistical significance was evaluated by one-way ANOVA test, n.s: not significant. **(D)** Western blot shows restoration of normal VRK1 levels in patients' fibroblasts after proteasome blockade by MG132. VRK1/GAPDH levels are quantified in **(E)** before and after MG132 treatment. After MG132 treatment, VRK1 levels in patients are restored to controls' values. The results are statistically significant as evaluated by 2way ANOVA test, **p <0.005, n.s: not significant. The graph represents the mean of three independent experiments ± SEM. NT: Not Treated, Ctrl: Control, Ctrl1F/2F: Fibroblasts Controls, Ctrl1L/2L: Lymphocytes controls.

Figure 3. Mislocalization of VRK1 in patients' fibroblasts is corrected by proteasome blockade.

(A) Immunolabeling of VRK1 in fibroblasts from patient II.2 shows a shift of VRK1 from the nucleus to the cytoplasm as compared to control (Control 1F). Nuclear localization is restored in patient's fibroblasts after MG132 treatment. Arrows, arrowheads and cross signs indicate examples of strictly nuclear, strictly cytoplasmic, nuclear and cytoplasmic localization respectively. **(B)** These results have been quantified by measurement of R, the ratio of fluorescence intensity of nuclear VRK1/cytoplasmic VRK1, by ImageJ software. MG132 treatment induces relocalization of VRK1 to the nucleus in patients' cells as attested by the drastic decrease of R to values below 0.9 in both patients II.2 and II.3. Three range values were determined for R, corresponding to three possible VRK1 subcellular localizations: ($R>2$) = strictly nuclear, ($0.9<R<2$) = nuclear and cytoplasmic, ($0.1<R<0.9$) = strictly cytoplasmic. Mean percentages of cells in each range are presented in the table for controls and patients. For each individual, we scored 300 cells from three independent experiments. NT: Not treated.

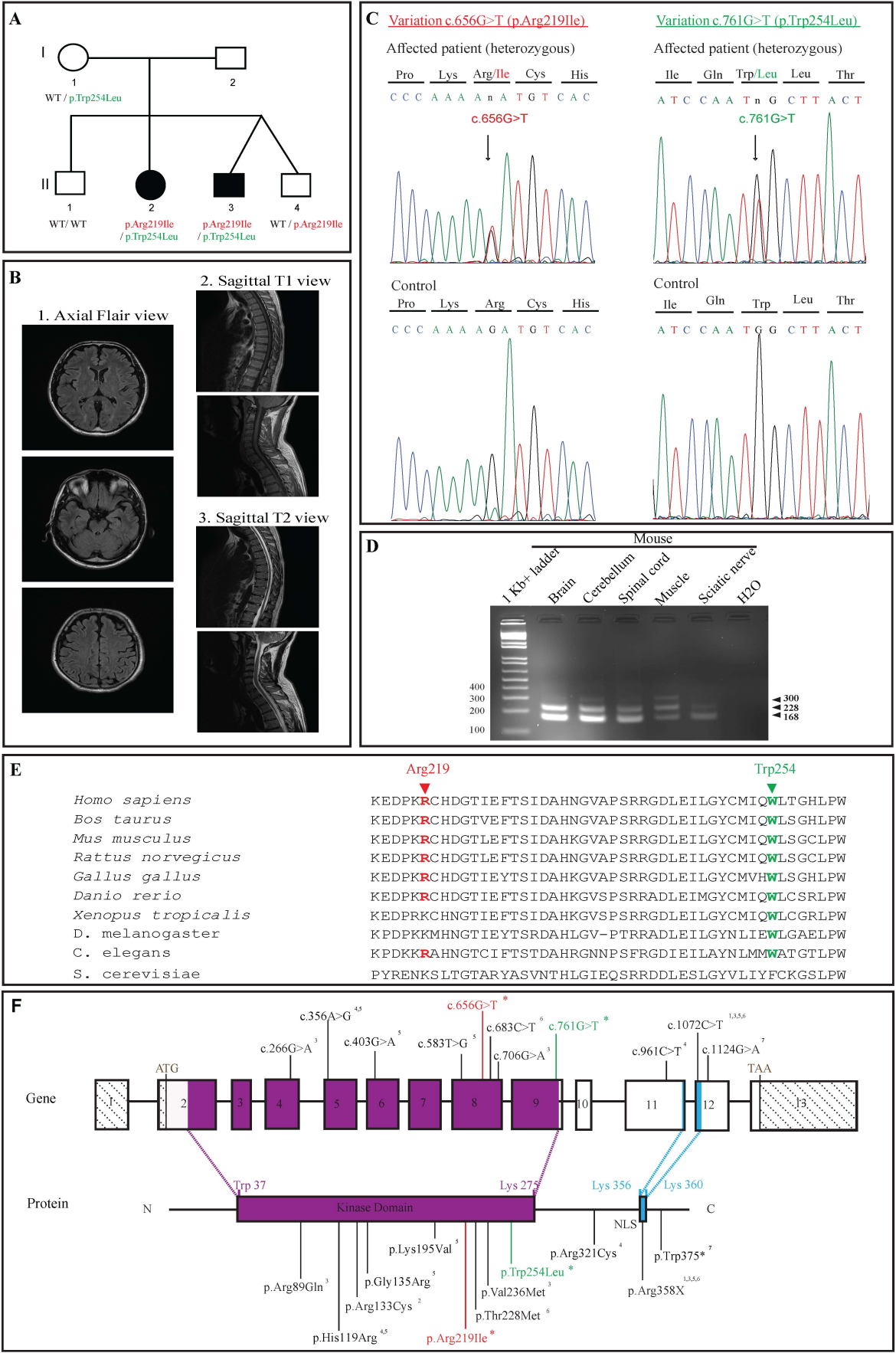
Figure 4. Depletion of VRK1 in patients' fibroblasts lead to altered coilin expression

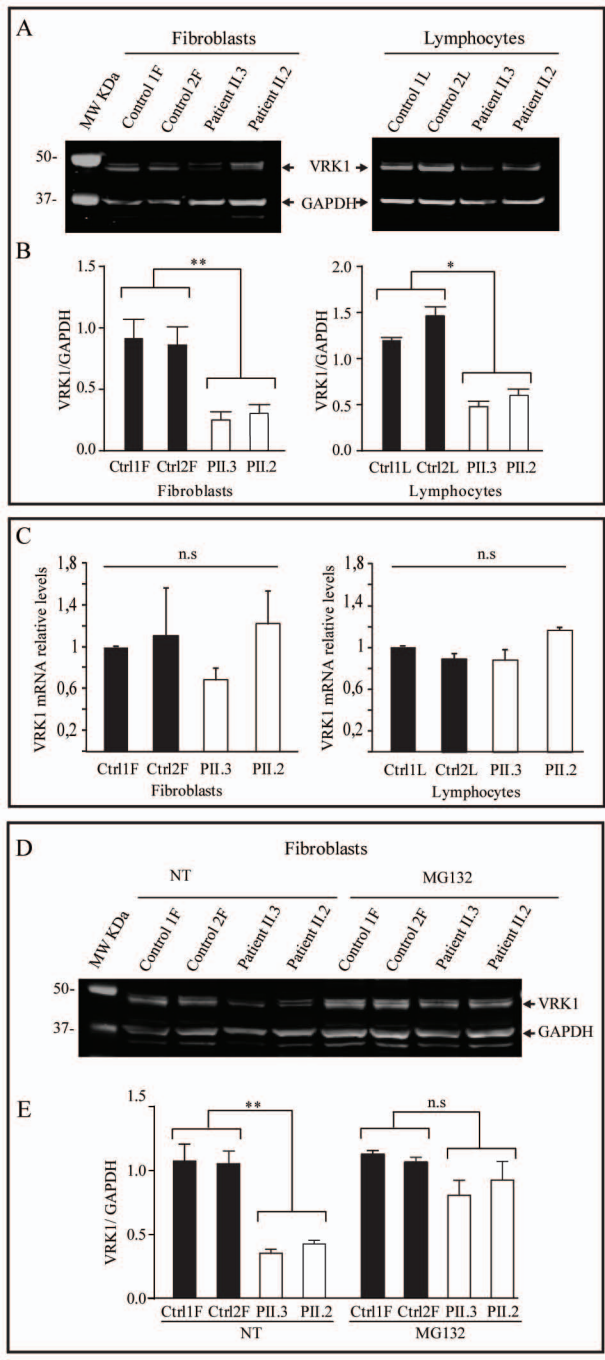
(A) Immunolabeling of coilin in fibroblasts from patients and controls shows a drastic decrease of coilin levels in the nucleus of patients II.2 and II.3 as compared to controls 1F and 2F. Here, we have presented the results for one patient (II.2) and one control (1F). Normal expression levels are restored after MG132 treatment. **(B)** Relative coilin fluorescence intensity was measured in 300 cells for each sample (scored from three independent experiments); the median for each sample is shown on the graph. Statistical significance was evaluated by Mann-Whitney test, ****$p<0.0001$, n.s: not significant. NT: Not treated.

Figure 5. VRK1 is essential for Cajal Bodies (CBs) assembly and for neurite elongation and branching in motor neurons.

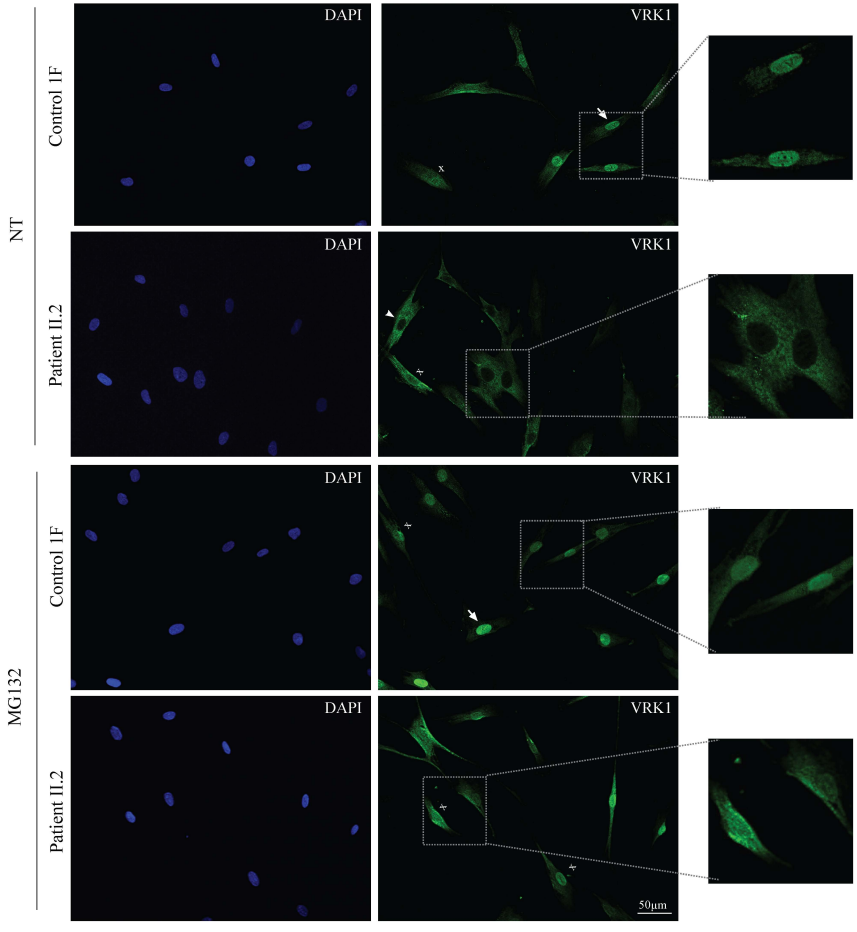
(A) Immunolabeling of VRK1 and coilin shows reduced VRK1 expression and CBs disintegration in hiPSC-derived Motor Neurons (MNs) from patient II.2 as compared to control hiPSC-derived MNs (1M). NFM: Neurofilament Medium Chain. **(B-C)** This reduction was confirmed and quantified by western blot experiments and statistically significant decrease of 64% was found in patient's MNs. Quantification was performed on data from three independent experiments. Statistical significance was evaluated by unpaired t-test, *$p<0.05$. **(D)** Percentage of MNs containing CBs was evaluated in at least 800 cells counted from three independent experiments (828 cells for control 1M and 1269 cells for patient II.2). Statistical significance was evaluated by unpaired t-test, ***$p<0.0005$. **(E)** In CBs positive cells (CBs^+), a tendency to decreased CBs size was observed in hiPSC-derived MNs from patient when compared to control. CBs size was calculated, for each individual, by measuring CBs area on 50 CBs^+ hiPSC-derived MNs, using ImageJ software. Data were recorded from three independent experiments. **(F)** Images of AAV 2/6-CMV-GFP infected hiPSC-derived MNs from patient II.2 and control 1M. **(G-I)** Graphs presenting neurite length and branching studies in AAV 2/6-CMV-GFP infected hiPSC-derived MNs in patient II.2 versus control (median values from two independent experiments are presented). Neurite length was calculated by summing up the length of all neuritic branches per cell (G). For branching studies, we evaluated two parameters: number of branches (H) and number of branching points (I) per cell. All 3 parameters are decreased in patient II.2's hiPSC-derived MNs: i) neurite length is decreased by 36%, ii) number of branches per cell drops from 5 (in control) to 3 (in patient), and iii) number of branching points decreases from 2 (in control) to 1 (in patient) (median values from two independent experiments). 50 MNs cells were counted

for each individual. Statistical significance was evaluated by Mann-Whitney test, *$p<0.1$, **$p<0.01$, ***$p<0.001$.


A
I
1
WT / p.Trp254Leu
2
II
1
WT/ WT
2
p.Arg219Ile / p.Trp254Leu
3
p.Arg219Ile / p.Trp254Leu
4
WT / p.Arg219Ile
B
1. Axial Flair view
2. Sagittal T1 view
3. Sagittal T2 view
C
Variation c.656G>T (p.Arg219Ile)
Affected patient (heterozygous)
Pro Lys Arg/Ile Cys His
C C C A A A A n A T G T C A C
c.656G>T
Control
Pro Lys Arg Cys His
C C C A A A A G A T G T C A C
Variation c.761G>T (p.Trp254Leu)
Affected patient (heterozygous)
Ile Gln Trp/Leu Leu Thr
A T C C A A T n G C T T A C T
c.761G>T
Control
Ile Gln Trp Leu Thr
A T C C A A T G G C T T A C T
D
Mouse
1 Kb+ ladder
Brain
Cerebellum
Spinal cord
Muscle
Sciatic nerve
H2O
400
300
200
100
300
228
168
E
Arg219
Trp254
Homo sapiens KEDPKRCHDGTIEFTSIDAHNGVAPSRRGDLEILGYCMIQWLTGHLPW
Bos taurus KEDPKRCHDGTVEFTSIDAHNGVAPSRRGDLEILGYCMIQWLSGHLPW
Mus musculus KEDPKRCHDGTLEFTSIDAHKGVAPSRRGDLEILGYCMIQWLSGCLPW
Rattus norvegicus KEDPKRCHDGTLEFTSIDAHNGVAPSRRGDLEILGYCMIQWLSGCLPW
Gallus gallus KEDPKRCHDGTIEYTSIDAHKGVAPSRRGDLEILGYCMVHWLSGHLPW
Danio rerio KEDPKRCHDGTIEFTSIDAHKGVSPSRRADLEIMGYCMIQWLCSRLPW
Xenopus tropicalis KEDPRKCHNGTIEFTSIDAHKGVSPSRRGDLEILGYCMIQWLCGRLPW
D. melanogaster KPDPKKMHNGTIEYTSRDAHLGV-PTRRADLEILGYNLIEWLGAELPW
C. elegans KPDKKRAHNGTCIFTSTDAHRGNNPSFRGDIEILAYNLMMWATGTLPW
S. cerevisiae PYRENKSLTGTARYASVNTHLGIEQSRRDDLESLGYVLIYFCKGSLPW
F
Gene
ATG
TAA
1 2 3 4 5 6 7 8 9 10 11 12 13
c.356A>G 4,5
c.266G>A 3
c.403G>A 5
c.583T>G 5
c.656G>T *
c.683C>T 6
c.706G>A 3
c.761G>T *
c.961C>T 4
c.1072C>T 1,3,5,6
c.1124G>A 7
Protein
N
C
Kinase Domain
NLS
Trp 37
Lys 273
Lys 356
Lys 360
p.His119Arg 4,5
p.Arg89Gln 3
p.Arg133Cys 2
p.Gly135Arg 5
p.Lys195Val 5
p.Thr228Met 6
p.Arg219Ile *
p.Val236Met 3
p.Trp254Leu *
p.Arg321Cys 4
p.Arg358X 1,3,5,6
p.Trp375* 7

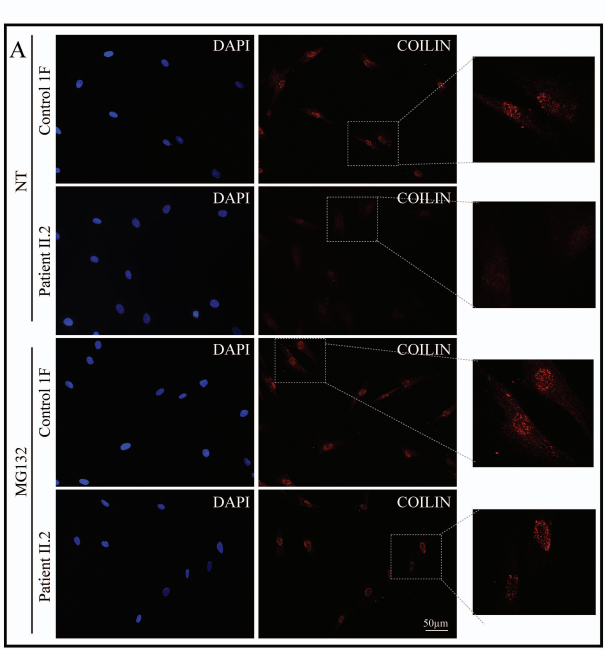

A
Fibroblasts
Lymphocytes
MW KDa
Control 1F
Control 2F
Patient II.3
Patient II.2
Control 1L
Control 2L
Patient II.3
Patient II.2
50-
37-
VRK1
GAPDH
B
VRK1/GAPDH
**
*
Ctrl1F Ctrl2F PII.3 PII.2
Fibroblasts
Ctrl1L Ctrl2L PII.3 PII.2
Lymphocytes
C
VRK1 mRNA relative levels
n.s
Ctrl1F Ctrl2F PII.3 PII.2
Fibroblasts
Ctrl1L Ctrl2L PII.3 PII.2
Lymphocytes
D
Fibroblasts
NT
MG132
MW KDa
Control 1F
Control 2F
Patient II.3
Patient II.2
50-
37-
VRK1
GAPDH
E
VRK1/ GAPDH
**
n.s
Ctrl1F Ctrl2F PII.3 PII.2
NT
MG132

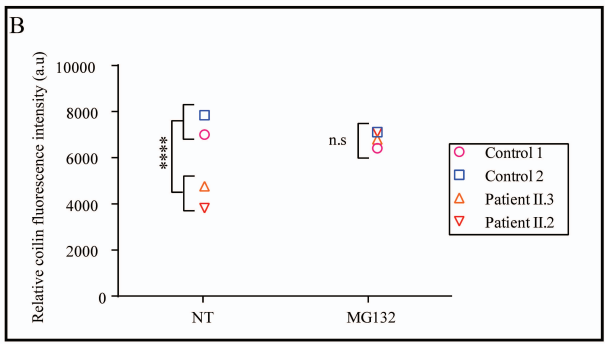
A



B

Fibroblasts / VRK1 localization	Control 1		Control 2		Patient II.3		Patient II.2	
	NT	MG132	NT	MG132	NT	MG132	NT	MG132
Strictly Nuclear (R>2)	60%	53%	39%	50%	-	17%	-	17%
Nuclear and Cytoplasmic (0,9<R<2)	40%	47%	61%	50%	20%	80%	13%	81%
Strictly Cytoplasmic (0,1<R<0,9)	-	-	-	-	72%	3%	73%	2%
Absent	-	-	-	-	8%	-	14%	-


A
NT
MG132
Control 1F
Patient II.2
Control 1F
Patient II.2
DAPI
COILIN
50µm


B
Relative coilin fluorescence intensity (a.u)
10000
8000
6000
4000
2000
0
****
n.s
NT
MG132
Control 1
Control 2
Patient II.3
Patient II.2

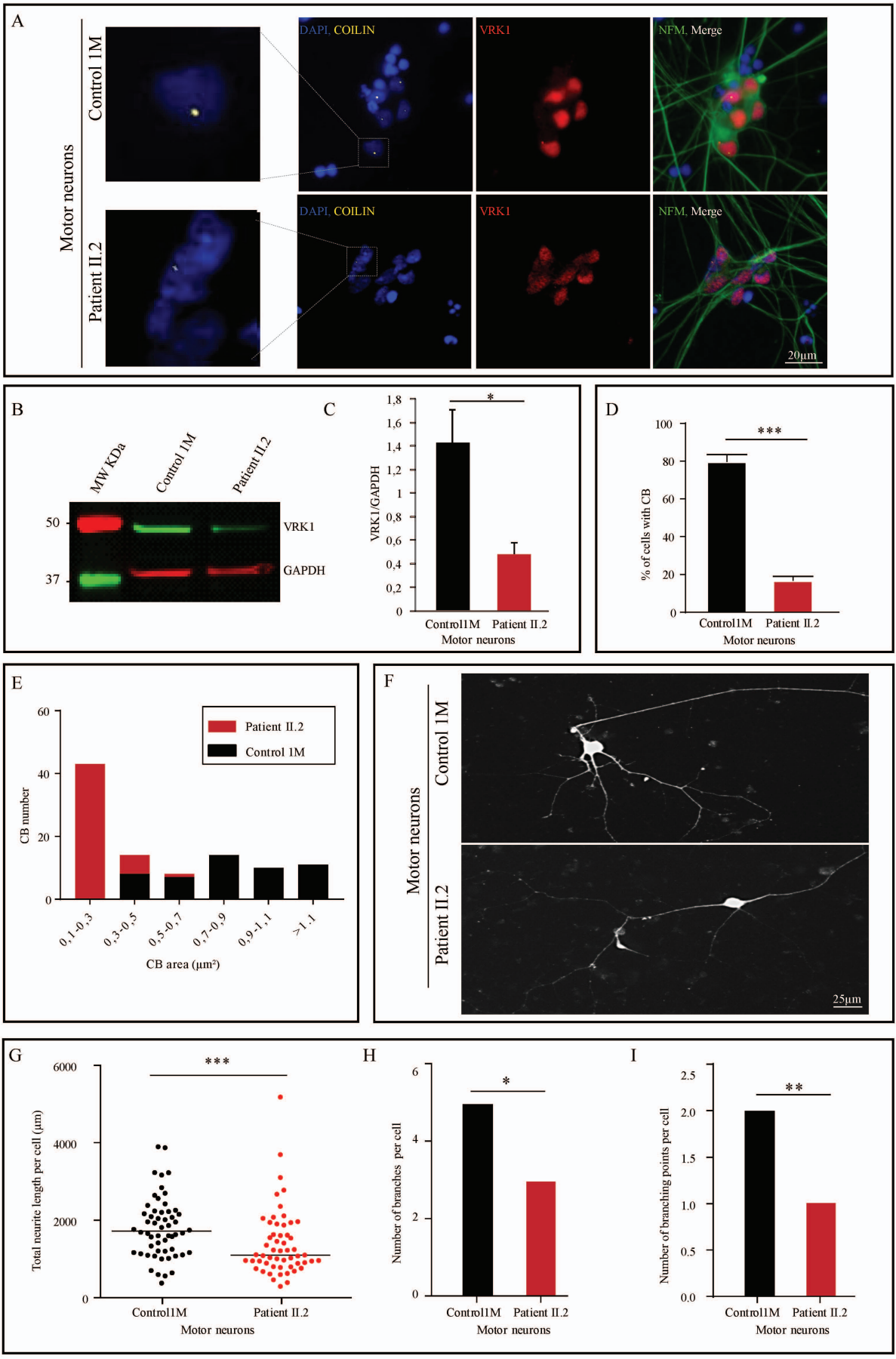

A
Motor neurons
Control 1M
Patient II.2
DAPI, COILIN
VRK1
NFM, Merge
20µm
B
MW KDa
Control 1M
Patient II.2
50
37
VRK1
GAPDH
C
VRK1/GAPDH
Control1M
Patient II.2
Motor neurons
D
% of cells with CB
Control1M
Patient II.2
Motor neurons
E
Patient II.2
Control 1M
CB number
CB area (µm²)
0,1-0,3
0,3-0,5
0,5-0,7
0,7-0,9
0,9-1,1
>1.1
F
Motor neurons
Control 1M
Patient II.2
25µm
G
Total neurite length per cell (µm)
Control1M
Patient II.2
Motor neurons
H
Number of branches per cell
Control1M
Patient II.2
Motor neurons
I
Number of branching points per cell
Control1M
Patient II.2
Motor neurons

Table 1. Nerve conduction studies in Patient II.3 at age 16 years

Motor Conduction studies				
Nerve		*DL*	*CMAP*	*CV*
Peroneal L	Ankle	6,2	3,2	
	Below fibula	15.5	2	40
	Above Fibula	17.3	1.6	53
Peroneal R	Ankle	6.5	4.6	
	Below fibula	16.7	3.5	39
	Above Fibula	18	3.2	46
posterior tibial L	Ankle	6.3	1.07	
	Knee	16.6	0.1	38
posterior tibial L	Ankle	5.5	2.2	
	Knee	17.3	1.3	36
Median R	Wrist	4	8.7	
	Elbow	8.6	8.5	57
Ulnar R	Wrist	2.9	7.8	
	Below elbow	7.5	7.3	50
	Above Elbow	8.5	7.4	43
Sensory Conduction studies				
Nerve		*DL*	*SNAP*	*CV*
Median R		3,1	18,9	50
Sural L		4.6	16.8	40
Sural R		2.6	18.7	42

DL=Distal Latencies. CMAP=Compound Motor Action Potential.

SNAP=Sensory Nerve Action Potential. CV=Conduction Velocities.

Normal values: DL< 4 ms, CMAP>4, CV>40 m/s

Table 2. Comparison of the genetic, clinical and electrophysiology findings of all patients described to date with VRK1 mutations

Reference	This study	This study	Renbaum *et al.* , 2009[7]	Renbaum *et al.* , 2009[7]	Najmabadi *et al.* , 2011[53]	Gonzaga-Jauregui *et al.* , 2013[12]	Gonzaga-Jauregui *et al.* , 2013[12]	Gonzaga-Jauregui *et al.* , 2013[12]	Nguyen *et al.* , 2015[8]	Stoll *et al.* , 2016[9]	Stoll *et al.* , 2016[9]	Stoll *et al.* , 2016[9]	Tomaselli *et al.*, 2016[10]	Li *et al.*, 2016[11]	Li *et al.*, 2016[11]
Family	**Family 10**	**Family 10**	**Family 1**	**Family 1**	**Family 2**	**Family 3**	**Family 3**	**Family 4**	**Family 5**	**Family 6**	**Family 6**	**Family 7**	**Family 8**	**Family 9**	**Family 9**
Patients	Patient II.2	Patient II.3	Proband IV.12	Patient IV.6	M017N	BAB3022	BAB3280	BAB5311		Family 1. Patient II.4	Family 1. Patient II.1	Family 2. Patient III.3		Patient II.7	Patient II.9
Clinical Diagnosis	dHMN+upper motor neuron signs	dHMN+upper motor neuron signs	SMA with PCH	SMA with PCH	PCH	Complex Motor and Sensory Axonal Neuropathy + microcephaly	Complex Motor and Sensory Axonal Neuropathy + microcephaly	Complex Motor and Sensory Axonal Neuropathy + microcephaly	Possible ALS	Adult onset distal SMA	Adult onset distal SMA. Similar to patient II.4, but later onset, more progressive	Childhood motor neuron disease, juvenile ALS	dHMN+upper motor neuron signs	adult-onset dSMA	adult-onset dSMA
Mutations Allele 1	c.656G>T	c.656G>T	c.1072C>T	n.d. (deceased)		c.266G>A	c.266G>A	c.1072C>T	c.961C>T	c.356A>G	c.356A>G	c.403G>A	c.683C>T	c.1124G >A	c.1124G >A
	p.Arg219Ile	p.Arg219Ile	p.Arg358*		p.Arg133Cys	p.Arg89Gln	p.Arg89Gln	p.Arg358*	p.Arg321Cys	p.His119Arg	p.His119Arg	p.Gly135Arg	p.Thr228Met	p.W375*	p.W375*
Allele 2	c.761G>T	c.761G>T	c.1072C>T			c.706G>A	c.706G>A	c.1072C>T	c.356A>G	c.1072C>T	c.1072C>T	c.583T>G	c.1072C>T	c.1124G >A	c.1124G >A
	p.Trp254Leu	p.Trp254Leu	p.Arg358*	n.d. (deceased)	p.Arg133Cys	p.Val236Met	p.Val236Met	p.Arg358*	p.His119Arg	p.Arg358*	p.Arg358*	p.Lys195Val	p.Arg358*	p.W375*	p.W375*
Genotype	Compound Heterozygous	Compound Heterozygous	Homozygous	n.d. (deceased)	Homozygous	Compound Heterozygous	Compound Heterozygous	Homozygous	Compound Heterozygous	Compound Heterozygous	Compound Heterozygous	Compound Heterozygous	Compound Heterozygous	Homozygous	Homozygous
Type of mutation	Missense	Missense	Nonsense	Nonsense	Missense	Missense	Missense	Nonsense	Missense	Missense	Missense	Missense	Missense	Nonsense	Nonsense
Inheritance pattern	AR	AR	AR		AR	AR	AR	AR	Sporadic	AR	AR	AR	AR	AR	AR
Consanguinity	-	-	+	+	+	-	-	-	-	-	-	-	-		
Origin	Lebanese	Lebanese	Ashkenazi Jewish	Ashkenazi Jewish	Iranian	n.a	n.a	Ashkenazi Jewish	Hispanic	Ashkenazi Jewish	Ashkenazi Jewish	n.a	n.a.	Chinese	Chinese
sex/age, Y	F/47	M/38	F/died at 11,5 Y	F/died at 9,5 Y (sisters)	n.a.	F/10	F/2	M/9	M/32 Y	F/35 Y	M/older	F/3Y	F/34Y	M/47	F/
Onset	Childhood (~10 years)	Childhood (~10 years)	Early	Early	n.a.	Early	Early	Early	Late	teenage (15 years)	late teenage	Late	Childhood (~10 years)	adulthood	adulthood
Initial symptoms	Motor	Motor	Motor +early-onset ataxia	Motor +early-onset ataxia	n.a.	Motor	Motor	Motor	Motor	Motor	Motor	Motor	Motor	Motor. Difficulties in walking on tiptoe	Wasting and weakness of distal symmetrical lower limbs
Progression	Slow	Slow	Rapid	Rapid	n.a.	Rapid	Rapid	Rapid	Slow	Moderate	Moderate	Rapid	Slow	Slow	Slow
Functional impairment	+. Minor	+. Minor. Sometimes use a walking stick.	++. Unable to sit at 1yr or walk without support at age 2 y. Wheelchaired.	++	n.a.	++. Never able to walk alone.Wheelchaired at 4y, limited use of arms and hands at 6Y	++. Unable to sit or walk without support at age 20months	++. Wheelchaired from 6Y	+. Minor. Falls, walking difficulties	+. Long arm crutch and mobility scooter	+.Wheelchair at 40Y	++. Loss of ambulation at 10Y. Complete dependance for daily activities at 18Y	+. Minor	++. Requires wheelchair to mobilize at age 42	+. High steppage gait due to bilateral foot drop at age 40 years
Muscle atrophy	+. Distally in hands and feet	+. Distally in hands and feet	+. Distal muscle atrophy, Progressive muscle wasting	+. Distal muscle atrophy, Progressive muscle wasting	n.a.	+. Distal progressive muscle wasting progressing to proximal	-	+	+. Profound proximal and distal muscle atrophy with fat replacement (seen at muscle MRI)	+. Profound symmetric distal muscle wasting at upper and lower limbs. Weakness more severe at lower limbs.	n.a.	+. Distal amyotrophy of the lower limbs	+. Distal amyotrophy of the lower limbs	++. severe distal muscle wasting in legs and arms	+. Severe and moderate wasting and weakness in lower extremities. Slight wasting and weakness of the ulnar intrinsic muscle of both hands
Sensory disturbance	Normal in all modalities	Normal in all modalities	+/-	+/-	n.a.	-	-	-	+ (pain)	-	n.a.	-	-	-	-
DTR	Brisk	Brisk	Brisk	Brisk	n.a.	Loss	normal	Decreased	Brisk	Brisk in the upper limbs Absent in the lower limbs	n.a.	Brisk	Brisk	Absent in lower and upper limbs	Absent in lower limbs. Brisk in upper limbs.
Other Upper motor neuron signs	+. Bilateral Babinski signs	+. Bilateral Babinski signs	-	-	-	-	-	-	-	+	n.a.	+	n.a.	-	-
Intellectual disability	-	-	Mild	+	n.a.	-	-	-	-	-		-	-	-	-
Ataxia	-	-	++	++	n.a.	-	-	-	-	-	n.a.		-	-	-
Hypotonia	-	-	+	+	n.a.	+	+ (mild)	+ (worsening)	-	-	n.a.		+	+(all limbs)	+(all limbs)
Speech problems	-	-	++ (dysarthria)	++ (dysarthria)	n.a.	-	-	+ (hypophonia and dysarthria)	-	-	n.a.	+ (hoarse and soft)	-	-	-
Skeletal deformations	-	-	+. Bilateral equinovarus	+. Bilateral equinovarus	n.a.	+. Scoliosis (worsening)	-	+. Scoliosis (worsening)	+. pes cavus and hammer toes	-	n.a.	+. severe scoliosis	n.a.	n.a.	+.Pes arcuatus
Feeding problems	-	-	++ (gastrostomy)	++ (gastrostomy)	n.a.	Gastric feeding tube	-	Gastric feeding tube	-	-	n.a.	-	-	-	-
Respiratory problems	-	-	-	-	n.a.	+	-	-	-	+(mild)	n.a.	+ (dyspnea and weak cough)	-	-	-
Microcephaly	-	-	+ (progressive)	+ (progressive)	n.a.	+	+(*in utero* , nonprogressive)	+ (*in utero* , severe)	-	-	n.a.	-	-	-	-
MRI (Brain and spinal cord)	Normal	Normal	Cerebellar hypoplasia. Normal spine	Cerebellar hypoplasia. Mild cerebral atrophy	n.a.	Microcephaly with simplified gyral pattern	Microcephaly with simplified gyral pattern	Microcephaly with simplified gyral pattern and underdevelopped cerebellar vermis	Normal	Mild to moderate generalized atrophy+atrophy of the spinal cord. **No PCH. No CA**	n.a.	Normal	Normal	Normal	Normal
NCS measurements motor nerves*	Decreased Amplitudes. Slightly reduced MNCVs: Upper and lower	Decreased Amplitudes. Slightly reduced MNCVs: Upper and lower	Axonal	Demyelinating?	n.a.	amplitudes and MNCVS n.r. (upper and lower)	Upper: normal to slighlty decreased MNCVs. Decreased amplitudes Lower: Normal MNCVs. Decreased amplitudes	Upper: normal MNCVs. Decreased amplitudes Lower: Normal MNCVs. Decreased amplitudes	Upper and lower: normal MNCVS and amplitudes	Severely decreased amplitudes. Relatively preserved MNCVs.	n.a.	Severely decreased amplitudes.	Motor	Upper and lower limbs: MNCVs not recordable, severe reduction in CMAP amplitudes.	Normal MNCVs at upper and lower limbs. Normal CMAPs at upper limbs. Reduction of CMAPs at lower limbs, which remain normal except at the peroneal nerve
NCS measurements sensory nerves*	Normal	Normal	Axonal	Demyelinating?	n.a.	Upper: normal NCVs. Slighltly decreased amplitudes Lower: n.r. NCVs. amplitudes n.a.	Upper: slighlty decreased NCVs. Normal amplitudes Lower: n.r.	Upper: normal MNCVs. Decreased amplitudes Lower: n.r.	Upper and lower: normal NCVS and amplitudes	Upper and lower: normal MNCVS and amplitudes	n.a.	Initially normal. From 19Y: decreased amplitudes	n.a.	Normal sensory NCVs and amplitudes.	Normal Sensory NCVs and amplitudes at upper and lower limbs.
EMG	Mild neurogenic motor units action potentials with preserved recruitment . Denervation of calf muscles	Mild neurogenic motor units action potentials with preserved recruitment	Fasciculations and fibrillations	Fasciculations and fibrillations	n.a.	n.a.	n.a.	n.a.	Fasciculations and fibrillations +chronic denervation/re-innervation in lower muscles	Chronic and active denervation	Chronic and active denervation	Chronic and active denervation	n.a.	Chronic neurogenic changes of high amplitude, long duration motor units	Chronic neurogenic changes
Nerve Biospy	n.a.	n.a.	Normal (sural)	n.a.	n.a.	Axonal neuropathy	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Muscle biospy	n.a.	n.a.	Neurogenic atrophy	Neurogenic atrophy	n.a.	Neurogenic myopathy	n.a.	n.a.	Neurogenic atrophy without evidence of reinnervation	n.a.	n.a.	Neurogenic atrophy	n.a.	n.a.	Neurogenic atrophy (right quadriceps)

Motor stands for distal lower limbs weakness/atrophy or gait disturbance. n.a.=not available. n.r.=not recordable, n.d.=not determined dHMN=distal Hereditary Motor Neuropathy, AR=Autosomal Recessive, F=Female, M=Male, ,DTR=Deep tendon Reflexes , CMAP=Compound Muscle Action Potential
SMA=Spinal Muscular Atrophy, PCH=PontoCerebellar Hypoplasia
Genbank accession number for VRK1: NM_003384

Supplementary Table 1: exome sequencing statistics and filtering

Exome sequencing coverage statistics

	Patient II.2	**Patient II.3**
Mean coverage	43X	45X
% of capture targets covered at 20X	74	75

Details of filtering steps in variant analysis for compound heterozygous variants

Filter	**Patient II.2**	**Patient II.3**	**Shared heterozygous**
Total variants	20911	21438	7526
Total variants in CDS	17837	18134	6509
Apparently compound heterozygous	7866	7978	4265
With global MAF≤1% in gnomAD	205	235	134
Additional filtering (GME, local database)	143	157	94
Non synonymous, nonsense, frameshift, insertion/deletion, splicing variants	94	96	57

Details of filtering steps in variant analysis for homozygous variants

Filter	**Patient II.2**	**Patient II.3**	**Shared homozygous**
Total variants	7801	8251	5954
Total variants in CDS	6660	6984	5137
With global MAF≤1% in gnomAD	5	20	3
Additional filtering (GME, local database)	0	16	0
Non synonymous, nonsense, frameshift, insertion/deletion, splicing variants	0	15	0

Supplementary Table 2. Final list of variants

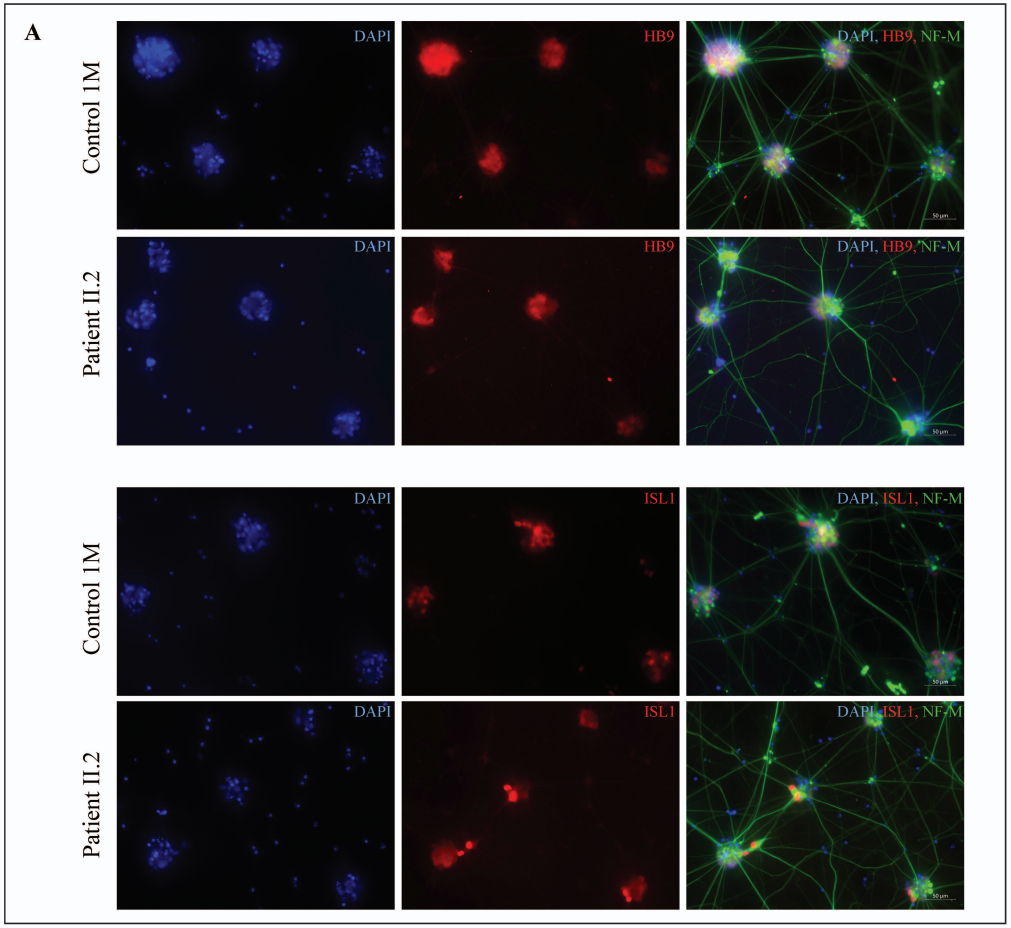
Chr	Position hg19	Ref	Alt	Genotype	Depth	Frequency	SNV Score	Gene	Type of variation	Variation	dbSNP147
17	74003216	C	T	het	16	0.44	154.77	EVPL	nonsynonymous SNV	NM_001988:exon22:c.6070G>A:p.A2024T	rs1495123
17	74004307	C	T	het	20	0.55	321.77	EVPL	nonsynonymous SNV	NM_001988:exon22:c.4979G>A:p.R1660Q	rs1996863
17	74004928	G	A	het	111	0.56	1956.77	EVPL	nonsynonymous SNV	NM_001988:exon22:c.4358C>T:p.P1453L	rs1131359
17	74007865	G	C	het	21	0.48	286.77	EVPL	nonsynonymous SNV	NM_001988:exon20:c.2556C>G:p.S852R	rs1411343
18	7037641	G	A	het	74	0.51	1113.77	LAMA1	nonsynonymous SNV	NM_005559:exon12:c.1673C>T:p.A558V	rs1439978
18	6997817	G	A	het	102	0.5	1461.77	LAMA1	nonsynonymous SNV	NM_005559:exon33:c.4730C>T:p.S1577F	rs7725448
2	1.79E+08	A	G	het	152	0.51	2569.77	TTN	nonsynonymous SNV	NM_001267550:exon322:c.68525T>C:p.I22842T	rs3683015
2	1.79E+08	C	T	het	144	0.52	2280.77	TTN	nonsynonymous SNV	NM_001267550:exon357:c.100226G>A:p.C33409Y	rs2011120
14	97321640	G	T	het	52	0.54	831.77	VRK1	nonsynonymous SNV	NM_003384:exon8:c.656G>T:p.R219I	.
14	97322518	G	T	het	126	0.5	1955.77	VRK1	nonsynonymous SNV	NM_003384:exon9:c.761G>T:p.W254L	.

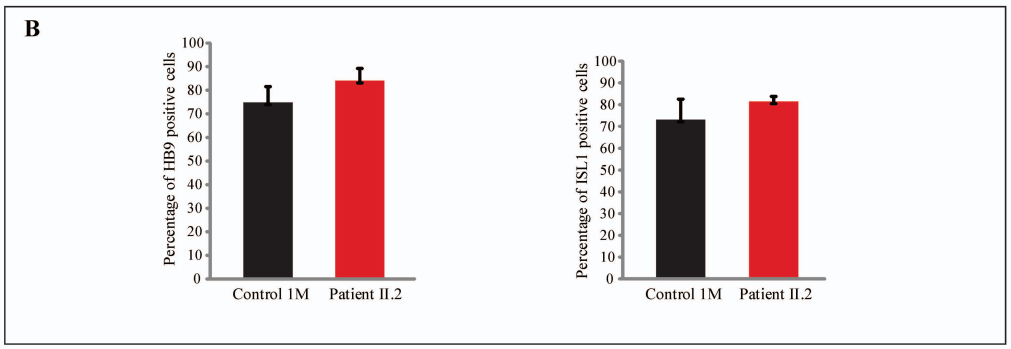
Chr = Chromosome; Ref = Reference allele; Alt = alternate allele; Depth = number of reads at position; gnomAD_exome MAF = minor allele frequency in gnomAD exome dataset; gnomAD_genome MAF = minor allele frequency in gnomAD genome dataset; GME MAF = minor allele frequency in the GME dataset.

Supplementary Legends

Supplementary Figure 1. **Analysis of hiPSC to spinal Motor Neuron differentiation efficiency by immunostaining of two transcription factors of spinal motor neurons: HB9 and ISLET1.**

(A) Representative images of immunostainings for ISL1 and HB9 and the pan neuronal marker neurofilament (NF) in hiPSC-derived MNs at day 30 of the differentiation protocol for patient II.2 and one control. NFM= Neurofilament Medium Chain. **(B)** Quantitative analysis of the differentiation efficiency on day 30 for patient II.2 and control 1M for the three experiments presented in Figure 5. For each experiment, these results have been quantified as the percentage of HB9 or ISL1 positive cells among DAPI positive cells. For each experiment, we scored at least 300 cells for each individual.


A
Control 1M
Patient II.2
DAPI
HB9
DAPI, HB9, NF-M
ISL1
DAPI, ISL1, NF-M
50 µm


B
Percentage of HB9 positive cells
Percentage of ISL1 positive cells
0
10
20
30
40
50
60
70
80
90
100
Control 1M
Patient II.2

Supplementary Table 1: exome sequencing statistics and filtering

Exome sequencing coverage statistics

Individual	Patient II.2	Patient II.3
Mean coverage	43X	45X
% of capture target covered at 20X	74	75

Details of filtering steps in variant analysis for compound heterozygous variants

Filter	Patient II.2	Patient II.3	Shared heterozygous
Total variants	20911	21438	7526
Total variants in CDS	17837	18134	6509
Apparently compound heterozygous	7866	7978	4265
With global MAF≤1% in gnomAD	205	235	134
Additional filtering (GME, local database)	143	157	94
Non synonymous, nonsense, frameshift, insertion/deletion, splicing variants	94	96	57

Details of filtering steps in variant analysis for homozygous variants

Filter	Patient II.2	Patient II.3	Shared homozygous
Total variants	7801	8251	5954
Total variants in CDS	6660	6984	5137
With global MAF≤1% in gnomAD	5	20	3
Additional filtering (GME, local database)	0	16	0
Non synonymous, nonsense, frameshift, insertion/deletion, splicing variants	0	15	0

Supplementary Table 2. Final list of apparently compound heterozygous variants

Chr	Position hg19	Ref	Alt	Genotype	Depth	Frequency	SNV Score	Gene
17	74003216	C	T	het	16	0.44	154.77	EVPL
17	74004307	C	T	het	20	0.55	321.77	EVPL
17	74004928	G	A	het	111	0.56	1956.77	EVPL
17	74007865	G	C	het	21	0.48	286.77	EVPL
18	7037641	G	A	het	74	0.51	1113.77	LAMA1
18	6997817	G	A	het	102	0.5	1461.77	LAMA1
2	179442717	A	G	het	152	0.51	2569.77	TTN
2	179401248	C	T	het	144	0.52	2280.77	TTN
14	97321640	G	T	het	52	0.54	831.77	VRK1
14	97322518	G	T	het	126	0.5	1955.77	VRK1

Chr = Chromosome; Ref = Reference allele; Alt = alternate allele; Depth = number of reads at positic

Type of variation	Variation	dbSNP147
nonsynonymous SNV	NM_001988:exon22:c.6070G>A:p.A2024T	rs149512337
nonsynonymous SNV	NM_001988:exon22:c.4979G>A:p.R1660Q	rs199686324
nonsynonymous SNV	NM_001988:exon22:c.4358C>T:p.P1453L	rs113135904
nonsynonymous SNV	NM_001988:exon20:c.2556C>G:p.S852R	rs141134319
nonsynonymous SNV	NM_005559:exon12:c.1673C>T:p.A558V	rs143997842
nonsynonymous SNV	NM_005559:exon33:c.4730C>T:p.S1577F	rs772544804
nonsynonymous SNV	NM_001267550:exon322:c.68525T>C:p.I22842T	rs368301580
nonsynonymous SNV	NM_001267550:exon357:c.100226G>A:p.C33409Y	rs201112096
nonsynonymous SNV	NM_003384:exon8:c.656G>T:p.R219I	.
nonsynonymous SNV	NM_003384:exon9:c.761G>T:p.W254L	.

ɔn; gnomAD_exome MAF = minor allele frequency in gnomAD exome dataset; gnomAD_genome MAF = m

Frequency in databases			Pathogenicity predic				
gnomAD_exome MAF	gnomAD_genome MAF	GME MAF	CADD	Mutation Taster	Mutation Assessor	Polyphen2	SIFT
0.0006	0.0011	0.004032	18,16	N	M	B	T
0.0005	0.0005	.	25,5	D	M	B	T
0.0006	0.0010	0.005539	22,5	D	M	B	T
0.0006	0.0010	0.004532	23,8	D	M	B	T
0.0037	0.0031	.	5,576	N	N	B	T
1.218e-05	.	.	15,54	N	M	B	T
1.643e-05	.	.	14,17	D	L	D	D
9.4e-05	0.0004	0.001514	21,8	D	L	P	D
.	.	.	32	D	L	D	D
.	.	.	29,8	D	L	D	D

ıinor allele frequency in gnomAD genome dataset; GME MAF = minor allele frequency in the GME dataset

tions	
UMD Predictor	
UMD score	**UMD Prediction**
66	Probably pathogenic
60	Probable polymorphism
60	Probable polymorphism
78	Pathogenic
5	Polymorphism
69	Probably pathogenic
93	Pathogenic
90	Pathogenic
99	Pathogenic
90	Pathogenic