

Medizinische Fakultät
der
Universität Duisburg-Essen

Aus dem Institut für Humangenetik

Characterization of intronic pentanucleotide expansions in neurological disorders

I n a u g u r a l - D i s s e r t a t i o n
zur
Erlangung des Doktorgrades der Medizin
durch die Medizinische Fakultät
der Universität Duisburg-Essen

Vorgelegt von
Rahel Theres Bettges
geb. Florian
aus Düsseldorf
2021

Dekan: Herr Univ.-Prof. Dr. med. J. Buer

1. Gutachter/in: Frau Univ.-Prof. Dr. rer. nat. Ch. Depienne

2. Gutachter/in: Herr Prof. Dr. G. Lesca

3. Gutachter/in: Frau Univ.-Prof. Dr. med. D. Timmann-Braun

Tag der mündlichen Prüfung: 15 Juli 2022

Parts of this work have been published in:

Florian et al. (2019) Unstable TTTTA/TTTCA expansions in *MARCH6* are associated with Familial Adult Myoclonic Epilepsy type 3

and:

Corbett et al. (2019) Intronic ATTTC repeat expansions in *STARD7* in familial adult myoclonic epilepsy linked to chromosome 2

DuEPublico

Duisburg-Essen Publications online



UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken



ub | universitäts
bibliothek

Diese Dissertation wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

DOI: 10.17185/duepublico/78827
URN: urn:nbn:de:hbz:465-20230922-072648-5

Alle Rechte vorbehalten.

Table of Contents

1	Introduction	7
1.1	Familial Adult Myoclonic Epilepsy (FAME)	7
1.1.1	Clinical features.....	8
1.1.2	Genetics.....	12
1.1.3	Neuropathological findings	15
1.2	Essential tremor (ET)	16
1.3	Aim.....	19
1.3.1	FAME.....	19
1.3.2	Screening for expansions in <i>MARCF6</i> and <i>STARD7</i> in ET-patients....	20
2	Materials und Methods.....	21
2.1	Materials.....	21
2.1.1	General Equipment.....	21
2.1.2	Chemicals, Buffer and Enzymes	22
2.1.3	Software	25
2.1.4	Primers	25
2.1.5	Patients and Probes.....	26
2.2	Methods.....	29
2.2.1	Isolation of genomic DNA from Blood.....	29
2.2.2	Concentration Measurement of Nucleic Acids	29
2.2.3	Polymerase Chain Reaction (PCR)	29
2.2.4	Agarose gel electrophoresis	34
2.2.5	Fragment length analysis (GeneScan).....	34
2.2.6	Sanger Sequencing	35
2.2.7	Mini Prep/TA cloning	38
2.2.8	Long-Read Based Technology	40

2.2.9	Clinical/Neurological Testing	43
3	Results	47
3.1	FAME.....	47
3.1.1	Identification of the Expansion	47
3.1.2	Characterization of the expansion	52
3.2	Essential Tremor	62
3.2.1	Amplification of the region of interest.....	62
3.2.2	Visualization through Sanger-Sequencing.....	63
3.2.3	TA-cloning	66
3.2.4	Investigation of allele-origin in <i>STARD7</i>	67
4	Discussion	69
4.1	FAME.....	69
4.1.1	Expansions	69
4.1.2	Pathological mechanisms	70
4.1.3	Tandem repeat expansions in neurological diseases	75
4.1.4	Technical realisation and difficulties	76
4.1.5	Perspectives and future studies	78
4.2	Essential tremor.....	79
4.2.1	Repeat expansions and Essential Tremor.....	80
5	Conclusion.....	82
6	Zusammenfassung.....	83
7	References	84
8	Abbreviations and Units.....	93
8.1	Abbreviations	93
8.2	Units	96
9	Figures.....	97

10	Tables	99
11	Supplementary Data	100
11.1	Supplementary Figures.....	100
11.2	Supplementary Tables	113
12	Acknowledgements	119

1 Introduction

1.1 Familial Adult Myoclonic Epilepsy (FAME)

Familial Adult Myoclonic Epilepsy (FAME) is a very rare autosomal dominant disorder characterized by adult-onset cortical myoclonus and seizures. It has been published under numerous names, including autosomal dominant cortical myoclonus and epilepsy (ADCME), benign adult familial myoclonic epilepsy (BAFME) and familial cortical myoclonic tremor associated with epilepsy (FCMTE) as the most common.

Abbreviation	Name
ADCME	autosomal dominant cortical myoclonus and epilepsy
BAFME	benign adult familial myoclonic epilepsy
CrtTr	cortical tremor
FAME	familial adult myoclonic epilepsy
FCMT	familial cortical myoclonic tremor
FCMTE	Familial cortical myoclonic tremor with epilepsy
FCTE	familial cortical tremor with epilepsy
FEME	familial essential myoclonus and epilepsy
FMEA	familial benign myoclonus epilepsy of adult onset
THE	heredofamilial tremor and epilepsy

Table 1. List of abbreviations and names used for FAME.

FAME was first described in the 1980s and 1990s in Japan (Inazuki et al. 1990; Ikeda et al. 1990; Kudo, Kudo, and Yamauchi 1984; Yasuda 1991). Since then, over a hundred pedigrees have been published in different countries all over the world with linkage studies proving the disease to be genetically heterogenic.

1.1.1 Clinical features

The clinical features and diagnostics described in this paragraph have been adapted from (van den Ende et al. 2018; Cen et al. 2016) and (van Rootselaar et al. 2005) if not stated otherwise.

1.1.1.1 Cortical tremor

The first symptom in most FAME patients is a tremor-like, cortical myoclonus of the distal limbs (most often the distal upper extremities) and less frequent facial or axial muscles, with onset in the second or third decade. These movements, often described as tremor, are in truth small, high-frequency myoclonic jerks caused by a cortical hyperexcitability as a rhythmic variant of a cortical myoclonus (Latorre et al. 2020). They most frequently are induced by posture or action, whereas in most patients cortical tremor is not present at rest. The tremor varies in severity between individuals and can also change over time where patients with longer disease duration sometimes have a higher level of cortical tremor severity. Clinical anticipation has also been described, even though mainly in Asian families. Potential stimuli include fatigue, emotional stress, sleep deprivation, alcohol, tactile stimuli, photic or phonic stimulation, glucose deprivation and vibration.

1.1.1.2 Epilepsy

At least 40 percent of the patients affected with FAME show some form of epilepsy in their life. In some individuals, epileptic seizures can also be the first symptom. Compared to Asian patients, European patients are often more severely affected by seizures, requiring poly-antiepileptic therapy more often.

The epilepsy occurring in these FAME patients is mainly characterized by generalized tonic-clonic seizures (GTCS) but focal seizures with impaired awareness or myoclonic seizures can also occur in some patients. GTCS can be preceded by auras or signs such as headache, dizziness, chest distress, palpitation or blurred vision and can be triggered by similar factors as the cortical tremor described in 1.1.1.1. Cortical Tremor.

The frequency of seizures in FAME patients is usually low, with more than five seizures per year being uncommon, even if the patient is untreated. As seizures are often the greatest handicap in FAME patients and the most common reason for them to request medical treatment, the treatment of FAME mostly consists of prevention of epileptic seizures (and reduction of troublesome myoclonus) using anti-epileptic-drugs (AEDs). Valproic acid with or without Clonazepam and Phenobarbital are used most commonly. Most FAME patients respond well to these AEDs and show significantly reduced frequency of GTCS or remain seizure free.

The low frequency of seizures and the good benefit of AEDs are the reason why seizures associated with FAME are often referred to as a benign epilepsy.

1.1.1.3 List of less common clinical features

Less frequently or only associated with single families, other (neurological) symptoms may also occur in FAME patients.

Atypical Seizures	Partial Seizures; Myoclonic Seizures; Intractable seizures; Absences; Frontotemporal and Generalized Interictal EEG Abnormalities
Cerebellar Symptoms	Mild ataxia; Gait instability; Nystagmus; Dysarthria
Psychiatric Symptoms	Depression; Anxiety; Personality Disturbances
Other	Mild cognitive decline; Migraine/Headaches; Visual Intolerance; Frontal Dysfunction; Night Blindness; Motionless State; Parkinsonism; Reduced Verbal Fluency; Visuospatial Impairment

Table 2. List of less common clinical features associated with FAME

1.1.1.4 Diagnostics

1.1.1.4.1 Electrophysiology

To characterize the tremor type and confirm the diagnosis of FAME electrophysiological examinations are used to detect the general cortical hyperactivity and the cortical origin of the tremulous myoclonic movements.

The electromyography (EMG) of a typical cortical myoclonus shows arrhythmic high frequent (> 10 per second) bursts of about 13-20 HZ/50 ms, often synchronous between agonist and antagonist muscles. Furthermore, FAME patients show significant corticomuscular coherence between the EMG and the contralateral EEG, on average peaking at 17 Hz with a consistent phase difference. Phase estimates demonstrate that the cortex leads the EMG abnormalities. The intermuscular coherence mirrors the corticomuscular coherence with a significant coherence between the first dorsal interosseus (FDI) and forearm extensors (Ext) with a peak frequency of about 17 Hz. This intermuscular coherence reflects rhythmic cortical output into the innervated muscles resulting in involuntary tremulous movements (van Rootselaar et al. 2006).

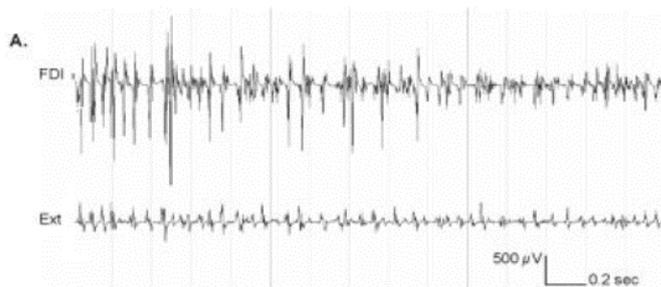


Fig. 1, Bipolar electromyogram (EMG) from right first dorsal interosseus (R FDI) and forearm extensors (Ext) during posture (after applying a high-pass filter of 10 Hz, before rectification). **A:** Familial cortical myoclonic tremor with epilepsy (FCMTE) 4: high frequent bursts of 0.05 seconds (13–18 Hz) (Picture and description taken from: van Rootselaar et al. 2006).

Other features of cortical reflex myoclonus may also be present, including giant Sensory-Evoked-Potentials (SEPs), where the SEP amplitudes are significantly higher in patients at a higher level of cortical tremor severity and compared to those FAME patients using AEDs. These SEP amplitudes can enhance in FAME patients with aging. Furthermore, a consecutive C-reflex can be evoked or enhanced Long-Latency-Reflexes (LLR) can be found. AEDs could sometimes normalize these findings.

Electroencephalography (EEG) findings include epileptiform abnormalities such as generalized (poly) spikes and waves, burst of diffuse slow waves, photoparoxysmal responses and photomyogenic responses. It is described that even in up to 48% of the affected family members who do not have epilepsy, EEG changes can be seen. However, some patients suffering from epilepsy also show normal interictal EEG and as these EEG changes are not specific for FAME they are of limited diagnostic value.

1.1.1.4.2 Structural (and functional) imaging

The structural imaging of the brain of FAME patients using an MRI-scan are normal in most cases, but generalized brain atrophy or slight cerebellar atrophy have been described in a few cases, which may indicate a cerebellar involvement.

Functional MRI-EMG show a reduction in short interval cortical inhibition leading to a cortical hyperactivity in some FAME patients.

1.1.1.4.3 Other diagnostics

Blood, urine, and spinal fluid analysis; funduscopy; electrocardiography; and muscle biopsy in FAME-patients are usually without pathological findings when tested.

1.1.2 Genetics

Up to date six autosomal dominant loci and one autosomal recessive locus, identified through linkage analysis, have been described in FAME.

FAME subgroup	Inheritance	Locus	Families	Confirmed locations
FAME1	autosomal dominant	8q23.3-q24.13	Mainly Japanese pedigrees (Mikami et al. 1999)	Expansion later confirmed
FAME2	autosomal dominant	2p11.1-q12.2	Mainly European pedigrees (Guerrini et al. 2001)	Expansion later confirmed
FAME3	autosomal dominant	5p15.31-p15	European and Chinese pedigrees (Depienne et al. 2010)	Expansion later confirmed
FAME4	autosomal dominant	3q26.32-3q28	One Thai family (Yeetong et al. 2013)	Expansion later confirmed
FAME5	autosomal recessive	Contactin 2 (<i>CNTN2</i>) on 1q32	One Egyptian family (Stogmann et al. 2013)	Not screened for expansion yet/location not confirmed
FAME6	autosomal dominant	<i>TNRC6A</i> on 16p21.1	One Japanese family (Ishiura et al. 2018)	Expansion confirmed
FAME7	unclear/likely autosomal dominant	<i>RAPGEF2</i> on 4q32.1	One Japanese individual (Ishiura et al. 2018)	Expansion confirmed

Table 3. List of FAME loci

Several possible gene mutations have been proposed as causative on the different FAME loci.

This mixed expansion was present in 48 of their 51 studied families and one other family showed an expansion in intron 4 of *SAMD12* where the TTTCA-repeats were placed in the middle of the TTTTA-expansion. In the two Japanese families without *SAMD12* expansion, similar TTTTA/TTTCA-expansions were identified in two other genes: *RAPGEF2* on chromosome 4 and *TNRC6A* on chromosome 16. Again, the TTTCA-repeats were located in the middle of the TTTTA-expansion.

Disease	Chromosome	Gene	Number of patients (families)	Reference sequences in hg19	Configuration of expanded repeats
BAFME1	8q24.11-24.12	<i>SAMD12</i> (repeat configuration 1)	82 (48 families)	(TTTTA) ₇ TTA(TTTTA) ₁₃	(TTTTA) _{exp} (TTTCA) _{exp}
		<i>SAMD12</i> (repeat configuration 2)	3 (1 family)	(TTTTA) ₇ TTA(TTTTA) ₁₃	(TTTTA) _{exp} (TTTCA) _{exp} (TTTTA) _{exp}
BAFME6	16p21.1	<i>TNRC6A</i>	5 (1 family)	(TTTTA) ₁₈	(TTTTA) ₂₂ (TTTCA) _{exp} (TTTTA) _{exp}
BAFME7	4q32.1	<i>RAPGEF2</i>	1 (1 family)	(TTTTA) ₅ TATTA(TTTTA) ₁₂	(TTTTA) _{exp} (TTTCA) _{exp} (TTTTA) _n

Subscripted 'exp' indicates 'expansions'.

Fig 3. Table showing the 2 different expansion-motifs at 3 different expansion-sites in FAME-affected individuals (Table taken from: Ishiura et al. 2018).

Furthermore, the expansion in *SAMD12* showed substantial variation in length and structure between each tested FAME-patient, where the length ranges from approximately 2 to almost 20 kb. Correlation of the mean length of the expansion to the age at onset of epilepsy in each individual showed an inverse correlation. Going further, the length of the expansion also varied within each individual depending on the chosen method of investigation and also in different runs of the same method, showing the difficulties of great expansions as well as raising the question of somatic instability (Ishiura et al. 2018).

This discovery led to the description of two new FAME-loci (FAME6: *TNRC6A* on 16p21.1 and FAME7: *RAPGAF2* on 24q32.1 (Table 3) and the thesis that FAME is caused by intronic TTTTA/TTTCA pentanucleotide repeats in different, apparently unrelated genes.

Confirming this hypothesis, Cen et al. published a report in April of 2018, linking eleven Chinese FAME1 pedigrees and seven unmapped FAME-pedigrees to a TTTTA insertion of at least 150 bp into intron 4 of *SAMD12*. Moreover, they found that a core haplotype containing the TTTCA insertion was shared between their Chinese and three of Japanese pedigrees previously published by Ishiura and collaborators, indicating a founder effect between FAME1 pedigrees from different countries (Cen et al. 2018).

Previous to this discovery, 3 different mutations had been described as causative for 3 FAME loci.

In 2014 De Fusco et al identified a novel in-frame insertion/deletion in the α 2-adrenergic receptor subtype B gene (α 2B-AR; *ADRA2B*) in two apparently non-related Tuscan families previously mapped to the FAME2 locus, that was absent in 575 Tuscan controls. Unfortunately, they failed to detect any *ADRA2B*-mutations in other FAME2-families originating from south Italy in their study (De Fusco et al. 2014). Later these families were linked to a TTTTA/TTTCA repeat expansion in *STARD7*, suggesting the *ADRA2B*-variant to likely be benign (Corbett et al. 2019).

In 2017 Van Rootselaar et al. identified a *CTNND2* missense mutation in a Dutch pedigree (van Rootselaar et al. 2017). Although *CTNND2* was comprised within the FAME3 locus, no mutation had been previously found in the family, which had allowed the description of the locus (Depienne et al. 2010), questioning the causality between the described mutation and the FAME phenotype. In a follow-up study in this family showed a TTTTA/TTTCA repeat expansion in *MARCHF6* (also known as *MARCH6*), again suggesting the *CTNND2*-variant to likely be benign.

In 2013 Stogmann et al identified an autosomal recessive mutation in *CNTN2* segregating with the ET-phenotype in one Egyptian family (Stogmann et al. 2013). The family so far has not been screened for a TTTTA/TTTCA expansion.

1.1.3 Neuropathological findings

Van Rootselaar et al. autopsied the brain of a FAME patient in 2007. They found no morphological changes not corresponding to physiological changes during aging in the cerebrum. The cerebellar cortex and vermis on the other hand demonstrated severe and diffuse loss of Purkinje cells as well as Bergmann gliosis and the remaining Purkinje cells showed an abnormal morphology (van Rootselaar et al. 2007). In 2017 they studied 3 additional brains from individuals carrying a *CTNND2* missense variant who were later on found to carry a repeat expansion in *MARCHF6*. They again observed a loss of Purkinje cells in the cerebellar cortex and abnormal morphology of Purkinje cells during pathological examination (van Rootselaar et al. 2017). This isolated cerebellar pathology in combination with the cortical tremor and seizures were

discussed to be the result of decreased cortical inhibition by cerebello-thalamo-cortical projections. This could be caused by a changed beta-aminobutyric acid (GABA) receptor function, which would be supported by the good response to AEDs such as clonazepam and valproic acid, both affecting GABA receptor function (van Rootselaar et al. 2005).

In 2018 Ishiura et al. autopsied six brains of FAME-affected individuals carrying a TTTTA/TTTCA expansion in *SAMD12*. They found that one individual who had a homozygous expansion showed a mild and diffuse loss of Purkinje-cells and halo-like amorphous materials around the cytoplasm, whereas patients with heterozygous expansions did not show any obvious signs of neurodegeneration. The level of *SAMD12*-transcripts and *SAMD12*-protein was slightly but significantly decreased in all of the autopsied brains. However, since the identification of the expansion in unrelated genes suggested that the decrease in *SAMD12* expression was not the pathogenic mechanism, they also showed UUUCA RNA-foci as well as short reads filled with 5'-TTTCA and 3'-TGAAA repeats accumulated in the brain-tissue and were not found in the liver of the same individuals. This tandem repeat expansion (TRE) and its specific expression in tissue of the central nervous system (CNS) brought new insights into the possibility of the involvement of RNA editing or other yet unknown causative mechanisms in FAME (Ishiura et al. 2018).

1.2 Essential tremor (ET)

Essential tremor is a frequent condition affecting 0,9% in people of all ages and 4,6% in people over 65 years world-wide. It is clinically characterized by bilateral, symmetric posture- and/or action-tremor of the distal (upper) extremities (Louis and Ferreira 2010).

Since ET is much more prevalent than FAME and is clinically very close, it represents the most frequent differential diagnosis to FAME (Bourdain et al. 2006). FAME and ET can be differentiated based on clinical features as well as pathological findings, but many FAME patients can be wrongly diagnosed with ET during their medical history.

Approximately 60% of ET patients have a positive family history, usually suggestive of dominant inheritance, making the distinction between ET and FAME even more difficult. However, ET can also affect the head/face and the voice as the tremor progresses with

age, sometimes leading to a more severe outcome and restriction to daily life compared to FAME-patients. In addition, there are no well-described triggers for the tremor of ET-patients, but in contrary to ET-patients the consumption of alcohol is described to temporarily reduce its severity. Finally, although in some more rare cases ET can be accompanied with ataxia of the extremities or a general impairment of walk, apart from these symptoms ET patients usually don't show any other accompanying neurological or general disabilities. The occurrence of epileptic seizures is therefore not likely to be related to the ET-symptoms (Deuschl 2012: 4 - 7).

Many ET patients are initially diagnosed through their clinical features, which as described can (on first glance) be similar to those of FAME-patients. The instrumental diagnostic procedures on the other hand show the difference between the diseases more clearly. In contrast to FAME, the EMG of most ET patients show rhythmical bursts of > 50 msec with a frequency between 5 and 7 Hz, showing peak of corticomuscular and intermuscular coherence at about 5.9/5.8 Hz, i.e. at tremor frequency. This coherence is exaggerated compared with healthy controls but does not indicate the same cortical mechanism as FAME. It rather reflects a more unspecific mixture of afferent and efferent influences on the sensorimotor cortex, leaving the origin of ET uncertain (van Rootselaar et al. 2006).

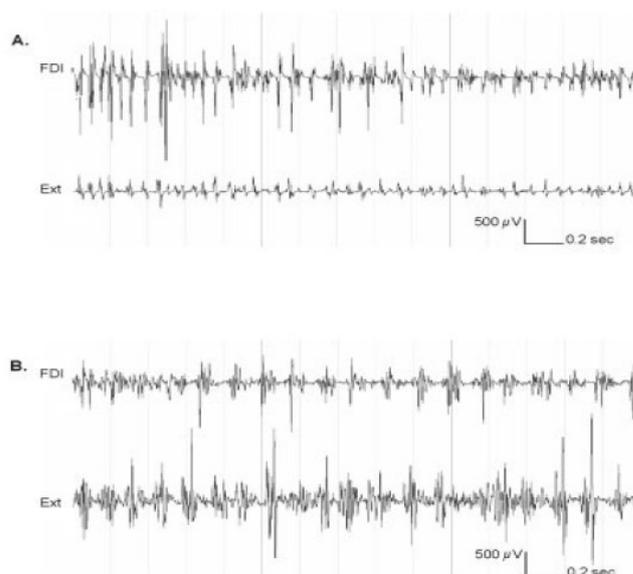


Fig. 4, Bipolar electromyogram (EMG) from right first dorsal interosseous (R FDI) and forearm extensors (Ext) during posture (after applying a high-pass filter of 10 Hz, before rectification). **A:** Familial cortical myoclonic tremor with epilepsy (FCMTE) 4: high frequent bursts of <0.05 seconds (13–18 Hz). **B:** Essential tremor (ET) 6: rhythmic bursts at a frequency of 6 Hz; burst duration is >0.05 seconds (Image and Description from: van Rootselaar et al. 2006).

Neuropathological findings in ET suggest cerebellar changes in Purkinje-cells, such as axonal and dendritic swellings (Yu et al. 2012; Louis 2011; Kuo et al. 2011).

As another main difference beta-blockers, which can be successfully used to treat ET, do not affect cortical tremor. Conversely, anti-epileptic drugs such as Valproate or Clonazepam are no first-line treatment in ET-patients (van Rootselaar et al. 2005).

Characteristics	FAME	ET
Frequency	About 100 families world-wide	0,9 % of the population
Inheritance	Clearly autosomal-dominant	Inheritance likely in 60%
Age at onset	Second to third decade	Fourth decade
Tremor	Posture/Action, Myoclonic jerks, Arrhythmic, High frequent 13-20 HZ, Duration about 50 msec	Posture/Action, Rhythmic, Low frequent 5-7 Hz, Duration > 50 msec
Triggers	Many (1.1.1.1 Cortical tremor)	None
Improvement under alcohol	No	Yes
Related symptoms	Epilepsy, 1.1.1.3 List of less common features	Rarely Ataxia/Impairment of walk
Instrumental diagnostics	Cortical origin	No clear cortical origin
Neuropathology	Changes in Purkinje-cells, Changes in relation to TTTTA/TTTCA-expansion (1.1.1.4.3 Neuropathological findings)	Changes in Purkinje-cells
Therapy	AEDs	Beta-blockers

Table 4. Comparing factors between FAME and ET (Bourdain et al. 2006; van Rootselaar et al. 2006; Deuschl 2012).

1.3 Aim

1.3.1 FAME

Over the course of 20 years of research regarding FAME, the scientific group lead by Prof. Dr. Christel Depienne (Institute of Human Genetics, University-Hospital Essen, University Duisburg-Essen, Germany) gathered a collection of eleven FAME families, the two largest families (Family 1 and Family 5) linked to the FAME2 and FAME3 loci. Before the study, these two largest families had been extensively studied by sequencing the coding regions as well as the whole regions contained in both loci and by analysing the mRNA of blood cells and lymphoblasts using RNA-sequencing. However, these strategies had failed to identify pathogenic mutations in both families suggesting an unconventional or noncoding causative mutation.

When I was introduced to the FAME project in the summer of 2017, applying for a position as a doctoral candidate starting after my second state-exam in April 2018, the aim of my study was to identify the causative variant responsible for FAME, by studying non-coding regions.

The discovery of Ishiura et al. published in January of 2018, that FAME is caused by an intronic expansion of TTTTA/TTTCA pentanucleotide repeats in different, apparently unrelated genes on chromosomes 8 (*SAMD12*), 4 (*RAPGAF2*), and 16 (*TNRC6A*) in Japanese families (Ishiura et al. 2018), allowed to accelerate the identification of the pathogenic expansions causing FAME in our families and modified the course of this study .

Going back to the Illumina genome data generated in 2017, expansions similar to those identified by Ishiura et al. were identified using bioinformatic approaches in intron 1 of *STARD7* for FAME2 and intron 1 of *MARCHF6* for FAME3 in our two families just before the beginning of my internship. Based on these findings, the main aim of the study was to confirm the presence of the expansion in the families provided with the genome data and to screen the remaining 9 families for an expansion in *MARCHF6* or *STARD7*, hereby confirming the theory of Ishiura et al. that FAME is caused by an intronic expansion of TTTTA/TTTCA pentanucleotide repeats in different, apparently unrelated

genes (Ishiura et al. 2018) not only in Asian families, but possibly in all families with the FAME phenotype.

To further prove the role of the TTTTA/TTTCA expansions in FAME, we aimed to characterize the length and structure of the expanded region in more detail and detect possible variations within and between individuals. Furthermore, the length of the regions were correlated to the severity of the phenotype.

1.3.2 Screening for expansions in *MARCHF6* and *STARD7* in ET-patients

As described in 1.2.1 Essential Tremor, ET and FAME show similarities, especially in earlier stages of the disease and we thought it to be likely that in some cases FAME patients are possibly misdiagnosed with ET as it is the far more frequent and thus more known disease. Therefore, we aimed to investigate whether ET patients had an expansion in one of the two main European FAME loci.

2 Materials und Methods

2.1 Materials

2.1.1 General Equipment

Name	Manufacturer
Centrifuges:	
Color Sprout Mini-Zentrifuge	Biozym
Mikro 120 Zentrifuge	Hettich Zentrifugen
Avanti™ J-20 XP Centrifuge	Beckman Coulter™
Vortex Machines:	
Vortex Genie 2	Bender & Hobein AG
Vortex Mixer	neoLab
Vortex mixer	Yellowline
Precisa 1212 MSCS	Göntgen
Thermomixer Comfort	Eppendorf
G25 Incubator Shaker	New Brunswick Scientific CO. INC.
Herasafe Safety Cabinet	Thermo Fisher Scientific
NanoDrop 1000	Thermo Fisher Scientific
NanoDrop 2000	Thermo Fisher Scientific
Veritti 96 Well Thermal Cycler	Applied Biosystems
Gel Chamber H5	BRL
Gel-Documentation M035206	Bachofer/Peqlab
3130XL Genetic Analyzer	Applied Biosystems
Micro Amp Reaction Tubes	Applied Biosystems
Micro Amp 8 Cap-Strip	Applied Biosystems
Pipettes	GILSON
Pipetting Tips:	
Tip 10/20 µl XL Graduated Filter Tip	StarLab
Tip 20 µl Bevelled Filter Tip	StarLab
Tip 200 µl Graduated Filter Tip	StarLab

Tip 1000 µl Filter Tip	StarLab
DISTRIPTIPS Micro ST	GILSON
DISTRIPTIPS Mini ST	GILSON
DISTRIPTIPS Maxi ST	GILSON
PCR Plate 96 Well	Thermo Fisher Scientific
Plate Septa, 96 well	Applied Biosystems
PRECISION Wipes	KIMTECH

2.1.2 Chemicals, Buffer and Enzymes

2.1.2.1.1 Standart Chemicals, Buffer and Enzymes

Name	Components	Manufacturer
TAE-Buffer	40 mM Tris-Acetat; 1 mM EDTA	Appli Chem
H ₂ O	H ₂ O Bidest Merck	Lichrosolv®
MgCl ₂	-	Life Technologies
EtOH	-	Merck
Hi-Di Formamide	-	Life Technologies
GeneAmp dNTPs	-	Life Technologies
peqGOLD Universal Agarose	-	Peqlab
Ethidium Bromide Solution	-	Invitrogen
Loading Dye: 6x Xylenblau	0,25% Xylenblau, 15% Ficoll 400; 10 mM EDTA	-
DNA/Protein-marker: MassRuler™ DNA Ladder, Low Range	-	Thermo Fisher Scientific

2.1.2.2 Specific Chemicals, Buffer and Enzymes

2.1.2.2.1 PCR

Name	Manufacturer
GoTaq G2 DNA-Polymerase Kit	Promega
HotStar Taq Master Mix Kit	QIAGEN
GeneAmp dNTPs 4x 320µl	Life Technologies

2.1.2.2.2 Fragment Length Analysis:

Name	Manufacturer
GeneScan -1200 LIZ Size Standard	Life Technologie

2.1.2.2.3 Sequencing:

Name	Manufacturer
ExoSap-IT, 5000 Reactions	Affymetrix
BrilliantDye Terminator Cycle Sequencing Kit	Nimagen

2.1.2.2.4 Sephadex-plate:

Name	Manufacturer
Sephadex G-50 100g Fine	GE Healthcare
MultiScreen 45µl Columns Loader	Millipore
MultiScreen-HV Plate	Millipore

2.1.2.2.5 TA-cloning: pcDNAtm 3.1/V5-His TOPO TA Expression Kit (Invitrogen)

Kit-Name	Reagent	Components	Manufacturer
pcDNA tm 3.1/V5-His TOPO TA Expression Kit	Salt Solution	1,2 M NaCl; 0,06 M MgCl ₂	Invitrogen
	TOPO vector	10ng/µl Plasmid in: 50% glycerol; 50	

		mM Tris-HCl pH 7,4; 1 mM EDTA; 2 mM DTT; 0,1% Triton® X-100; 100 Mg/ml BSA; 30 µM Phenol red	
	SOC-medium	2% Tryptone; 0,5% Yeast Extract; 10 mM NaCl; 2,5 mM KCl; 10 mM MgCl ₂ ; 10 mM MgSO ₄ ; 20 mM glucose	
	TOP10 cells 50 µl		
-	LB medium LB-medium pulver	25 g/l	ROTH
-	LB plates	40g/l LB Agar and Ampicillin 100µg/ml	ROTH

2.1.2.2.6 Mini-preparation:

Reagent	Components	Manufacturer
P1-Buffer	50mM Tris-Cl, pH 8.0, 10mM EDTA, 100ug/mL RNase A	QIAGEN
P2-Buffer	200mM NaOH, 1% SDS	QIAGEN
P3-Buffer	3.0M potassium acetate, pH 5.5	QIAGEN
BamHI-HF		NEB
EcoRV-HF		NEB
CutSmart		NEB

2.1.2.2.7 Molecular Combing:

Name	Manufacturer
DNA Clean & Concentration 5	Zymo

2.1.3 Software

Name	Manufacturer
Geneious	Biomatters
GeneMapper	Thermo Fisher Scientific
Thermo Fisher Cloud	Thermo Fisher Scientific

2.1.4 Primers

The following primers were designed by our collaborators, the group of Mark Corbett and Jozef Gecz, University of Adelaide, Australia., or using the Primer 3 software (<http://primer3.ut.ee/>) and synthesised and ordered via Metabion (Planegg).

2.1.4.1 Primers for Fragment Length Analysis and Sanger Sequencing

2.1.4.1.1 Human Primers

Locus	Name	Sequence
Chromosome 2 <i>/ STARD7</i>	FAM-RP_FAME2-P1-F	GGCTACTTACGTGCCAGATAAC
	RP_FAME2-P1-F	GGCTACTTACGTGCCAGATAAC
	RP_FAME2-P2-R	TGCCCAGCTACACTGTCTCTT
Chromosome 5 <i>/ MARCHF6</i>	FAM-RP_FAME3-P2-F	GGAAAAGGGAGGGTTATAGAGGA
	RP_FAME3-P2-F	GGAAAAGGGAGGGTTATAGAGGA
	RP_FAME3-P1-R	CCATCAGAGGCAAGCAATGT
Unspecific to Locus	P3-PU	TACGCATCCCAGTTTGAGACG
	FAM-P3-PU	TACGCATCCCAGTTTGAGACG
	RP_FAME-P5.1-F	TACGCATCCCAGTTTGAGACGTTTTATT TTATTTTATTTTATTTTATTTTATTTA
	RP_FAME-P5.2-F	TACGCATCCCAGTTTGAGACGTTTCATTT CATTTCATTTCATTTTC

2.1.4.1.2 Monkey Primers Chromosome 2 / *MARCF6*

Species	Name	Sequence
Chimpanzee	RP_FAME2-P1-F	GGCTACTTACGTGCCAGATAAC
	RP_FAME2-P2-R	TGCCCAGCTACACTGTCTCTT
Gorilla	RP_FAME2-P1-F	GGCTACTTACGTGCCAGATAAC
	RP_FAME2-P2-R	TGCCCAGCTACACTGTCTCTT
Orangutan	RP_FAME2-P1-F	GGCTACTTACGTGCCAGATAAC
	RP_FAME2-P2-R- Ora_Gib	CGCCCAGCTACACTGTCTCTT

2.1.4.1.3 TA-clones

These primers were provided in the pcDNA tm 3.1/V5-His TOPO TA Expression Kit by Invirogen.

Direction	Name	Sequence
Foreward:	T7	TAATACGACTCACTATAGGG
Reverse:	BGH	TAGAAGGCACAGTCGGAGG

2.1.4.2 Primers for Molecular Combing

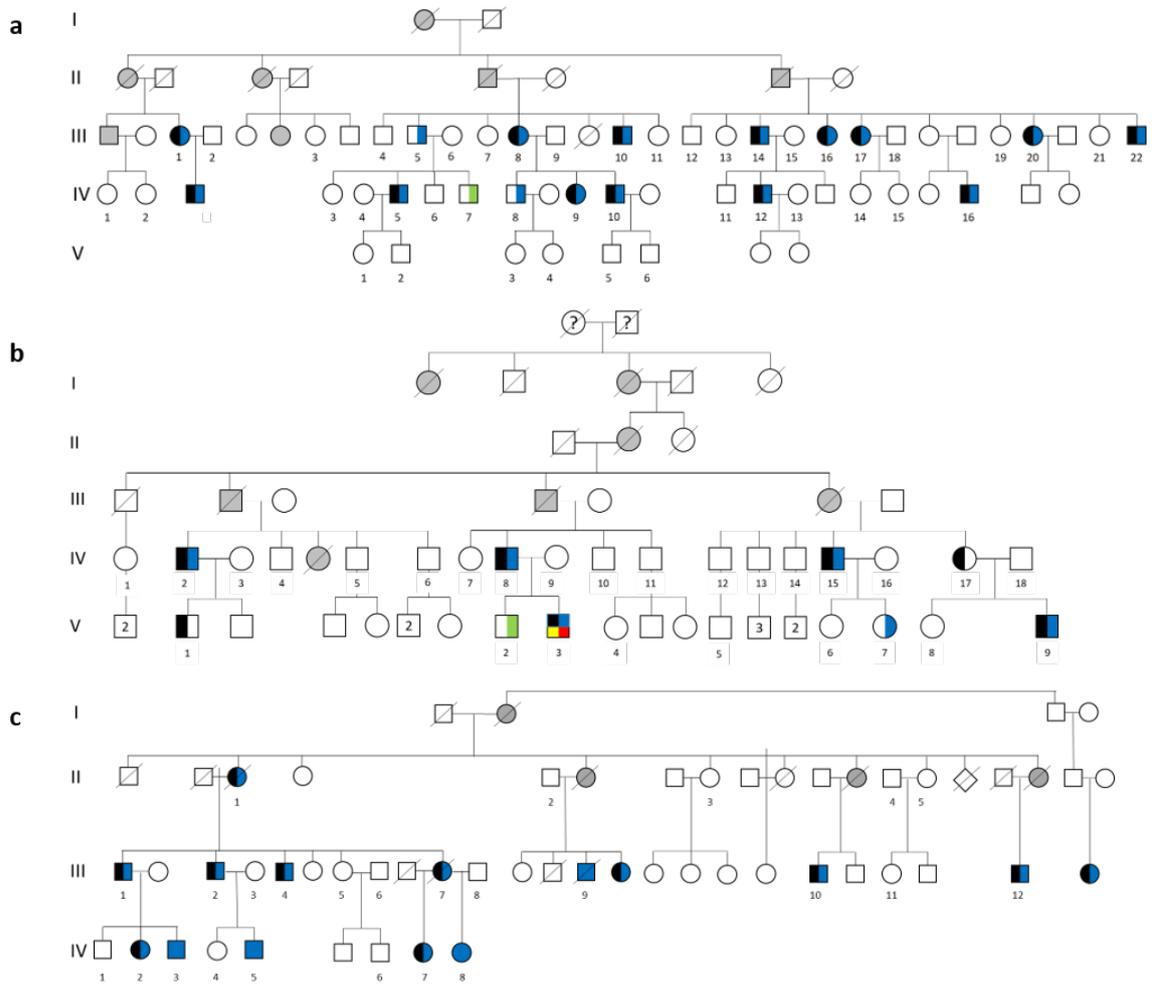
Ten pairs of overlapping primers flanking the candidate region identified on chromosome 5 were used for molecular combing.

2.1.5 Patients and Probes

All of the samples studied were obtained from blood or saliva. Most of them were kindly provided from the cell bank of ICM in Paris or collaborators of the Institute of Human Genetics, University Hospital Essen, University Duisburg-Essen, Essen.

One hundred-twenty-seven individuals from 12 families, linked to FAME were analysed. There were 4 larger families with 15, 26, 34 and 44 members and 8 smaller families ranging from 2 to 7 individuals per family. The samples of these families were provided

by the DNA and cell bank ICM in Paris, Prof. Arn van den Maagdenberg (University of Leiden, Netherlands) and Dr. Gabrielle Rudolf (University of Strasbourg, France). Four of the 44 members of the biggest family were resampled by Dr. Eloi Magnin (Hôpital de Besancon, France) during the study. Twenty-three of the 34 individuals of the second largest family were resampled by Prof. Stephan Klebe (University of Essen, Germany) and myself in Chartres, France and 21 of these were examined neurologically in collaboration with Prof. Dr. Stephan Klebe at this occasion.



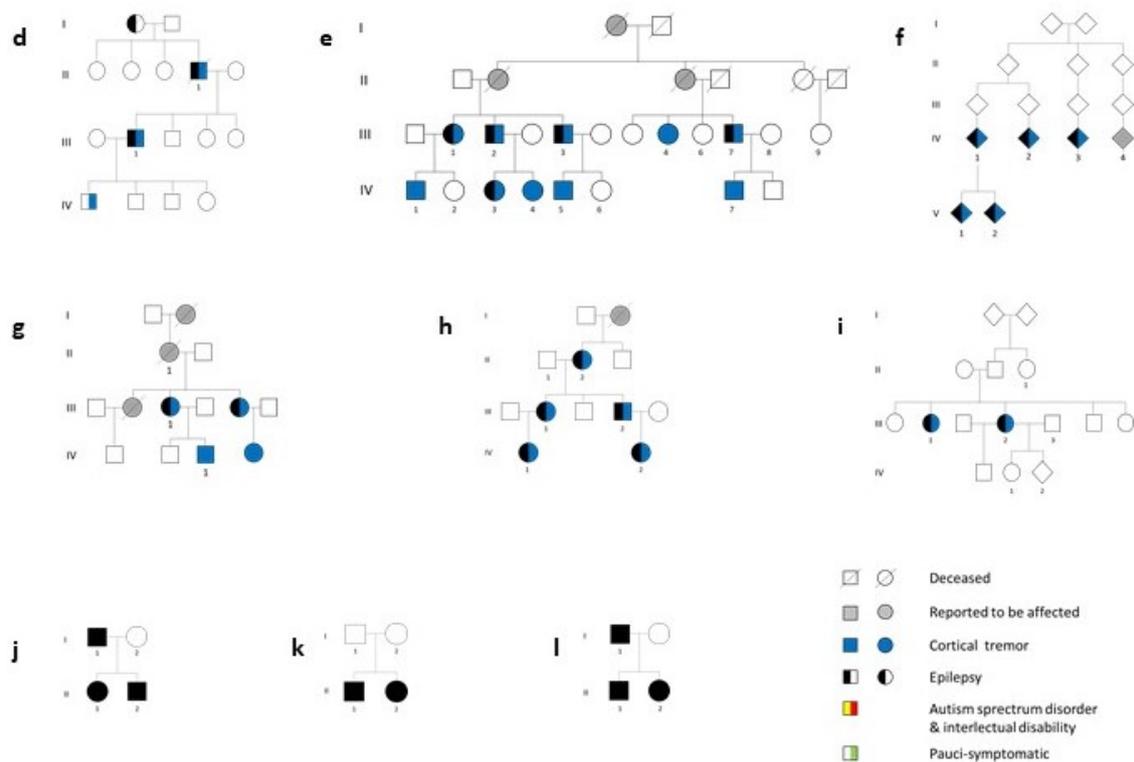


Fig 5, Pedigrees of all included FAME-families. **a)** Family 1, **b)** Family 2, **c)** Family 3, **d)** Family 4, **e)** Family 5, **f)** Family 6, **g)** Family 7, **h)** Family 8, **i)** Family 9, **j)** Family 10, **k)** Family 11, **l)** Family 12

Furthermore, we obtained 129 ET-samples from 29 families provided by Dr. Giovanni Stevanin (ICM). (Supplementary Fig.1)

Five orangutan-, 5 chimpanzee- and 2 gorilla-samples were used to analyze the conservation of our candidate region during recent evolution. The monkey-samples were taken from the blood-stock of the Institute of Human Genetics, University-Hospital Essen, University Duisburg-Essen.

As control individuals, 30 blood-donor samples from the Institute of Human Genetics, University-Hospital Essen were used.

2.2 Methods

Most of the described methods, if not stated differently, were performed in shared work with Sabine Kaya of the Institute of Human Genetics, University-Hospital Essen, University Duisburg-Essen, Essen.

2.2.1 Isolation of genomic DNA from Blood

Most of the DNA used was kindly provided by collaborators of the cell bank of ICM in Paris. The preparation of the blood-samples of the newly sampled individuals from the LEC-family was done by Sabine Kaya and Claudia Mertel in the Institute of Human Genetics, University-Hospital Essen, University Duisburg-Essen, Essen.

2.2.2 Concentration Measurement of Nucleic Acids

To determine the concentration of a DNA-probe the photometric absorption of an aliquot of a probe was measured at $\lambda = 260$ nm and $\lambda = 280$ nm in a Spectrophotometer. An optic density (OD) of 1 at 260 nm correlates to a concentration of 50 $\mu\text{g/ml}$ double-stranded DNA, therefore the OD_{260} -score was multiplied by 50 to obtain the DNA-concentration in $\mu\text{g/ml}$.

In order to evaluate the purity of the probe the ratio $\text{OD}_{260} / \text{OD}_{280}$ was used. A ratio of 1.8 indicates pure DNA, lower values indicate the presence of substances such as proteins in the probe.

2.2.3 Polymerase Chain Reaction (PCR)

The amplification of the regions of interest from genomic DNA was performed by PCR.

To set up the PCR-Reaction the HotStarTaq Master Mix Kit (Qiagen) or the GoTaq G2 DNA-Polymerase Kit and a dNTP-Mix was used. The reaction-mixtures were prepared as following, with total reaction-volumes of 25 μl :

1.)

Component	Volume
dNTP (1,25 mM)	2 μ l
GoTaq Reaction Buffer (5x)	5 μ l
Primer F (20 pmol/ μ l)	0,5 μ l
Primer R (20 pmol/ μ l)	0,5 μ l
GoTaq DNA-Polymerase (5 U/ μ l)	0,25 μ l
H ₂ O	11,75 – 12,75 μ l
DNA (25 ng/ μ l)	2-3 μ l

2.)

Component	Volume
HotStarTaq	12,5 μ l
Primer R (20 pmol/ μ l)	1 μ l
Primer F (20 pmol/ μ l)	1 μ l
H ₂ O	7,5 – 8,5 μ l
DNA (25 ng/ μ l)	2-3 μ l

When used for fragment length analysis, one of the primers was labelled with FAM fluorescence. Primers used for Sanger sequencing were not labelled.

The following two sets of conditions were used for the PCR:

Set 1:

- 1.) Initial Denaturation: 95°C – 2 Min
- 2.) x35 Denaturation: 95°C – 30 Sec
Primer-Annealing: 58°C – 30 Sec

	Elongation:	72°C – 50 Sec
3.)	Final Elongation:	72°C – 7 Min
	Cooling:	4°C – ∞

Set 2 (TouchDown):

1.)	Initial Denaturation	95°C – 2 Min
2.) x12	Denaturation:	95°C – 30 Sec
	Primer-Annealing:	66°C – 30 Sec (-0,5°C per cycle starting from cycle 2)
	Elongation:	72°C – 50 Sec
3.) x38	Denaturation:	95°C – 30 Sec
	Primer-Annealing:	60°C – 30 Sec
	Elongation:	72°C – 50 Sec
4.)	Final Elongation	72°C – 7 Min
	Cooling:	4°C – ∞

2.2.3.1.1 Repeat-Primed PCR

To confirm the presence of the pentanucleotide-repeat expansions in affected individuals repeat-primed PCR (RP-PCR) was performed targeting TTTCA or TTTTA.

The P1F/P2F Forward-primers as well as the P1R/P2R Reverse-primers were region-specific, flanking the pentanucleotide repeat region and labelled with FAM fluorescence. Primers P4 and P5 consisted of a part containing TTTTA/TTTCA repeats at the 3'-end and a part corresponding to a universal sequence not present in any part of the human genome at the 5'-end. These primers annealed randomly with repeats present in the *MARCF6/STARD7*-genes and therefore amplified products of different length depending on where these primers primed within the expansion. The P3-PU primer was identical to the universal tail of the P4/P5 primers and used to improve the amplification of the products generated with the P4/P5 primers. As amplification of larger fragments is less efficient compared to smaller fragments, the resulting PCR leads to multiple amplicons decreasing with size, visualized on GeneScan (Ishiura et al. 2018; Singh et al. 2014).

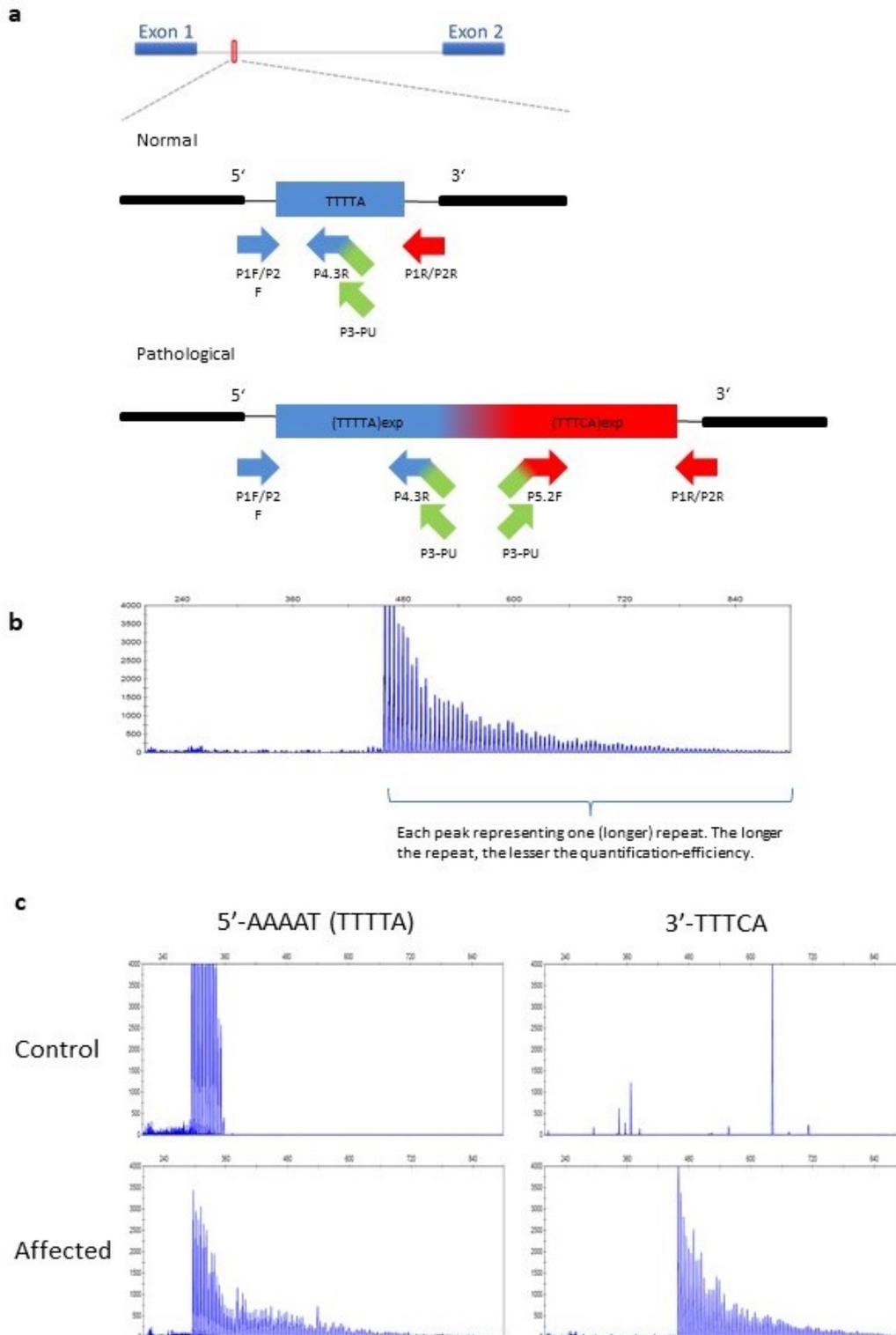


Fig 6. a Schematic representation of primers binding within the normal or the expanded allele, P1F/P2F and P1R/P2R are region-specific, P4 and P5 binding within the TTTTA/TTTCA repeats, P3-PU is used to improve the amplification of the products generated with the P4/P5 primers; **b** amplicons decreasing with size due to inefficiency of amplification of larger fragments, visualized with GeneScan; **c** Examples of affected and control individuals, control individuals show view repeats in TTTTA-amplification and no repeats in TTTCA amplification, affected individuals show multiple repeats in TTTTA and TTTCA amplification, visualize with GeneScan.

The PCR reaction, with a total reaction-volume of 25 μ l, contained:

Component	Volume
HotstarTaq	12,5 μ l
MgCl ₂ (25 mM)	2 μ l
Gene-specific flanking primer (P1F/P2F/P1R/P2R) (20 pmol/ μ l)	1 μ l
P3-PU (20 pmol/ μ l)	1 μ l
Unspecific primer (P4/P5) (20 pmol/ μ l)	1 μ l
H ₂ O	5,5 – 6,5 μ l
DNA (25 ng/ μ l)	2-3 μ l

The performed PCR-conditions were established as following by Prof. Dr. Christel Depienne and Sabine Kaya through repeated testing in various conditions in combination with the conditions provided by the manufacturer-manuals and the RP-PCR established by (Ishiura et al. 2018).

- | | | |
|---------|----------------------|--|
| 1.) | Initial Denaturation | 95°C – 5 Min |
| 2.) x10 | Denaturation: | 95°C – 30 Sec |
| | Primer-Annealing: | 48°C – 45 Sec (+1°C per cycle starting from cycle 2) |
| | Elongation: | 65°C – 50 Sec (+1°C per cycle starting from cycle 4) |
| 3.) x30 | Denaturation: | 95°C – 30 Sec |
| | Primer-Annealing: | 58°C – 1 Min |
| | Elongation: | 72°C – 5 Min |
| 4.) | Final Elongation | 72°C – 7 Min |
| | Cooling: | 4°C – ∞ |

2.2.4 Agarose gel electrophoresis

PCR-products were controlled on 1% to 2% agarose gels to verify the concentration of the products and the specificity of the PCR. The agarose was heated in 1x TAE-buffer and 0,4 µg/ml ethidium bromide (EtBr) was added after cooling. 5 µl of the samples amplified with the HotStar Taq were added with 5 µl of a loading dye (1x Cylene-Blau) prior to loading. 5-7 µl of the mix or the samples amplified with the GoTaq Green were loaded on the gel. The gel was run in 1x TAE-buffer with 0,4 µg/ml EtBr at 120 V. The visualisation with a UV-transilluminator ensued and was documented through photodocumentaries (Gel-Documentation M035206; Bachofer/Peqlab).

The concentration was estimated in comparison a specific ladder (MassRuler™ DNA Ladder, Low Range; Thermo Fisher Scientific).

2.2.5 Fragment length analysis (GeneScan)

2.2.5.1 PCR

All of the previously described PCRs have been used for Fragment length analysis, always with a FAM-marked Primer present.

2.2.5.2 Capillary electrophoresis

Analysis of the length of the sequences containing the possibly expanded region was done using fragment length analysis using Applied Biosystems 3130xl Genetic Analyzers.

In preparation the reaction-mixtures of 10 µl of Hi-Di Formamide and 0,5-2 µl of the Size Standard GS1200LIZ (68 fragment ranging from 20 bp to 1200 bp) were added to each well of a 96 Well PCR Plate (Thermo Fisher Scientific) containing 1 to 5 µl of PCR-product depending on the individual signal-intensity of the samples on the agarose-gel under ultraviolet-light.

The following Instrument Protocol was used: GS_POP6_20s_8000s

Fragment length analysis separated the fragments, amplified through PCR, by size using capillary electrophoresis. A high-voltage charge was applied to the PCR products which

forced the negatively charged fragments to move towards the anode and into the capillaries filled with a polymer (POP6, Thermo Fisher Scientific). As the size of the fragment determines its total charge, the speed of the travel of a fragment is inversely proportionate to its size. Therefore, the DNA-fragments were separated by size while moving through the capillaries. Shortly before reaching the anode, they passed a laser beam, which caused the fluorescent markers attached to the fragments to fluoresce. These fluorescent signals were detected and converted into digital data automatically (Applied Biosystems 2009).

The analysis software afterwards established a relative size for each sample through comparison of the standard curve determined using the LIZ1200 applied to each probe.

2.2.5.3 Data analysis

The analysis of the raw data of the fragment length analysis as well as the preparation of graphs that visualised the GeneScan results were performed using the Software-Tool GeneMarker from SoftGenetics as well as the ThermoFisherCloud from Thermo Fisher Scientific.

2.2.6 Sanger Sequencing

2.2.6.1 Sephadex-plate

In Order to perform Sanger Sequencing a Sephadex-plate was prepared, filling the Sephadex-powder into a MultiScreen-HV plate using the MultiScreen 45 µl Loader. To each well 300 µl sterile, nuclease free H₂O were added, the columns were incubated at 4 °C for at least 3 hours and centrifuged at 900 x g for 5 minutes to remove the excess of water before use.

2.2.6.2 Sanger Sequencing PCR's

PCR-Reactions were performed as described in 2.2.3 Polymerase Chain Reaction using the HotstarTaq and Primers that were not labelled with FAM with the TouchDown PCR protocol to amplify the target sequence before proceeding with cycle sequencing. Following the approximate concentration was determined with an agarose gel.

To remove the excess of primers and unincorporated nucleotides, 5 µl PCR-products were purified with 2 µl of ExoSAP-IT in a ThermalCycler with the following conditions provided by the manufacturer manuals:

- 1.) Enzyme incubation: 37 °C 15 Minutes
- 2.) Enzyme inactivation: 80 °C 15 Minutes
- 3.) Cooling: 4 °C ∞

The purified products were used in a sequencing PCR reaction which was prepared with the BrilliantDye Terminator Cycle Sequencing Kit as follows: 2 µl BrilliantDye, 2 µl reaction buffer, 1 µl Primer (Forward and Reverse primers separately in an own reaction each) and 1-5 µl purified product depending on the approximate concentration of each probe in the first PCR. Up to 4 µl of sterile, nuclease free water were added to a total reaction-volume of 10 µl. The conditions used in the ThermalCycler which were established by Prof. Dr. Christel Depienne and Sabine Kaya through repeated testing of various conditions in consideration of the manufacturer manuals, consisted of:

- 1.) Initial Denaturation: 96°C – 1 Min
- 2.) x25 Denaturation: 96°C – 10 Sec
- Primer-Annealing: 59°C – 5 Sec
- Elongation: 60°C – 4 Min
- 3.) Cooling: 4°C – ∞

Sanger Sequencing uses 2'-3'-dideoxynucleotides (ddNTPs), each base labelled with a different fluorescent dye, alongside deoxynucleotides (dNTPs). In a conventional PCR or other DNA amplification the product of amplification grows in 5' to 3' direction by forming a phosphodiester bridge between the 3'-hydroxyl group of a deoxynucleotide at the growing end of the primer and 5'-phosphate group of the next nucleotide.

The 2'-3'-dideoxynucleotides lack the 3'-hydroxyl group and therefore the next nucleotide cannot bind, putting an end to the elongation. Adding both kinds of nucleotides, each new base added can either terminate the elongation process for the individual fragment or not resulting in multiple fragments of various lengths. Each fragment with a labelled nucleotide at the 3'-end (Applied Biosystems 2009).

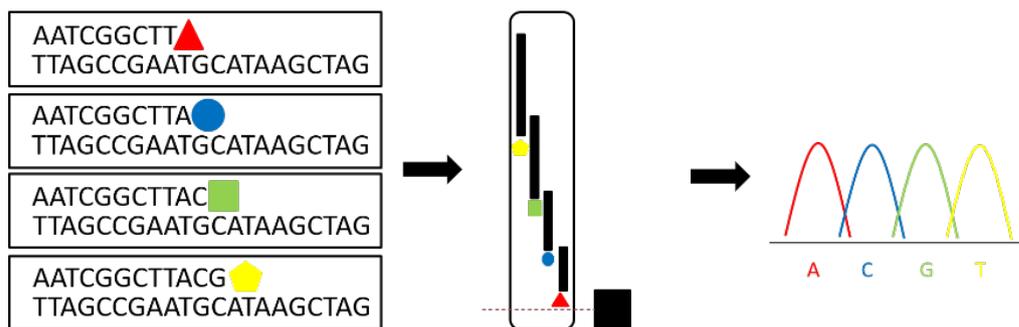


Fig 7. Schematic visualization of sanger sequencing. ddNTPs hereby labelled as red square (A), blue circle (G), green square (G) and yellow pentagon (T). Each fragment passing the capillary tube sorted by size and therefore passing the laser beam in the correct (sizedependent) order from smallest to longest.

2.2.6.3 Capillary electrophoresis

The sanger sequencing analysis was performed with one of Applied Biosystems 3130xl Genetic Analyzers (Thermo Fisher Scientific).

In preparation the 10 μ l PCR-product of the sequencing PCR were loaded onto the sephadex columns and centrifuged at 900 x g for 5 minutes into a Thermo-Fast 96 PCR Detection Plate (Thermo Fisher Scientific) already loaded with 10 μ l of Hi-Di formamide per well. This again, was used to purify the PCR-products.

During the sanger sequencing analysis, the individual DNA fragments per well were separated by size using capillary electrophoresis according to the same principle explained in 2.2.5.2 capillary electrophoresis. The different sized fragments each labelled with a marker specific to a base therefore move past the laser-beam in order, from the shortest to the longest fragment and the software can put an order to the bases in a sequence.

2.2.6.4 Data analysis

Data analysis of the raw data from the Sanger sequencing analysis and the preparation of graphs was performed using the Geneious-software.

2.2.7 Mini Prep/TA cloning

2.2.7.1 Cloning Reaction and Transformation

During our Mini-Preparation/TA-cloning, a Taq polymerase-amplified PCR product (our region of interest) was inserted into a plasmid vector. To do so, the Taq polymerase adds a single deoxyadenosine (A) to the 3' end of the PCR-product during amplification. This product is introduced to a plasmid vector (pcDNA[™] 3.1/V5-His TOPO) which is supplied with a single 3' thymidine (T) overhang and a covalently bound Topoisomerase 1. Hereby the tyrosyl residue of the Topoisomerase 1 forms a covalent bond with the phosphodiester of the thymidine of the vector. When the PCR-product with the adenosine overhang meets the vector, the bond between the vector and the Topoisomerase is attacked by the free 5' hydroxyl of the PCR-product, resulting in the releasing of the Topoisomerase and the ligation of the PCR product into the vector (Invitrogen 2009).

TA cloning was performed using the pcDNA[™] 3.1/V5-His TOPO TA Expression Kit (Invitrogen). Fresh PCR-products (as described in 2.2.3. Polymerase Chain Reaction) were used as a source for DNA.

The TOPO cloning reaction was performed with 2 µl PCR product added with 1 µl Salt Solution, 1 µl TOPO vector and sterile water to an overall volume of 5 µl. The mixture was mixed gently on ice and afterwards incubated at room temperature (22-23°C) for 5 minutes. 2 µl of this cloning reaction were then added into 50 µl of a "One Shot TOP10 Chemically Competent E.coli"-solution for chemical transformation. The mixture incubated for 30 minutes on ice, was afterwards heat-shocked for 30 seconds at 42°C and then immediately transferred on ice, where it rested for 5 minutes. 250 µl of SOC-medium were added, the reaction was incubated at 37°C and shaken at 200 rpm for 1 hour. Last 50 and 150 µl of the reaction were spread on an LB-plate with a Drigalski spatula and incubated at 37°C overnight. The next day 3-5 colonies were picked, inoculated with 3 ml LB-Medium and 3 µl Ampicillin per colony and incubated at 37°C and 230 rpm overnight.

2.2.7.2 Mini-preparation of plasmids

The preparation of the plasmids was performed as following: 1,5 ml of the overnight culture were pipetted into an Eppendorf-tube and centrifuged at 13.000 rpm for 1 minute. The supernatant was discarded, the pellet was pipetted into 300 μ l of the Resuspension Buffer (P1-buffer) and resuspended through vortexing. 300 μ l of the Lysis Buffer (P2-buffer) were added and incubated for 3-4 minutes at room-temperature. Then 300 μ l of the Neutralization buffer (P3-buffer) were added and mixed in carefully before centrifuging for 10 minutes at 13.000 rpm at room temperature. The supernatant was then pipetted in a new tube containing 500 μ l EtOH and again centrifuged for 10 minutes at 13.000 rpm at room temperature. The new supernatant was then discarded and the pellet was washed with 300 μ l of 70% EtOH. Afterwards it was centrifuged at 13.000 rpm for 5 minutes. The supernatant was discharged and the pellet dried at room temperature for up to 30 minutes. The pellet was applied in 30 μ l H₂O, rested for 15 minutes and was vortexed to blend at last.

2.2.7.3 Enzymatic Digestion

For the enzymatic digestion the following components were incubated at 37 °C for 120 minutes in a Thermal Cycler.

Component	Volume
Mini-Preparation Product	2 μ l
Bam H1 (20 U/ μ l)	0,1 μ l
EcoR5 (20 U/ μ l)	0,1 μ l
CutSmart (10x)	1 μ l
H ₂ O	6,8 μ l

2.2.7.4 Sequencing of clones

The results of the enzymatic digestion was used in a sequencing PCR with the BrilliantDye Terminator Cycle Sequencing Kit, mixed as described under 2.2.6.2 Sanger

Sequencing PCR's. T7 and BGH were used as primers in a specific reaction each, with a concentration of 0,1 µg/µl in TE Buffer. Each primer was used with its own reaction-conditions a ThermalCycler as following:

T7:	1.)	Initial Denaturation:	96 °C – 1 Min
	2.) x25	Denaturation:	96 °C – 10 Sec
		Annealing:	47 °C – 5 Sec
		Elongation:	60 °C – 4 Min
	3.)	Cooling:	4 °C - ∞
BGH:	1.)	Initial Denaturation:	96 °C – 1 Min
	2.) x25	Denaturation:	96 °C – 10 Sec
		Annealing:	55 °C – 5 Sec
		Elongation:	60 °C – 4 Min
	3.)	Cooling:	4 °C - ∞

The sequencing reactions were performed as described under 2.2.6 Sanger Sequencing.

2.2.8 Long-Read Based Technology

Long read based technology such as Nanopore Sequencing and Molecular Combing was performed to estimate the size of the expanded sequences and the distribution of the TTTTA- or TTTC A-sequence within the expansion. The estimated length was then correlated to the severity of the phenotype.

Furthermore long-read based technology was used to make a statement about the presence of somatic mosaicism.

2.2.8.1 Nanopore Sequencing

The genome of 6 individuals of Family 1 and 2 were sequenced with nanopore sequencing technology. The sequencing was kindly performed through our collaborator Dr. Florian Kraft from the Institute of Human Genetics, University-hospital Aachen.

During nanopore sequencing the DNA of a probe is forced through a biological nanopore imbedded into an electrically resistant artificial membrane using an electric charge.

During the sequencing a helicase acts as a motorprotein and unzips the dsDNA to allow 1D reads. It also slows down the translocation of the DNA through the Nanopore, optimizing the reads. Every base has a specific current that causes a deflection in the current across the pore, therefore generating a specific electrical profile while moving through the pore. The pore can then generate a signal that corresponds to five bp at a time using a sensor. Each MinION flowcell consists of 512 sensors, which are connected to four pores each. This technology enables reads up to 50 bp (Leggett and Clark 2017).

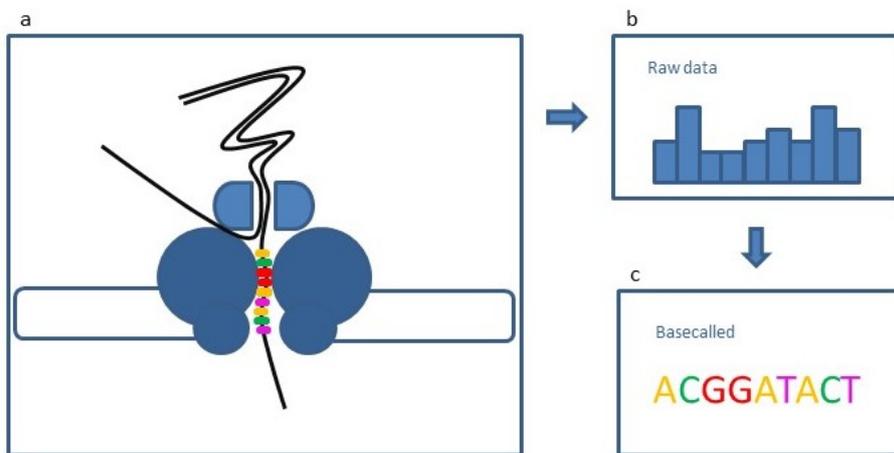


Fig 8, Schematic overview of the process of nanopore sequencing **a**) visualisation of a DNA-strand travelling through nano-pore, covering to **b**) Raw data and to **c**) a base-sequence after being basecalled

The Basecallers used by Dr. Florian Kraft were Guppy, Scrappy and Canu (in combination with Guppy) (all developed by ONT), the Alignment was performed via minimap 2 and the Analysis was executed with NanoSatellite.

2.2.8.2 Molecular Combing

Blood samples from 6 individuals of family 1 and 2 were analysed using molecular combing. The experiments were performed by the Cologne Center for Genomics (CCG), University Cologne, Cologne using the technology of Genomic Vision (Bagneux, France).

Molecular combing is a commercial method used to label specific regions of DNA with a fluorescent to point out variabilities to a referring pattern of markings. After preparation of the probands DNA, a special cover-slip is incubated into a solution with the probands DNA-mixture. Because of the pH and ionic strength conditions of the solution the DNA unwinds and the hydrophobic core of the DNA is exposed and attaches to the hydrophobic

surface of the cover-slip. One or two ends of a DNA-strand attach to the cover slip. As the slip is pulled from the solution the vertical force stretches the DNA in one direction, “combing” it into parallel strands. The DNA was immobilized in the surface by baking at 60 °C for 4 h.

In our study molecular combing was used to mark the flanking regions to the expansion in two different colours (5'-end in blue and 3'-end in green) and the TTTCA repeats in a third colour (red). TTTTA was not labelled with a colour and occurred as a gap in our regions of interest (ROI).

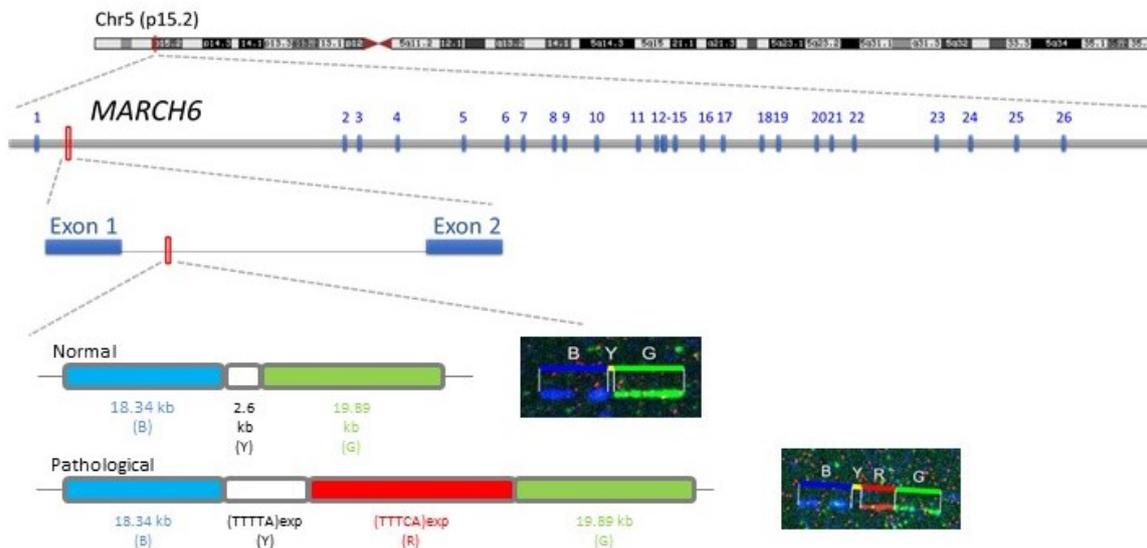


Fig. 9, Example of nanopore sequencing in intron 1 of *MARCH6*. The upper example hereby representative of a normal allele, the lower example representative of an expanded allele with a TTTCA-repeat insertion. The blue and the green sections mark the flanking regions of the region of interest, the unstained yellow part marks the TTTTA-repeats either in a “normal” size of about 2.6 kb and the red section marks the TTTCA-repeats if present.

In preparation the previously described primers (1.1.5.2 Molecular Combing) were ordered. The primers were flanking the region around the Expansion in an overlapping manner and were used to perform PCR’s on 7 different control-individuals. The PCR reaction, with a total reaction-volume of 25 µl, contained:

Component	Volume
HotstarTaq	12,5 µl
Forward Primer (20 pmol/µl)	0,5 µl
Reverse Primer (20 pmol/µl)	0,5 µl
H ₂ O	9,5 µl
DNA (25 ng/µl)	2 µl

The performed PCR-conditions were established as following by Prof. Dr. Christel Depienne and Sabine Kaya through repeated testing in various conditions.

- | | | | |
|---------|----------------------|-------------|----------------|
| 1.) | Initial Denaturation | 95°C | – 15 Min |
| 2.) x38 | Denaturation: | 95°C | – 30 Sec |
| | Primer-Annealing: | 56;60;64 °C | – 30 Sec |
| | Elongation: | 72°C | – 2 Min 30 Sec |
| 3.) | Final Elongation | 72°C | – 7 Min |
| | Cooling: | 4°C | – ∞ |

The PCR-products were cleaned with the DNA Clean & Concentrator-5 Kit (Zymo) and 2 µg of each of the seven PCR products were pooled and send to GV for labelling. These PCR-products corresponding to the region of interest were then used as templates for probe labelling by random priming by the CCG. DNA FISH probes were designed and labelled using a mix of red (TTTCA), green (3' flanking region) and blue (5' flanking region) probes. The Hybridization was carried out overnight and detected using fluorophore-coupled antibody-layers. This was done using the method developed by Genomic Vision (Bagneux, France).

2.2.9 Clinical/Neurological Testing

Twenty-one family members of family 2 were visited in France, resampled and examined neurologically in collaboration with Prof. Dr. Stephan Klebe, Institute of Neurology, University-Hospital Essen, University Duisburg-Essen, Essen.

The probands were asked about their medical history, age at onset for possible symptoms; medication and factors that had an influence on their tremor and/or Epilepsy.

As an assessment of the severity of the tremor, the Fahn-Tolosa-Marin Scale (FTM-Scale) for essential Tremor was used, as cortical tremor is very similar to essential tremor. The scale was developed by Stanley Fahn, Eduardo Tolosa and Concepcion Marin in 1988. To be able to rate the severity of the tremor, in this scale different components of the Phenotype, in 3 sections, were evaluated.

- Section 1: Quantification of the tremor at rest, in holding posture and with action and intention manoeuvres.
- Section 2: Assessment of action-tremor of the upper extremities with the use of Handwriting (only dominant hand), drawing an Archimede's spiral, drawing a straight line between narrow confines and pouring water from one glass to the other (without the glasses touching).
- Section 3: Quantification of functional disabilities, such as speaking, drinking, eating, hygiene, dressing, writing and working. The scores are provided by the subjective judgment of the proband.

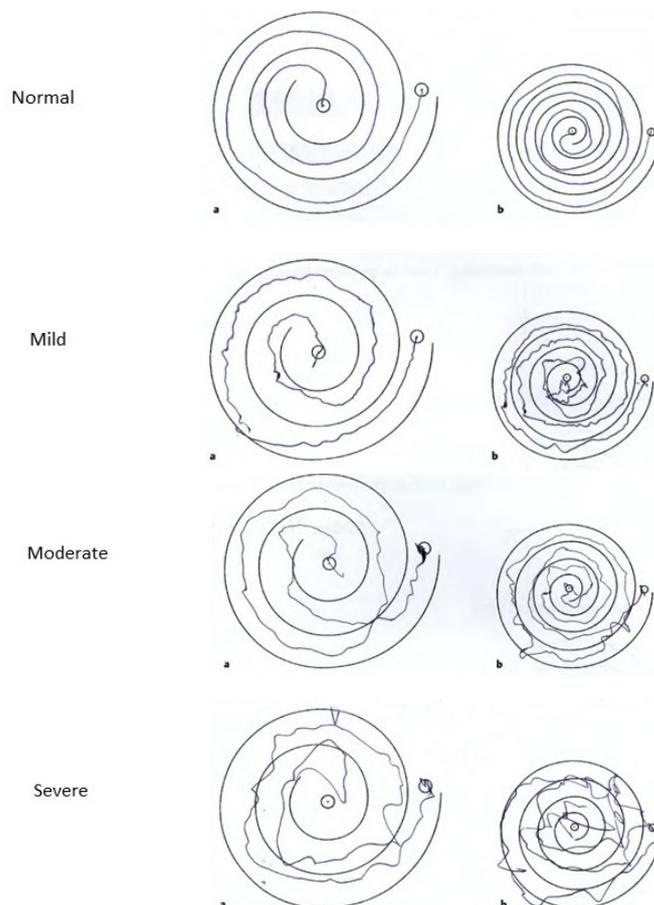


Fig. 10, Example of grading of drawing of an Archimedis spiral. In a drawing graded as normal, the line would not overstep the line of the archimedis spial, and the line itself would be without tremulous waves In a drawing classified as mildly pathological, the line would show slight tremulous waves, but overstep the lining of the archimedis spiral only slightly/rarely. In a drawing classified as moderately pathological, the line would show tremulous waves and cross the lining of the archimedis spiral regularly. In a drawing classified as severely pathological, the task of drawing within the archimedis spiral could only be implemented under great difficulty, resulting in lines that rarely stayed inside the spiral.

The definition of the scores was as following:

- Tremor
 - 0 = No tremor
 - 1 = slight tremor;
Amplitude < 0,5 cm
 - 2 = moderate tremor;
Amplitude 0,5 – 1 cm
 - 3 = marked tremor;
Amplitude 1 – 2 cm
 - 4 = severe tremor;
Amplitude > 2 cm

- Writing
 - 0 = Normal
 - 1 = mildly abnormal
 - 2 = moderately abnormal
 - 3 = markedly abnormal,
illegible
 - 4 = severely abnormal, unable
to keep the pencil on the paper
without holding the writing
hand down with the other hand

- Drawing (Figure 10):
 - 0 = Normal
 - 1 = May cross lines
occasionally
 - 2 = Crosses lines frequently
 - 3 = Many errors
 - 4 = Unable to complete the
drawing

- Pouring:
 - 0 = Normal
 - 1 = More careful than normal,
nothing is spilled
 - 2 = Small amounts spilled (up
to 10%)
 - 3 = Spills >10 – 50 %
 - 4 = Spills most of the water

- Speaking:
 - 0 = Normal
 - 1 = Tremulous when nervous
 - 2 = Constantly mildly
tremulous
 - 3 = Constantly moderately
tremulous
 - 4 = Severely tremulous, some
words can't be understood

- Feeding:
 - 0 = Normal
 - 1 = Mild spilling
 - 2 = Moderate spilling
 - 3 = Unable to cut food; needs
two hands to feed
 - 4 = Needs help to feed

- Drinking/Directing liquids to the mouth
 - 0 = Normal
 - 1 = Able to use a spoon; can't fill it all the way
 - 2 = Unable to use spoon; can drink from a glass or cup
 - 3 = Needs two hands to drink from a glass or cup
 - 4 = Must use straw to drink
- Hygiene:
 - 0 = Normal
 - 1 = More careful than usual
 - 2 = Errors occur
 - 3 = Needs two hands for fine tasks
 - 4 = severely abnormal, needs help
- Dressing:
 - 0 = Normal
 - 1 = More careful than usual
 - 2 = Errors occur
 - 3 = Needs help with fine tasks
 - 4 = Needs full assistance
- Working
 - 0 = Normal
 - 1 = More careful than usual
 - 2 = Errors occur
 - 3 = Must change profession; very limited in household tasks
 - 4 = Not able to perform any profession; needs help in the household

(Jankovik and Tolosa 1988: 225 - 34)

The decision of a cut-off for a proband to be labelled affected of FAME was made in collaboration with Prof. Dr. Stephan Klebe, Institute of Neurology, University-Hospital Essen, University Duisburg-Essen, Essen. It was decided without the knowledge of the individual test results of each examined proband to counteract adulteration of the results. The cut-off was set at 12 points. Therefore, a proband with a score equal or higher than 12 was handled as affected from a tremor similar to essential tremor and/or cortical tremor as occurring in FAME. The higher the score is, the more severely affected is the proband.

3 Results

Three of the four families published in Florian et al. in 2019 (Florian et al. 2019) and three of the families published in Corbett et al. in 2019 (Corbett et al. 2019) were analysed in the Institute of Human Genetics, University-Hospital-Essen for this thesis. Family 4 included in this thesis was first tested by the group of Mark Corbett and Jezef Gecez, University of Adelaide, Australia.

3.1 FAME

3.1.1 Identification of the Expansion

Going back to the Illumina genome data generated in 2017, bioinformatic approaches were used to search for TTTCA/TTTTA repeat-expansions within the linked regions of our two index-families. Family 1 hereby previously linked to the FAME3 locus on chromosome 5 and family 5 linked to the FAME2 locus on chromosome 2.

This analysis revealed TTTCA-repeats mapping to intron 1 of *MARCHF6* for family 1 and intron 1 of *STARD7* for family 5. These regions usually consist of a variable number from up to 20 TTTTA-repeats in the control population (reference genome: *MARCHF6* (TTTTA)₁₂, *STARD7* (TTTTA)₁₁). TTTCA-repeats at these loci were observed in all tested affected family-members and absent in the tested healthy spouses as well as the reference-genome-data.

Visualization of the mapped reads suggested the following expansion-structure:

5'-(TTTTA)_{exp}(TTTCA)_{exp}-3'.

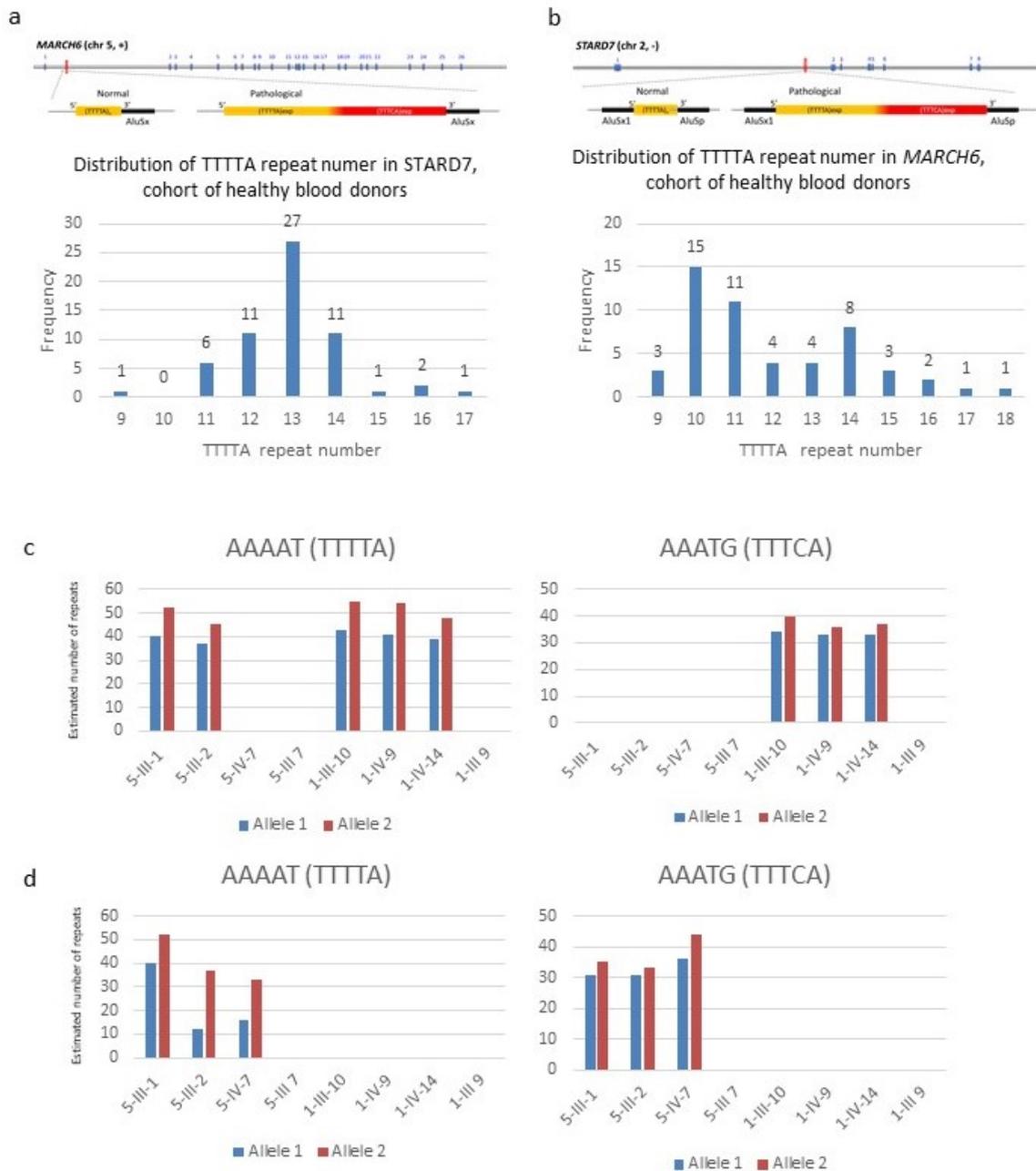


Fig. 11, Identification of TTTTA/TTTCA expansions in *MARCH6* and *STARD7*. **a+b**) Schematic representation of the region where the expansion occurs in intron 1 of *MARCH6* on chromosome 5 (**a**) and intron 1 of *STARD7* on chromosome 2 (**b**). The yellow rectangle indicates the TTTTA repeats while the red rectangle represents the TTTCA repeats. Below are the results of the distribution of TTTTA-repeat-numbers in this intron in the normal population. **c+d**) Number of TTTTA (actual repeated motif searched for: AAAAT) and TTTCA (actual repeated motif searched for: AAATG) repeats identified by ExpansionHunter from Illumina short-read genome data of three affected individuals (1-III-10, 1-IV-9, 1-IV-14) and one healthy spouse (1-III-9) of Family 1 and three affected individuals (5-III-1, 5-III-2, 5-IV-7) and one healthy spouse (5-III-7) of Family 5 in *MARCH6* (**c**) and *STARD7* (**d**). Dark and light bars indicate allele 1 and allele 2, respectively.

3.1.1.1 Amplification of the region of interest

To confirm the presence of pathogenic expansions in the affected family members of family 1 and family 5 we first amplified the region of interest by conventional PCR with a size-standard of 1200 bp in length.

Using this quantification-method we obtained results suggesting affected family members were homozygous with different alleles showing apparent nonmendelian inheritance. This suggested that the expanded allele had failed to amplify using conventional PCR due to its size. (Supplementary Table 1)

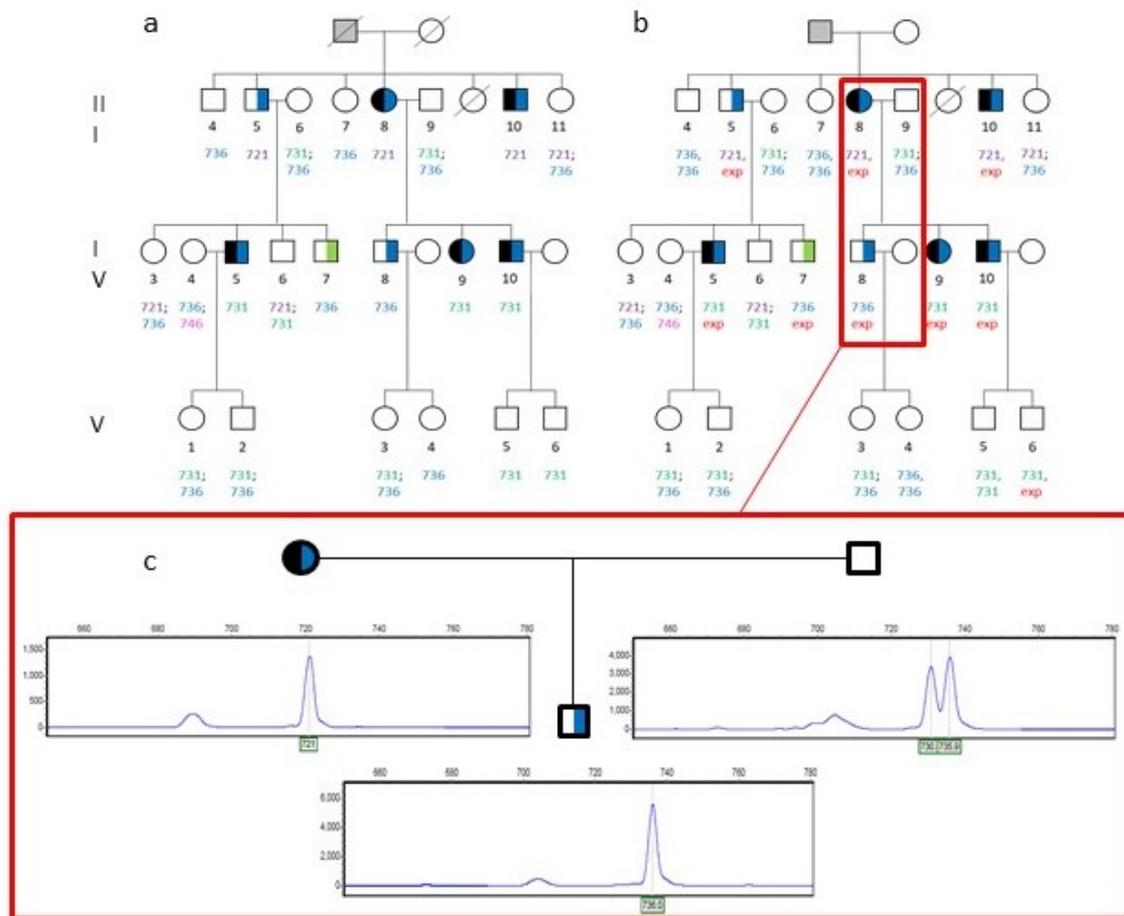


Fig. 12, Conventional PCR of the region of interest, Example of parts of Family 1 with amplification of *MARCF6*. **a+c)** Apparent non-mendelian inheritance observed from generation III (1-III-8 and 1-III-9) to generation IV (1-IV-8) with the affected mother 1-III-8 apparently being homozygous with allele 721 and her affected son being homozygous with allele 736. **b)** assuming the mother would inherit an expanded allele that could not be amplified by conventional PCR the inheritance would be mendelian as shown in **b**.

3.1.1.2 Confirmation using RP-PCR

To further confirm the presence of the expanded alleles, we set up 5'- and 3'-repeat-primed PCR essays using reverse and forward primers directly binding within the expansion (Fig 6). These assays confirmed the presence of 5'-TTTTA and 3'-TTTCA expanded motifs in all 16 affected family members and one unaffected family member (individual 1-V-6) from family 1 in intron 1 of *MARCHF6* and in all 10 affected family members and two unaffected members (individual 5-IV-2, 5-IV-6) from family 5 in intron 1 of *STARD7*.

We further used the RP-PCR to screen the remaining 10 families for expansions in *MARCHF6* or *STARD7*. This testing revealed 3 more families positive for an expansion in *MARCHF6* (Family 2, 3 and 4. Family 4 was tested by our collaborators Mark Corbett and Josef Gez University of Adelaide, Australia). Two additional families tested positive for an expansion in *STARD7* (Family 6 and 7). (Supplementary Fig. 2)

The expansion co-segregated with the disorder in which 59/60 of the affected individuals of all 7 families tested positive for an expansion. Individual IV-4 of family 6 was hereby reportedly affected with tremor uncertain in connection to a FAME phenotype but tested negative for the expansion.

The three individuals with questionable incomplete penetrance from family 1 and 5 were from younger generations, likely to be pre-symptomatic when first sampled. Individual 5-IV-2 was hereby tested positive for a (TTTTA)_{exp} pattern on either the 3'- and the 5'-End of the *STARD7* locus, suggesting that this late onset individual may had a different expansion-pattern on site. All other unaffected members from all families (59/59) were tested negative for any expansion. (Supplementary Table 2)

Families 8 to 12, that previously had not been linked to FAME2 or FAME3 showed no expansion in either locus.

3.1.1.2.1 Somatic mosaicism detected by RP-PCR

Individual 1-III-5 of family 1 as well as individual 5-III-1 and 5-III-3 of family 5 and five members of family 6 (individual 6-IV1, 6-IV-2, 6-IV-3, 6-V-1, 6-V-2) repeatedly tested positive for both TTTTA and TTTCA repeat expansions in the RP-PCRs using the reverse

primer FAME3-P1-R. This indicated that the expansion-structure in these individuals varied between: 5'-(TTTTA)_{exp}(TTTCA)_{exp}-3' and a different pattern putting the (TTTTA)_{exp} at the 3'-end of the structure.

To investigate this phenomenon, we sequenced intron 1 of *MARCHF6* in individual III-5 of family 1, knowing the sequenced allele would represent the “normal” allele to rule out a possibly elongated “normal” allele that corresponds to the allele amplified with a TTTTA-expansion. The analysis showed a TTTTA-repeat number of 12 repeats, not matching the number of TTTTA peaks present in the repeat-primed PCR. This confirmed that the TTTTA-expanded allele amplified through RP-PCR did not represent a slightly elongated “normal” allele and indicated that the RP-PCR-results may indeed reflect somatic mosaicism in individual 1-III-5 of family 1.

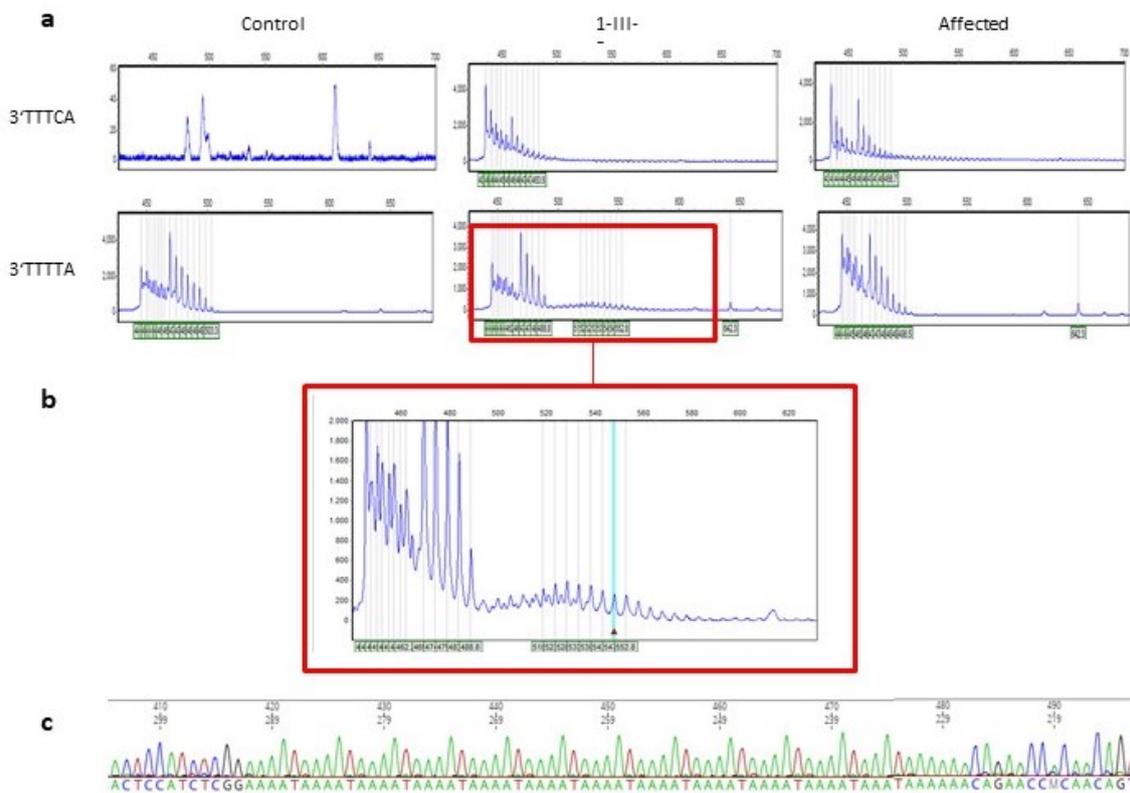


Fig 13, Somatic mosaicism in individual 1-III-5. **a)** Representative RP-PCR results of a control-individual on the left side showing non-expanded 3'TTTTA-repeats and no 3'TTTCA-repeats. Representative PCR-results of an affected individual with 3'TTTCA-expansion and non-expanded 3'TTTTA-repeats representing the normal allele. RP-PCR-results of individual 1-III-5 in the middle showing both 3'TTTCA and 3'TTTTA-repeat-expansions (enlarged in **b**). **c)** Sanger sequencing results of individual 1-III-5 showing a normal allele with 12 TTTTA-repeats, not matching the RP-PCR results.

3.1.2 Characterization of the expansion

To further characterize the expansions, we clinical reassessed members of Families 1 and 2 as well as resampling of four individuals from Family 1 and six individuals from Family 2.

To do so, I revisited Family 2 in collaboration with Prof. Dr. Stephan Klebe and examined and resampled all available family members. Furthermore, regarding Family 1, we received clinical data and new blood-samples from our collaborator Dr. Eloi Magnin (Hôpital de Besancon, France).

Since short-read based technology did not permit an accurate assessment of repeat number exceeding the corresponding read length, we decided to use long-read based technology to further characterize the length and structure of the expansions. For this purpose, we sent some of the newly sampled blood to our collaborator Dr. Florian Kraft from the Institute of Human Genetics, University-hospital Aachen for sequencing with Oxford Nanopore sequencing. The other samples were sent to the Cologne Center for Genomics (CCG), University Cologne, for performing molecular combing experiments. The detection and preparation of the regions of interest was hereby executed by us in the Institute of Human Genetics, University-hospital Essen.

3.1.2.1 Clinical data

3.1.2.1.1 Family 2

The clinical evaluation was based on an anamnesis regarding age at onset with symptom at onset, subjective symptoms currently observed, worsening factors, previous clinical evaluations and medication, as well as an objective neurological evaluation using the Fahn-Tolosa-Marin-Scale (Table 5). The clinical evaluation was performed, not knowing the previous clinical status or potential presence of the expansion of the evaluated individuals. The cut-off for the Fahn-Tolosa-Marin-Scale was picked at 12 points.

In comparison to previously performed evaluations, we could detect 2 individuals from younger generations that had still been pre-symptomatic when last seen, now showing signs of (myoclonic) tremor despite still being young (Individual 2-V-7 being 39 years old, individual 2-V-2 30 years old at examination). During evaluation individual 2-V-2 did not subjectively describe any tremor and had never been medicated but neurological examination revealed a moderate myoclonic tremor of the upper limbs with discrete signs in writing, drawing and pouring-tasks. Individual 2-V-7 had already been diagnosed with (myoclonic) tremor after being first evaluated and was medicated with beta blockers, but still showed a mild tremor and tremolos difficulties of the upper limbs during writing, drawing and pouring tasks and described slight restrictions in her daily life. Both individuals had expansions in *MARCF6* using RP-PCR. Six other individuals (Individual 2-IV-2, 2-IV-8, 2-IV-15, 2-IV-17, 2-V-1, 2-V-9) who had previously been described as FAME-affected, tested positive again during our clinical trial. All of them showed expansions in *MARCF6*, too.

3.1.2.1.2 Family 1

Regarding clinical evaluation (Table 6), the fourteen previously reported symptomatic individuals from family 1 tested positive for an expansion in *MARCF6*. Furthermore, individual 1-IV-7 and individual 1-V-6 who were previously reported asymptomatic harboured the expansion. At re-examination individual 1-IV-7 reported walking-difficulties possibly due to myoclonic tremor affecting the lower limbs and a single initially focal, evolving to bilateral convulsive seizure at the age of 46 years. Under

treatment with VPA he showed no further seizures. Neurological examination revealed a mild myoclonic asymmetrical tremor affecting the upper left extremity and the right lower extremity. Individual 1-V-6 was not available to re-examination, but belonged to the youngest generation of family 1, thus likely being pre-symptomatic when examined and sampled eleven years ago.

Patient-ID	Expansion in <i>MARCHF6</i>	Clinically Affected	Gender	Age at examination	Age at tremor onset	Age at seizure onset	Fahn Tolosa Marin Skala	Treatment
2-IV 1	-	-	f	71			2	
2-IV 2	+	+	m	71	19	19	23	VPA
2-IV 4	-	-	m	69			5	
2-IV 5	-	-	m	66			6	
2-IV 7	-	-	f	65			5	
2-IV 8	+	+	m	60	40	21	37	VPA; BZD
2-IV 10	-	-	m	58			3	
2-IV 11	-	-	m	58			3	
2-IV 12	-	tremor at higher age	m	75	NA		15	
2-IV 13	-	tremor at higher age	m	75	71		12	CBZ
2-IV 14	-	-	m	73			9	Keppra
2-IV 15	+	+	m	71	14	30	20	VPA
2-IV 17	+	+	f	67		30	17	VPA
2-V 1	+	+	m	42	12	12	15	VPA, Keppra, BZD, LAM
2-V 2	+	+	m	30	NA		15	
2-V 4	-	-	f	22			0	
2-V 5	-	-	m	49			8	
2-V 6	-	-	w	41			0	
2-V 7	+	+	f	39	14		13	BB
2-V 8	-	-	w	45			5	
2-V 9	+	+	m	44		30	9	VPA

Table 5. Summarized clinical data for family 2.

Ages are expressed in years. NA Not available, F female, M male, VPA sodium valproate, BZD Benzodiazepine, LAM Lamotrigin, LVT Levetiracetat, CBZ Carbamazepine, BB Betablocker

Patient-ID	Expansion in <i>MARCHF6</i>	Clinically affected	Gender	Age at examination	Age at tremor onset	Age at epilepsy onset	Treatment
1-III 1	+	+	F	60	41	41	VPA, BZD
1-III 5	+	+	M	72	25	-	VPA
1-III 8	+	+	F	70	35	-	VPA
1-III 10	+	+	M	63	30	30	VPA
1-III 14	+	+	M	67	35	25	VPA, BZD
1-III-16	+	+	F	65	35	30	VPA
1-III-17	+	+	F	66	25	25	VPA, BZD
1-III-20	+	+	F	54	47	26	VPA
1-III-22	+	+	M	48	40	30	VPA
1-IV-6	+	+	M	47	30	25	VPA, BZD
1-IV-7	+	+	M	53	52	46	VPA
1-IV-8	+	+	M	50	25	-	VPA, LVT
1-IV-9	+	+	F	49	30	32	VPA
1-IV-10	+	+	M	46	30	-	BZD
1-IV-12	+	+	M	44	30	-	CBZ
1-IV-16	+	+	M	35	25	25	VPA, LAM
1-V-6	+	-/NA	M				

Table 6. Summarized clinical data for family 1.

Ages are expressed in years. NA Not available, F female, M male, VPA sodium valproate, BZD Benzodiazepine, LAM Lamotrigine, LVT Levetiracetam, CBZ Carbamazepine, BB Betablocker

3.1.2.2 Long-read based technology

3.1.2.2.1 Oxford Nanopore sequencing

Long-read based Oxford-Nanopore sequencing at low (5-10X) coverage allowed us to retrieve one to four reads displaying the expansion per individual. We observed a substantial variability in reads covering the expansions, not only between individuals, but also for the same individual. Expansions ranged from a mean of 4-6.5 kb and between 791 and 1035 repeats in five of the six individuals. The reads sequenced in individual IV-5 of family 1 failed to cover the whole expansion with two of four reads already spanning a variable TTTCA-expansion-part of up to 5 kb. (Supplementary Fig. 3)

We observed multiple insertions and deletions of single nucleotides, sometimes resulting in new pentanucleotide patterns interrupting the expected $(TTTTA)_{exp}-(TTTCA)_{exp}$ -pattern, which were not consistent from read to read within the reads. These differences could be real or correspond to errors of the nanopore sequencing method.

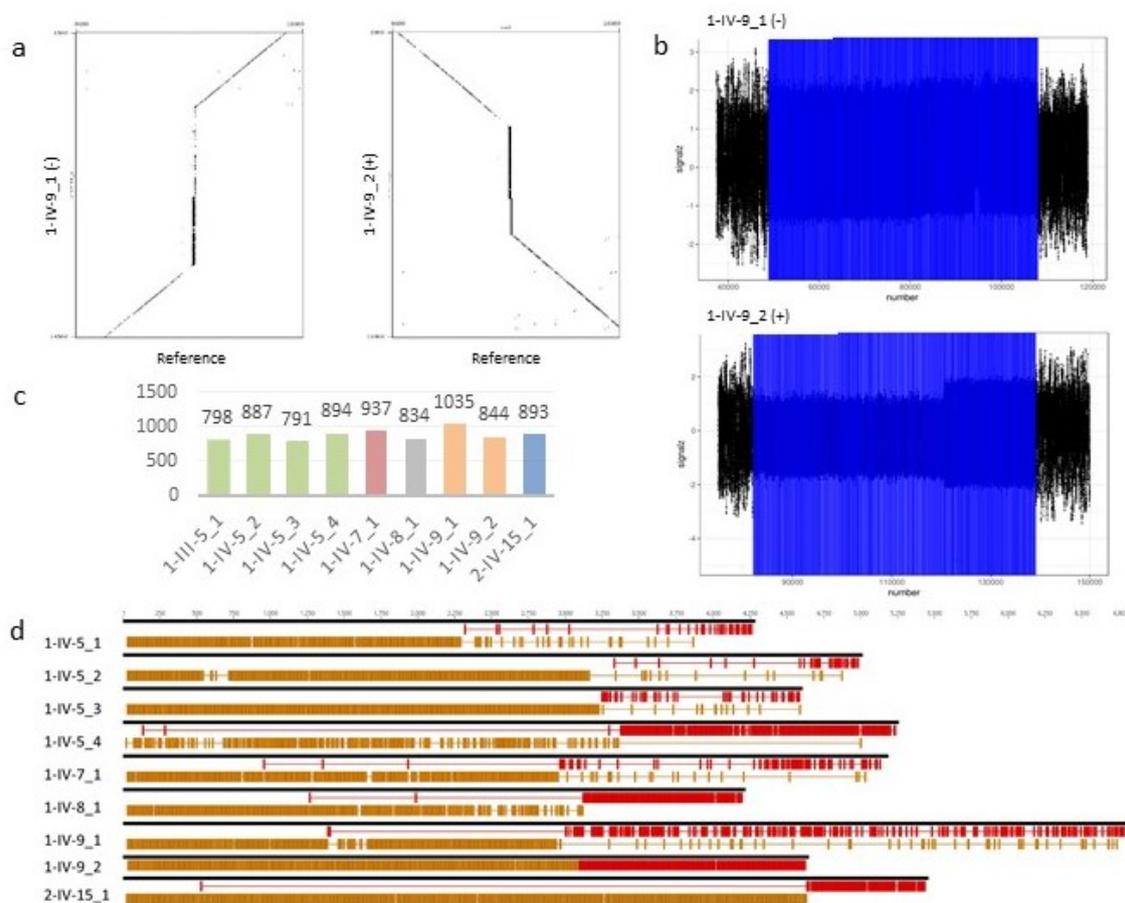


Fig. 14. Characterization of *MARCHF6* expansions by Nanopore sequencing. **a** Dot plots comparing two nanopore reads from individual 1-IV-9 displaying the expansion (Y-axis, scale: 13 kb) with the corresponding hg19 reference region (X-axis, scale: 8.1 kb). The expansions appear as vertical lines. Read 1-IV-9_1 is on the negative strand while read 1-IV-9_2 is on the positive strand. **b** Analysis of the same raw nanopore reads using NanoSatellite. The signals corresponding to the expanded repeats appear in blue. **c** Number of total repeats inferred by NanoSatellite for each extracted read covering the expansion. Data are displayed for the five individuals for whom reads covering the whole expansion have been detected. Four reads covering parts of the expansion and flanking regions were obtained for individual 2-IV-5 but are not included in this graph. **d** Schematic representation of the sequence of the same nanopore reads showing exact TTTTA motifs in yellow and exact TTTC A motifs in red. Gaps between exact repeats possibly correspond to interruptions or sequencing (base calling) errors.

3.1.2.2.2 Molecular combing

Our goal using molecular combing was to get a further idea about expansion sizes and to verify the existence of somatic mosaicism in the expanded alleles of FAME-affected individuals indicated by Nanopore sequencing and the analysis of individual III-5 of family 1. Molecular combing enabled us to stain long, stretched DNA-fibres by in situ hybridization (Fig. 8) and measure the length of every signal for every allele. Each coverslip for each individual hereby containing about 100 stained alleles.

The analysis of the data revealed a large variability of the length of expansion between individuals and within each individual, confirming the presence of somatic mosaicism. We did not observe any difference in mean length or configuration between blood-cells and fibroblasts. The overall expansion-size hereby ranging on average from 3.34 to 14.07 kb. (Table 7, Supplementary Fig. 4)

Conversely, we observed that within the individuals with the smallest expansions, the percentage of pathogenic alleles was lower than expected by chance, which suggests a higher overlap between normal and pathogenic alleles in these individuals and a possible underestimation of the size of the repeat expansion in molecular combing (4.1.4.2 characterization of the expansion).

3.1.2.2.3 Micro rearrangements in large expansions

Furthermore, as indicated in RP-PCR, some individuals tested with molecular combing showed recurrent staining patterns compatible with expansion configurations other than 5'-(TTTTA)_{exp}(TTTCA)_{exp}-3', some associated with complex micro-rearrangements at the expansion-sites. Especially the individuals with the largest expansion (individual 2-IV-8 with a mean expansion length of 14.07 kb and individual 2-V-3 with a mean expansion length of 13.3 kb) showed micro-rearrangements in up to 10% of the alleles present on the coverslip, indicating that 20% of expanded alleles showed rearrangements in these individuals. (Supplementary Fig. 5)

Going back to the nanopore reads covering the expansion of individual 2-IV-8, we could find one read covering the expansion of individual 2-IV-8 even showed the 3' flanking region and TTTCA part of the expansion on chromosome 5p15.2 fused to a region on chromosome Xp22.3. We concluded that this read corresponded to a micro-rearrangement involving another chromosome.

These findings indicate that the size of the expansion positively correlates with the degree of somatic instability thus more often resulting in micro-rearrangements.

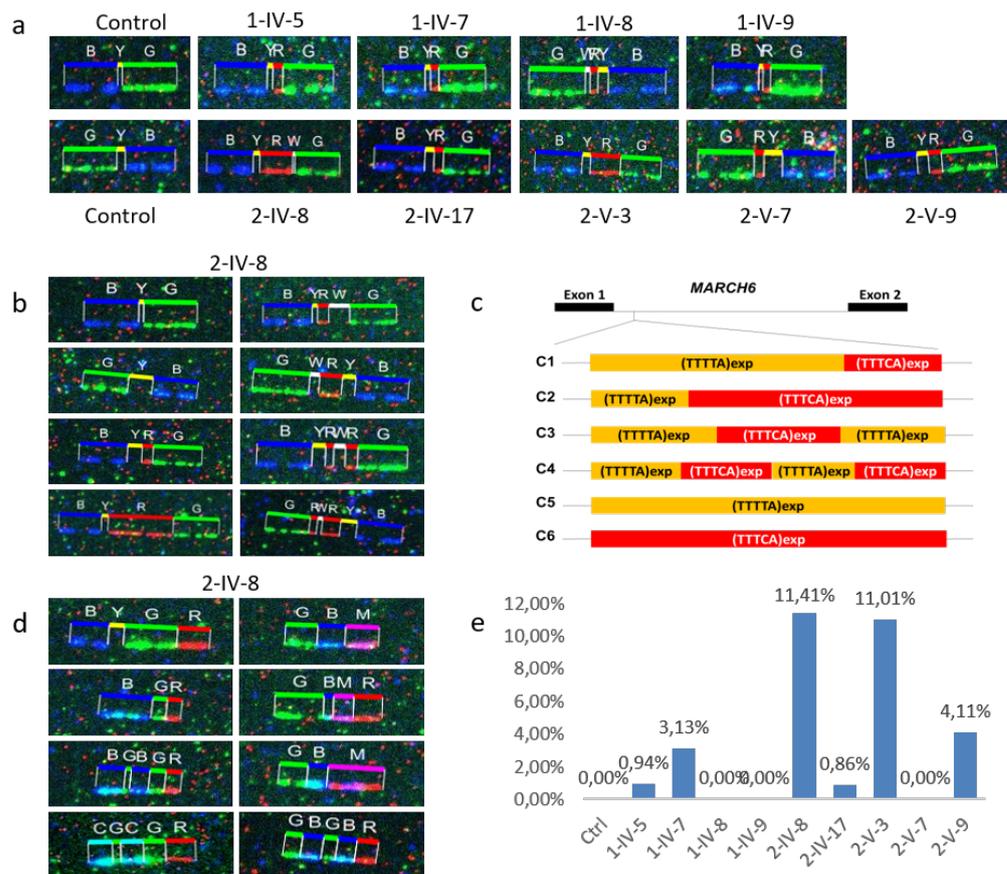


Fig. 15, Somatic mosaicism of *MARCHF6* expansions detected by molecular combing. **a)** Representative images seen in a control individual (two panels on the left) and in nine expansion carrier individuals for whom molecular combing was performed. Y refers to the unstained part between the blue and red signals; unstained parts detected between the red and green signals or in-between two red signals are referred to as W. **b)** Selected images observed at the expanded site in the proband of Family 2 (2-IV-8), showing extreme variability of the expansion length and structure in his blood. **c)** Schematic representation of the different expansion configurations (C1–C6) observed using molecular combing. **d)** Selected micro-rearrangements observed at the expanded site in individual 2-IV-8. M (magenta) and C (Cyan) correspond to the overlay of red and blue or green and blue probes, respectively, indicating an overlap of probes that should normally be separated. All images corresponding to micro-rearrangements observed in individuals 2-IV-8 and 2-V-3 are shown in Supplementary Fig. 3. **e)** Percentage of micro-rearrangements observed in the ten individuals analyzed by molecular combing. Individuals with the largest expansions (2-IV-8 and 2-V-3) exhibit a higher percentage of rearranged alleles than individuals with smaller expansions.

3.1.2.3 Genotype-phenotype correlation

Based on molecular combing data obtained from blood, we further explored the relationships between the repeat size and clinical features.

We could not detect a significant correlation between the age at tremor onset and either size of the expansion-components. On the contrary, we observed an inverse correlation between the age at seizure onset and the number of TTTCA-repeats as well as the overall expansion-size (including TTTTA and TTTCA repeats). The TTTTA expansion size alone did not significantly correlate with the age at seizure onset.

Accordingly, the two individuals with the largest expansions (individual 2-IV-8 and 2-V-3) were amongst the most severely affected individuals. Both started to have generalized seizures at 17-18 years of age. During our re-evaluation, individual 2-IV-8 at 60 years of age showed a moderate, asymmetric myoclonic tremor affecting the upper extremities (right side more affected than left side) and unspecific gait-difficulties despite being treated with VPA and Clobazam (Benzodiazepine). His 28 years old son (individual 2-V-3) had an autism spectrum disorder (ASD) and an intellectual disability (ID) in addition to FAME and lived in an institution for disabled people. Therefore, he was not available for re-evaluation. Analysis of trio exome had failed to reveal any other pathogenic variant in individual 2-V-3 and it remained unclear whether his ASD/ID phenotype was related to FAME or not. Furthermore, the two individuals without epilepsy (individuals 1-IV-8 and 2-V-7) had the smallest expansions.

Conversely, individual 1-IV-7 who had a late onset tremor and a single seizure at the age of 46 years (Clinical data) had an expansion of about 5 kb (5.13 calculated with Oxford nanopore and 5.06kb by molecular combing) and a TTTCA length (2.21 kb) comparable to those of close relatives (1-IV-5, 1-IV-8, 1-IV-9), who were all earlier and more severely affected (clinical data) raising further questions about the influence of other factors besides the expansion itself or about the ability to accurately predict the expansion sizes existing in the brain by analysis of peripheral tissues.

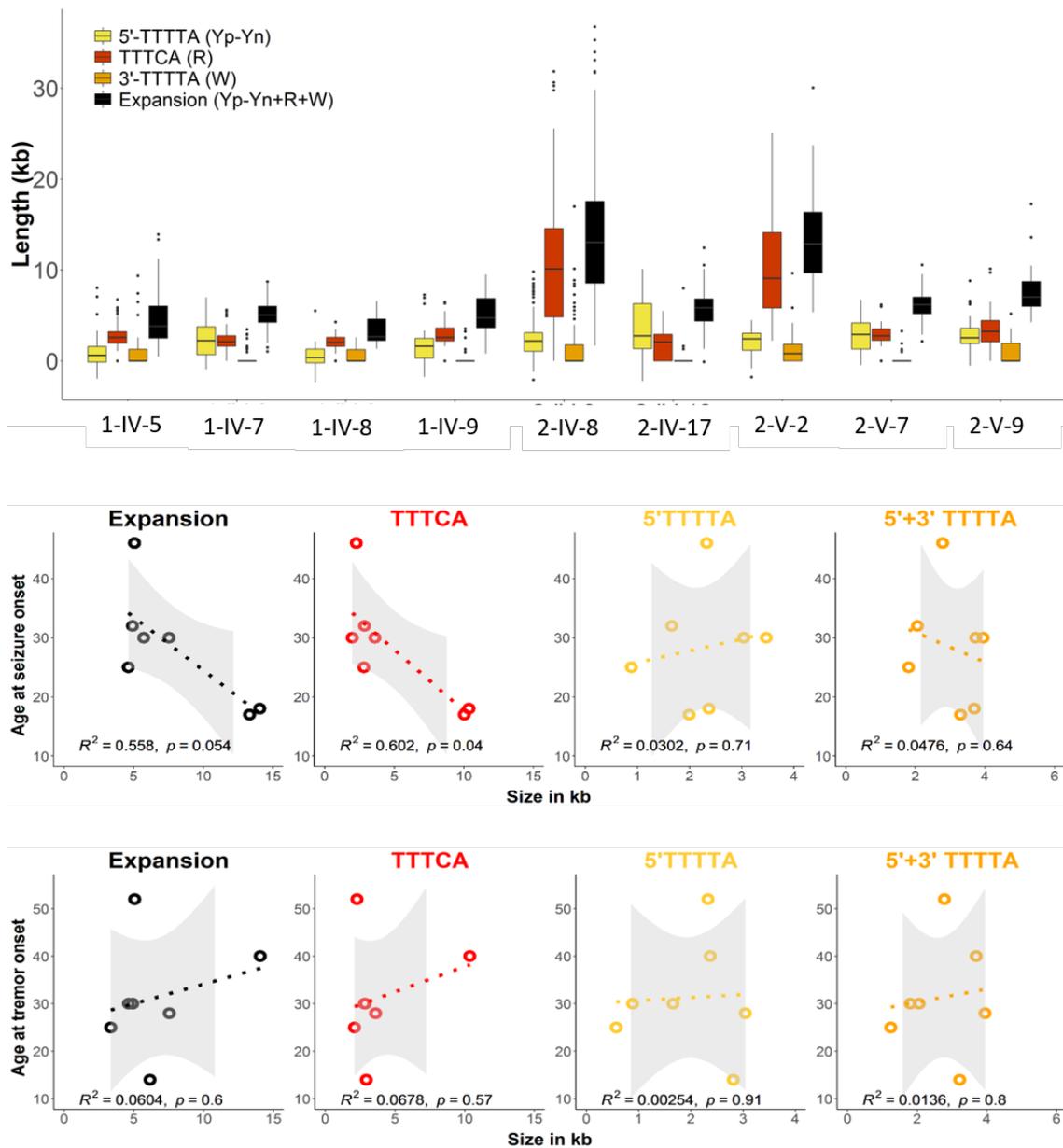


Fig. 16, Distribution of expansion lengths and genotype–phenotype correlations. **a** Box plots showing the distribution of the size of the overall expansion (in black), as well as the 5'-TTTTTA (yellow, Yp–Yn; see Methods for details) and TTTCA (red, R) parts in blood from the nine carrier individuals. Some alleles showed an unstained part between the red and the green signals, which is referred to as 3'-TTTTTA (W, in orange). Box plots elements are defined as follows: center line: median; box limits: upper and lower quartiles; whiskers: 1.5× interquartile range; points: outliers. **b** Correlations between the age at seizure onset and the mean size (in kb) of the overall expansion (left), the TTTCA (middle left), the 5'-TTTTTA (middle right), or the overall (5' + 3') TTTTA repeats region (right). Individuals with larger TTTCA repeat region have an earlier age at seizure onset. On the contrary, neither the size of 5'-TTTTTA or 5' + 3'-TTTTTA repeats correlate with the age at epilepsy onset. Individuals included in the graph are 1-IV-5, 1-IV-7, 1-IV-9, 2-IV-8, 2-IV-17, 2-V-3, and 2-V-9. Individuals without epilepsy also have the smallest TTTCA stretches although they are not included. R^2 is the square value of the Pearson coefficient; 95% confidence intervals appear in gray; Correlations between the age at tremor onset and the mean size (in kb) of the expansion and each part, showing no correlation with any of them. Individuals included in the graph are 1-IV-5, 1-IV-7, 1-IV-8, 1-IV-9, 2-IV-8, 2-V-7, and 2-V-9

ID	1-IV-5	1-IV-7	1-IV-8	1-IV-9	2-IV-8	2-IV-15	2-IV-17	2-V-3	2-V-7	2-V-9
<i>Clinical Data</i>										
Age at last examination	58	53	61	60	60	71	67	28	39	44
Age at tremor onset	30	52	25	30	40	14	-	NA	14	28
Age at seizure onset	25	46	-	32	18	30	30	17	-	30
<i>Nanopore sequencing</i>										
No. of P alleles	4	1	1	2	4	1	ND	ND	ND	ND
Mean expansion size	4.73	5.13	4.16	5.67	NA	5.40				
Mean 5'-TTTTA size	2.95	2.92	3.08	3.00	NA	4.60				
Mean TTTCA size	1.78	2.21	1.08	2.67	>5	0.80				
<i>Molecular combing</i>										
No. of P alleles	71	58	25	29	219	ND	50	54	30	38
Mean expansion size	4.62	5.06	3.34	4.92	14.07		5.72	13.33	6.16	7.55
Mean 5'-TTTTA size	0.88	2.33	0.57	1.66	2.37		3.47	1.99	2.81	3.04
Mean TTTCA size	2.82	2.28	2.10	2.86	10.37		1.99	10.04	2.93	3.60
Mean 3'-TTTTA size	0.92	0.46	0.67	0.40	1.32		0.27	1.31	0.41	0.90

Table 7. Summarized clinical features and expansion characteristics for the 10 resampled individuals. Ages are expressed in years and expansion sizes are in kb; *NA* not available; *ND* not done

3.2 Essential Tremor

Since ET is much more prevalent than FAME and is clinically very close, it represents the most frequent differential diagnosis to FAME. In theory, it is therefore likely that some patients diagnosed with ET are in truth affected with FAME.

During this study we investigated whether some ET patients could have an expansion at one of the two FAME loci identified during this study.

3.2.1 Amplification of the region of interest

To investigate the presence of pathogenic expansions in the affected members of the families of our ET cohort, we first amplified the region of interest (intron 1 of *STARD7* and intron 1 of *MARCHF6*) by conventional PCR in order to search for apparent non-mendelian inheritance as observed in expansion-carriers in our FAME families.

We observed apparent non-mendelian inheritance in only one of our families (Family 24) in *STARD7*.

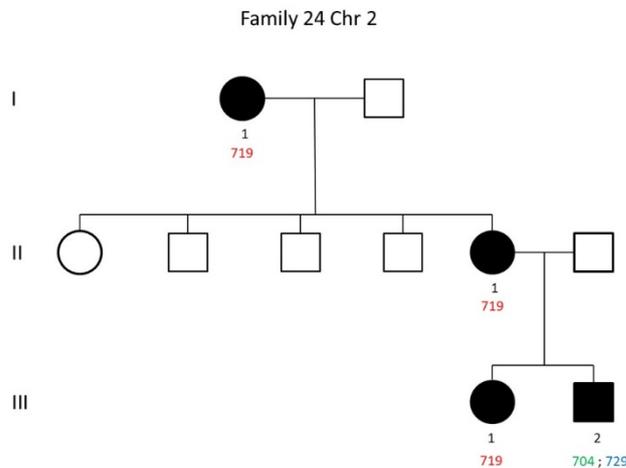


Fig 17. Apparent non-mendelian inheritance observed in family 24, where if individual 24-II-1 is homozygous for allele 719, individual 24-III-2 would need to carry the allele 719, too.

Nevertheless, we obtained some different, unexpected results, showing that in some individuals one allele was far less efficiently quantified than the other. We first thought that these results were due to technical difficulties during PCR, but even after optimizing the PCR-procedures, the results were reproducible. Furthermore, in some families, one allele seemed to be relatively frequent, possibly segregating with the ET phenotype.

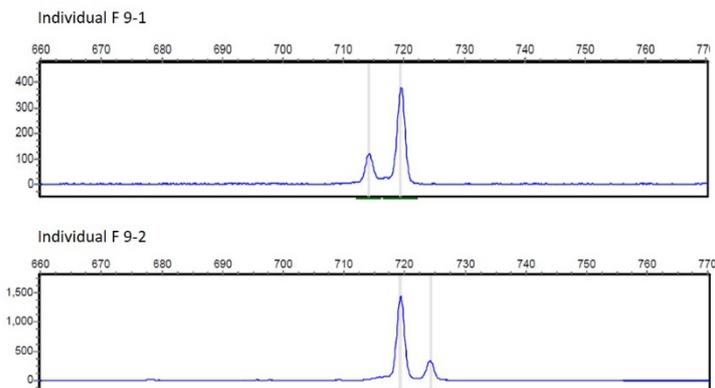


Fig. 18, Example for less efficiently quantified second allele in two individuals of Family 9.

3.2.1.1 RP-PCR

To further rule out the existence of expansions at the two FAME loci, we amplified the regions using RP-PCR targeting a TTTCA expansion. We decided to amplify *MARCHF6* in family 36 exemplary, because we observed the segregation of the allele 704 in this family. Unfortunately, both family-members did not show to carry a TTTCA expansion in *MARCHF6*.

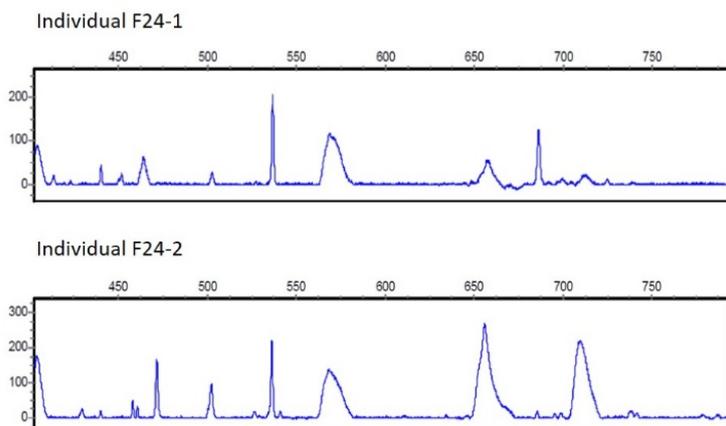


Fig. 19, Negative results for RP-PCR targeting TTTCA-repeats in *MARCHF6* on Chromosome 2 for individual 24-1 and 24-2 of family 24.

3.2.2 Visualization through Sanger-Sequencing

To go further, we wanted to visualize the region of interest to determine whether the differences in amplification-efficiency during PCR were due to technical difficulties, or if there was another explanation. To do so, we decided to test single individuals from families that had alleles that (mostly) segregated with the ET phenotype with Sanger sequencing.

Testing *STARD7*, we found that in many of the individuals who had one allele that was less sufficiently quantified in PCR, we could only quantify one allele during sanger-sequencing with a TR-motif other than TTTTA. Overall, 9 of the 22 families tested with sanger sequencing had family members with alternative motifs (about 40% of the tested families or around 31% of all families). The most frequently observed motif was (TGTTA)₈₋₁₆ in 6 families, but we also observed (TTTTG)₁₄ and (TTTGA)₁₉. The other individuals showed numbers of TTTTA-repeats that matched their PCR results. To investigate if the alternative motif segregated with the disease, we sequenced the previously untested family-members in 7 families (Family 13, 17, 20, 21, 22, 24, 27, 38). We found that the allele segregated with the disease in two families (family 17 and 27), but all other families had affected members that did not carry an alternative TR-motif in *STARD7*, questioning the connection to the ET-phenotype.

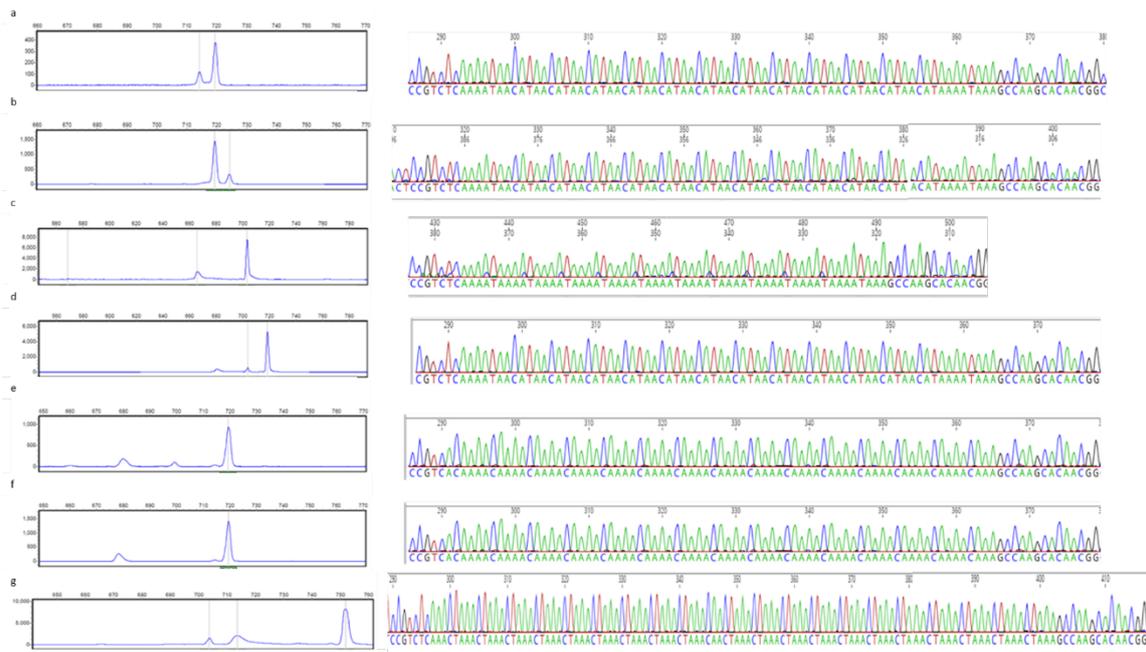


Fig 20. Examples of individuals with one less efficiently quantified allele, that carry an alternative TR-motif.

a) Individual 21-II-1 with TGTTA, **b)** Individual 21-II-2 with TGTTA, **c)** Individual 21-III-1 with TTTTA underlying with a small c every 5 bases (was not considered pathogenic), **d)** Individual 21-II-1 with TGTTA, **e)** Individual 24-I-1 with TTTTG, **f)** Individual 24-III-1- with TTTTG, **g)** Individual 38-III-1 with TTTGA

Indiviudal ID	Alleles	Sanger Sequencing
13-1	(663); 704	(TTTTA) ₁ (TTTA) ₁ (TGTTA) ₈ (TTTTA) ₁
13-2	704	(TTTTA) ₁₁
13-3	704	(TTTTA) ₁ (TTTA) ₁ (TGTTA) ₈ (TTTTA) ₁
13-4		(TTTTA) ₁₁
13-5	704	(TTTTA) ₁₁
14-III-1	704	(TTTTA) ₁₁
16-III-1	709	(TTTTA) ₁₂
17-II-1	719 ; (744)	(TTTTA) ₁ (TGTTA) ₁₂ (TTTTA) ₁
17-II-2	(704) ; 719	(TTTTA) ₁ (TGTTA) ₁₂ (TTTTA) ₁
17-III-1	(704) ; 719	(TTTTA) ₁ (TGTTA) ₁₂ (TTTTA) ₁
18-IV-1	709	(TTTTA) ₁₂
20-II-1	(704); 719	(TTTTA) ₁ (TGTTA) ₁₂ (TTTTA) ₁
20-III-1	(666); 703	(TTTTA) ₁₁
21-II-1	(714) ; 719	(TTTTA) ₁ (TGTTA) ₁₂ (TTTTA) ₁
21-II-2	719 ; (724)	(TTTTA) ₁ (TGTTA) ₁₂ (TTTTA) ₁
21-II-3	704 ; 724	(TTTTA) ₁₀ ; (TTTTA) ₁₄
21-III-1	(704) ; 719	(TTTTA) ₁ (TGTTA) ₁₂ (TTTTA) ₁
21-III-2	704 ; 724	(TTTTA) ₁₀ ; (TTTTA) ₁₄
23-1	699 ; 724	(TTTTA) ₁₀ ; (TTTTA) ₁₄
23-3	(699) ; 724	(TTTTA) ₁ (TGTTA) ₁₄
24-I-1	719	(TTTTG) ₁₄
24-II-1	719	(TTTTG) ₁₄
24-III-1	719	(TTTTG) ₁₄
24-III-2	704 ; 729	(TTTTA) ₁₀ ; (TTTTA) ₁₅
25-III-1	709	(TTTTA) ₁₂
26-IV-1	703	(TTTTA) ₁₁
27-I-1		(TTTTA) ₁ (TGTTA) ₁₂ (TGTTTA) ₁ (TGTTA) ₄
27-II-1	(704) ; 740	(TTTTA) ₁ (TGTTA) ₁₂ (TGTTTA) ₁ (TGTTA) ₄
27-III-1	(704) ; 740	(TTTTA) ₁ (TGTTA) ₁₂ (TGTTTA) ₁ (TGTTA) ₄
28-II-1	703 ; 734	(TTTTA) ₁ (TGTTA) ₁₁ (TGTTT) ₁ (TGTTA) ₄
29-I-1	704	(TTTTA) ₁₁
29-II-2	703	(TTTTA) ₁₁
33-II-1	(699) ; 704	(TTTTA) ₉ ; (TTTTA) ₁₀
33-III-1	704	(TTTTA) ₁₁
33-III-2	704	(TTTTA) ₁₁

33-III-3	704	(TTTTA) ₁₀
33-IV-1	699 ; 704	(TTTTA) ₉ ; (TTTTA) ₁₀
34-IV-1	709	(TTTTA) ₁₂
35-II-1	716 ; 755	(TTTTA) ₁₁ ; (TTTTA) ₁₆
35-III-1	(662); 704	(TTTTA) ₁₁ ; (TTTTA)?
36-III-1	704	(TTTTA) ₁₁
38-II-1	704 ; 709	(TTTTA) ₁₁ ; (TTTTA) ₁₂
38-II-2	704	(TTTTA) ₁₁
38-III-1		(TTTGA) ₉ (TTTGTTGA) ₁ (TTTGA) ₁₀
39-II-2	698 ; 703	(TTTTA) ₁₀ ; (TTTTA) ₁₁
40-II-3	709	(TTTTA) ₁₂
42-II-6	709	(TTTTA) ₁₂

Table 8. Summary of all ET-individuals tested with sanger sequencing in *STARD7*; individuals with TGTTA motif in green, with TTTTG motif in blue and TTTGA motif in red; individuals that had a small c underlying the sequencing are marked in yellow, we did not account these individuals as alternative motif; individuals with TTTTA motif are not marked.

Furthermore, we did not find any alternating motifs in *MARCHF6*. Individuals who showed a smaller, less efficiently amplified allele in PCR were usually homozygous for the efficiently amplified TTTTA-allele, suggesting the smaller peak to be an artefact. (Supplementary table 4)

3.2.3 TA-cloning

Going further, we wanted to investigate the possible presence of a second, less efficiently quantified allele in individuals with an alternating motif in *STARD7* hypothesising a possible allele drop out during quantification. To do so, we decided to perform TA-cloning on individual 21-II-2. We managed to clone the TGTTA motif, as well as an allele with a TTTTA motif in 2/5 clones. These results therefore confirmed, that this individual was heterozygous and that the introduction of a G/C in the otherwise very TA-rich region results in an amplification bias in favour of the alternative allele. This bias explained the apparent non-mendelian inheritance previously observed in family 24. We concluded that individual 24-II-1 carried either the (TTTTA)₁₀ or the (TTTTA)₁₅ allele, but due to the presence of the TGTTA motif in the other allele, it was not sufficiently quantified during PCR, masking its presence both in the conventional PCR and in sanger sequencing.

4 Discussion

4.1 FAME

4.1.1 Expansions

In this study, we provide further evidence that the FAME-disease is homogeneously caused by TTTTA/TTTCA repeat expansions at polymorphic microsatellite sites in different genes on different chromosomes, originally composed of TTTTA repeats. We hereby identified two new sites of this expansion in intron 1 of *MARCHF6* in four FAME3 associated families and intron 1 of *STARD7* in three FAME2 associated families.

4.1.1.1 Expansion sites

Ishiura et al. first described TTTTA/TTTCA-expansions causative for FAME in 2018 and described 3 genes harbouring FAME expansions in the process of their studies. Our study provides two additional loci for FAME-causative expansions and during the duration of our study another expansion site in *YEATS2* on Chromosome 3 has been published by Yeetong et al. All expansions are therefore located in different genes on different chromosomes. Furthermore, the six genes harbouring FAME-related expansions strikingly have different functions and although some are specifically expressed in the central nervous system the expression of others is more ubiquitous.

Gene	Function	Expression
<i>SAMD12</i> on Chromosome 8 (Sterile alpha motif domain containing 12)	Unknown function; encoding three different spliced isoforms	Mainly expressed in cortex and cerebellum
<i>RAPGEF2</i> on Chromosome 4 (Rap guanine nucleotide exchange factor 2)	Encodes several protein isoforms from RAS subfamily of GTPases; act to switch on Ras and/or ERK signal pathways in response to the activation cell surface receptors, such as dopaminergic receptors	Primarily expressed in neurons, highest expression in the cortex

<i>TNRC6A</i> on Chromosome 16 (Trinucleotide repeat-containing gene 6A protein)	Encodes a component of a cytoplasmic ribonucleoprotein complex involved in regulating mRNA silencing, stability and translation	Ubiquitously expressed, highest expression in the cerebellum
<i>MARCHF6</i> on Chromosome 5	Encodes an E3 ubiquitin ligase that mediates the degradation of misfolded or damaged proteins in the endoplasmic reticulum	Ubiquitously expressed
<i>STARD7</i> on Chromosome 2	Encodes a ubiquitous protein expressed in lipid transport and metabolism	Ubiquitously expressed
<i>YEATS2</i> on Chromosome 3	Encodes a ubiquitous subunit of the ADA2A-containing histone acetyltransferase complex	Ubiquitously expressed; high expression in the cerebellum

Table 9. Summary of FAME related genes carrying a TTTCA-expansion with function and location of the expression (Ishiura et al. 2018; Corbett et al. 2019; Florian et al. 2019; Yeetong et al. 2019)

4.1.1.2 Rearrangements and instability

FAME expansions are somatically unstable, leading to a wide range of expansion sizes and configurations, even within each individual. Expansions ranged on average from 3.34 to 14.07 kb at the *MARCHF6* locus. We further showcase that the observed instability also results in genomic rearrangements in individuals with very large (> 10kb) expansions. Accordingly, these genomic rearrangements are associated with the more severely affected individuals and likely have a deleterious impact on the corresponding cells. It remains unclear, whether they directly or indirectly contribute to the pathophysiology of FAME. However, we observe that the more rearranged the expansion site is, the more unstable and prone to further mutation is the expansion itself.

4.1.2 Pathological mechanisms

Both our study and the publication of Ishiura et al in 2018 show a negatively inversed correlation between the size of the expansion and the age at epilepsy-onset. We now demonstrate that the correlation is mainly driven by the size of the TTTCA part of the expansion whereas the length of the TTTTA expansion alone does not correlate with the age at disease onset. This matches the observation, that rare expanded TTTTA alleles

containing no TTTCA repeats exist in 5.9% of healthy individuals in the Japanese population (Ishiura et al. 2018), as well as the observation that some FAME1-individuals that carry the TTTCA-expansion show normal/only slightly elongated TTTTA repeats (Cen et al. 2018).

This clear phenotypical connection to the TTTCA motive as well as the detection of FAME-causative expansions in apparently unrelated genes strongly suggests that the pathological mechanisms are not fully dependent from the gene itself or its function and are more likely related to the type of expansion itself.

Furthermore, all FAME-related expansions are located within gene introns, suggesting that transcription is a key step in the pathogenic process. Up to date, there are three hypotheses about how repeat expansions in non-coding regions cause neurotoxicity resulting in neurological symptoms. Either, the expanded RNA forms nuclear RNA-foci that prevent interacting proteins from their normal function, or the expanded RNA is translated to short peptides which are toxic, or the expansion results in a gain or loss of function of the associated gene (Cen et al. 2018).

4.1.2.1 Repeat-associated non-AUG (RAN)-initiated translation

Repeat-associated non-AUG (RAN)-initiated translation enables elongation through a repeat strand in the absence of an AUG initiation codon producing multiple repeat-containing proteins in multiple reading frames. It can be initiated in multiple DNA-regions, including within 5' untranslated regions, protein-coding regions, or introns and non-coding RNAs. The mechanism by which RAN translation occurs remain largely unknown (Green, Linsalata, and Todd 2016).

First described by Zu et al. in 2011 (Zu et al. 2011), RAN-translation has been associated with ten different neuropathological diseases associated with different microsatellite motifs, including tri-, tetra-, penta-, and hexanucleotide repeat expansions. (SCA8 (Zu et al. 2011), SCA31 (Ishiguro et al. 2017), HD (Banez-Coronel et al. 2015), DM1 (Zu et al. 2011), DM2 (Zu et al. 2017), ALS/FTD (Gendron et al. 2013), fragile X tremor/ataxia syndrome (FXTAS) (Krans, Kearse, and Todd 2016), fragile X-associated primary ovarian insufficiency (FXPOI) , (Buijsen et al. 2016)familial frontotemporal lobar

degeneration (FTLD) (Mori et al. 2013) and Fuchs' endothelial corneal dystrophy (FECD) (Soragni et al. 2018)).

Although disorders associated with RAN-translation have their own distinct clinical features and pathology, including individual RAN proteins, it is possible to distinguish common features in these diseases. For instance, RAN translation is often repeat-length-dependent, where the protein accumulation increases proportionally with repeat-length or where the detection of RAN-proteins is only possible if a certain number of repeats are present. In context, many RAN-positive disorders show repeat-length-dependent genetic anticipation. Furthermore, some brain areas in patients with RAN-positive disorders show clustered accumulation and high-dense staining of RAN-proteins while other areas within the same brain are much less affected, raising further questions about initiation processes of RAN-translation. HD, DM2, and SCA8 for example show RAN protein accumulation specifically in damaged white matter brain regions, accompanied by astroglyosis, demyelination, microglial, and/or caspase activation, further suggesting an involvement of the accumulated proteins in these neurodegenerative pathways by mechanisms such as ER stress, oxidative stress and DNA damage, protein translation abnormalities, nuclear transport deficits or stress granule formation (Banez-Coronel and Ranum 2019).

4.1.2.2 RNA accumulation

After the end of my internship in the Institute of human genetics, university hospital Essen, the teams of Prof. Dr. Christel Depienne in Essen and Dr. Mark Corbett in Adelaide further investigated the presence of altered mRNA or protein levels in blood cells and fibroblasts in individuals with an expansion in *MARCHF6* or *STARD7* compared with those of non-expansion-carrier controls. Conversely, although both genes are ubiquitously expressed, they could not detect either an increase in intron 1 retention or any read filled with UUUUA or UUUCA repeats as an expression of accumulation or formation of RNA foci. Furthermore, the expression levels of *MARCHF6* and *STARD7* were comparable to those of unaffected controls (Corbett et al. 2019; Florian et al. 2019). These results indicate that the expansion does not alter its expression in blood cells and fibroblasts of carrier individuals compared to those of unaffected controls.

This contrasts with previous observations made in post-mortem brains of Japanese patients with *SAMD12* expansions, where reads filled with UUUUA/UUUCA repeats were detected and RNA foci associated with abortive transcription following *SAMD12* expansions were observed in postmortem brains of Japanese patients (Ishiura et al. 2018).

This discrepancy could be the reflect of processes happening only in neuronal cells, limiting the accumulation of UUUCA repeats to neuronal tissues. However, the number of possible RNA foci found in the post-mortem brains was very limited, showcasing two foci for the homozygous brain and one focus for the heterozygous brain (Ishiura et al. 2018). In contrast other diseases associated with expansion RNA-foci, such as myotonic dystrophy, usually show many larger aggregates (Urbanek and Krzyzosiak 2016). Additionally, no control treatments with RNase or DNase were provided to prove that the detected dots corresponded to RNA, not DNA. We think that this question clearly needs to be further addressed, ideally in additional human brain samples or appropriate cellular or organoid models.

4.1.2.2.1 The role of the cerebellum

As previously described, even though the origin of the tremor seen in FAME-patients is a cortical hyperexcitability, it is evident that the cerebellum plays a part in the pathophysiology of FAME. Not only do many FAME-related genes (i.e. *SAMD12*, *MARCHF6* and *YEATS2*) show high expression in the cerebellum (Table 6), but neuropathological examinations show changes in the cerebellum (i.e. a diffuse loss of Purkinje cells, dendritic sprouts, neuronal loss in the dentate nucleus and halo-like amorphous material around the cytoplasm of remaining Purkinje cells) in various publications (van Rootselaar et al. 2007; van Rootselaar et al. 2004; Carr et al. 2007; Ishiura et al. 2018).

Electrophysiological findings (1.1.1.4.1 Electrophysiology) in the FAME-associated Tremor suggest that the movements are generated by an abnormal sensorimotor discharge arising the cerebral cortex. This translates to a sudden and brief activation of the corticospinal tract neurons, which results in a rhythmic oscillation of the involved muscles (mainly the upper, distal limbs). The burst of activity is hereby so brief, that it is

likely that there is a inhibitory mechanism that terminates the excitation (Latorre et al. 2020).

This rhythmicity that differs cortical tremor from cortical myoclonus could be explained by interactions between local (cortical) factors and synchronization by external sources such as local circuits linking corticospinal tract neurons with interneurons or more distant connections, such as unstable cortical loops. Involving the cerebellum it could be proposed, that abnormal cerebello-thalamo-cortical projections could change the gain of sensorimotor connections (Latorre et al. 2020).

This would be consistent with the hypothesis that decreased cerebellar drive from the hypoplastic cerebellum results in abnormalities of regulatory mechanisms involving the motor cortex resulting in cortical myoclonus (Rocchi et al. 2019) and pathophysiological findings showing the reduction of enhanced LLSR in patients with cerebellar atrophy by appliance of anodal transcranial direct current stimulation in order to increase cerebello-motor cortex inhibition (Diener et al. 1984).

4.1.2.3 Spinocerebellar ataxia type 37 (SCA 37)

While we can clearly prove a causative connection between the TTTCA/TTTTA expansion in our patients and their FAME disorder, SCA37 is an autosomal dominant neurodegenerative disorder clinically characterized by pure cerebellar ataxia and dysarthria which is also caused by an insertion of unstable ATTC repeats in the middle of an ATTTT repetition in the noncoding region of the neurodevelopmental *DABI*-gene. The total pentanucleotide repeat size varies from approximately 190 to 220 units, inversely correlating to the age at onset (Seixas et al. 2017).

Pathophysiological findings include AUUUC repeat aggregation in a human cell lines and an experimental study with zebra fish embryos suggests a toxic effect of these RNA accumulates (Seixas et al. 2017).

The difference in phenotype might be attributed to the highly specific expression of *DABI* in the cerebellum, but several genes where FAME expansion occurs (e.g. *MARCHF6*, *STARD7* and *TNRC6A*) are also highly expressed in the cerebellum. It could also be

determined by the size of the expansion, in which FAME expansion are much longer than those of SCA 37 members.

In any case, the presence of TTTCA-repeats as a causative mutation in another disease suggests that the expression profile of the gene where the expansion occurs might also be important but does not suffice by itself to determine the clinical presentation.

4.1.2.4 Intronic (TTTGA)_n Insertion in *SAMD12*

During the duration of this study, Cen et al identified an expanded intronic (TTTGA)_n insertion at the same site at the previously reported TTTCA expansion in *SAMD12* co-segregating with the FAME phenotype in one Chinese pedigree (Cen et al. 2019).

These findings raise further questions about the disease-causing mechanisms, suggesting that other repeat motifs than TTTCA could also be pathogenic although this has to be confirmed by additional studies.

4.1.3 Tandem repeat expansions in neurological diseases

Tandem repeats (TRs) as present in the described expansion are intrinsically unstable elements present ubiquitously in human genomes, contributing about 3% of genomic DNA (Gymrek et al. 2016). Expansion of TR length both in coding and non-coding regions is a well-known process that can result in genetic disorders affecting the CNS, such as Fragile X syndrome, Huntington disease (HD), spinocerebellar ataxias (SCA), frontotemporal dementia (FTD), and amyotrophic lateral sclerosis (ALS), as well as at least 20 others. The phenotypes of these diseases often show an inverse correlation between the number of tandem repeats and the age at disease onset or the severity of the disease (Hannan 2018).

4.1.3.1 Neurological diseases with intronic pentanucleotide repeat expansions

Disease	Phenotype	Inheritance	Mutation	Path. Mechanism
SCA 10	Cerebellar ataxia, Seizures	Autosomal dominant	ATTCT expansion in intron 9 of <i>SCA10</i>	Repeat motif does not differ from reference genome, number of repeats is pathogenic
SCA 31	Cerebellar ataxia, Purkinje cell degeneration	Autosomal dominant	TTCCA/TTTTA expansion in intronic region shared by <i>Nedd4</i> (<i>BEANI</i>) and <i>Kinase2</i> (<i>TK2</i>)	Repeat motif, gain- of-toxic-function
SCA 37	Cerebellar ataxia, Dysarthria	Autosomal dominant	TTTCA/TTTTA expansion in noncoding region of <i>DABI</i>	Repeat motif, RNA accumulates
CANVAS	Cerebellar ataxia, impaired vestibular function, sensory neuropathy	Autosomal recessive	AAGGG expansion in intron 2 of <i>RFC1</i>	Repeat motif, other than that unclear
FAME	Cortical tremor, Epilepsie	Autosomal dominant	TTTCA/TTTTA expansions in introns of several genes	Repeat motif, other than that unclear

Table 10. Neurological diseases with intronic pentanucleotide repeat expansions (Matsuura et al. 2000; Sato et al. 2009; Ishiguro et al. 2017; Seixas et al. 2017; Cortese et al. 2019; Ishiura et al. 2018; Ishiura and Tsuji 2020; Corbett et al. 2019; Florian et al. 2019)

4.1.4 Technical realisation and difficulties

4.1.4.1 Clinical assessment

During the duration of this study it has become evident that the clinical assessment of FAME patients is not without difficulties.

FAME is a late onset disease, where first symptoms can occur in ages over 40 years in less severely affected individuals. This leads to unidentified individuals in younger generations, that can be labelled as individuals with incomplete penetrance, despite being affected. We could showcase this with two individuals who had previously been reported

unaffected when examined in younger ages and now show clear signs of a FAME phenotype when re-examined. Unfortunately, this leads to the presence of individuals that are reported unaffected but carry an expansion, such as individuals 1-V-6, 2-IV-2 and 2-IV-6.

Furthermore, with higher age it is harder to differentiate a FAME-related tremor to a physiological tremor of the elderly. With limited access to instrumental testing, this can lead to individuals wrongly assessed as phenocopies. In this study, we could detect a tremor of uncertain origin in two individuals of family 2 (Individual 2-IV-12 and 2-IV-13). Both individuals were previously reported unaffected with FAME and only started to show tremulous signs at over 70 years of age, exceeding the usual age at onset of other family members by far. In close collaboration with the neurologist Prof. Dr. Stephan Klebe, university-hospital Essen, we therefore evaluated the tremor to not be FAME related.

As a result of these difficulties, we conclude that the optimal time for clinical evaluation of a possible FAME patient is between 30(40)-70 years of age. Outside of this timeframe it is more difficult to correctly diagnose the FAME disease by clinical evaluation and we need to rely more heavily on instrumental testing and possibly on genetical testing in the future.

4.1.4.2 Characterization of the expansion

Short read based technologies such as fragment length analysis or Sanger sequencing based on conventional PCR fail to amplify fragments as big as the analyzed expansions. We managed to establish a RP-PCR that enables to detect the presence of expansions and its configuration by visualization of each repeat. Still, it is not possible to predict the accurate size of the expansion because the efficiency of the replication during RP-PCR declines with the length of the fragment, leading to smaller and smaller peaks.

To get a more accurate insight into expansion length and structure we therefore used long-read based technologies. Low coverage sequencing using Oxford-Nanopore sequencing allowed us to retrieve one to four reads displaying the expansion per individual. Using this method, we first observed the substantial variability in reads covering the expansions. However, nanopore sequencing still leads to a high percentage of sequencing errors,

preventing its use to look at the single base level. We therefore used molecular combing combined with fluorescent in situ hybridization (FISH) to confirm our results. Molecular combing did not only provide about 100 ROIs per slide, allowing to calculate more accurate mean sizes of the expansion and its different parts, but could also confirm the somatic mosaicism and identify rearrangements at the expansion site. Nevertheless, the analysis of the ROIs is made by hand via computer software and is therefore prone to human errors, possibly over- or underestimating the exact expansion-sizes. In order to get a full insight into the FAME expansions we therefore still rely on a combination of tests.

4.1.5 Perspectives and future studies

Since there are still some families with an undistinguishable FAME-phenotype left that have yet failed to link to one of the known expansion sites, we hypothesize that other expansion sites remain to be identified. The goal of future projects should therefore be the development of a diagnostic strategy for FAME patients. We suggest the usage of established bioinformatics pipelines like ExpansionHunter to search for repeat expansions from Illumina genome data and confirmation through a combination of short-read and long-read based technologies. In addition, the remaining families could be screened for an expansion in one of the known expansion-sites using the already established and published methods. For further characterization of the expansion, the aim should be to analyse as many FAME patients as possible, to get more valid information about mean sizes and to be able to do more detailed genotype-phenotype correlation work.

In addition, further studies are needed to understand how the expansions occur and cause FAME. It is proven that the FAME expansions are associated with a common haplotype. Most recent, a Thai family, a Sri Lankan family and a Indian family were linked to the BAFME1 locus in intron 4 of *SAMD12*, spanning this mutation over 5 countries. Using haplotype analysis, a core haplotype was found containing the TTTTCA-insertion. The most recent ancestor was hereby estimated to have lived around 495/670 generations/an equivalent of about 12.375/17.000 years ago (Yeetong et al. 2020; Bennett et al. 2020). It remains unclear whether the expansion was already present on the haplotype at that time and has been preserved for all this time or if repeats would have expanded independently from the same predisposing haplotype more recently. In order to help

identify the mechanism by which these expansions form one could try to identify predisposing or protective haplotypes. Furthermore, the exact mechanism of TR expansion remains unknown. The coexistence of different motifs within the expansion indicates a highly complex mechanism that remains to be analysed.

To understand how expansions cause FAME, we need further analysis of neuronal tissue to compare other FAME loci to the findings in *SAMD12* patients and study possible effects of the expansions.

4.2 Essential tremor

In this study, we could not detect TR-expansions in *STARD7* or *MARCHF6* in ET patients and therefore failed to provide evidence that a substantial amount of ET diagnosed patients are in truth affected with the FAME-disease.

Beyond that, we detected alternative TR motifs in intron 1 of *STARD7* in members of 9 ET-families (31% of all families, 40% of the tested families) with TGTTA as the most frequent motif. Interestingly, these motifs oftentimes did not segregate with the ET phenotype. Going further, we could provide evidence that the TGTTA motif is also present in up to 13,33% of the normal population world-wide and can also be found in humanoid primates such as the orangutan.

It remains unclear why this variability seems to reoccur regularly in low percentages. Therefore, the presence of alternating motifs (in ET) should be assessed further in the future, with a larger cohort and extensive sequencing of all individuals. With the data gathered in this study, we cannot link the existence of alternative motifs to the ET phenotype. It would require a larger cohort of control individuals that ideally are neurologically examined to properly compare the frequency of the presence of alternative motifs between ET patients and the control population. For now, the comparison of frequency remains difficult, because of the comparatively small amount of individuals in our tested groups and the high prevalence of ET in the population (approximately 0.9% of individuals in all populations) that results in an unsure ET-status of blood donors.

Furthermore, we have faced technical difficulties in the manipulation and amplification of the TA-rich regions observed in intron 1 of *MARCHF6* and intron 1 of *STARD7* (Not

only during the evaluation of our ET-individuals, but also during the FAME-experiments). This phenomenon is well known in TA-rich regions and can result (as well as other underlying factors) in fragile sites (FS) defined as “loci that exhibit chromosome fragility as visible gaps and breaks on metaphase chromosomes or by physical or genetic assays of chromosome breakage” (Kaushal and Freudenreich 2019). In 2019 Irony-Tur Sinai et al found that the introduction of AT-dinucleotide rich sequences into previously stable sites can generate gene fragility under replication stress, providing clear evidence for the direct role of AT-DRSs in driving chromosomal fragility. One underlying mechanism driving the fragility is the tendency of AT-rich regions to form stable secondary structures under the single strand state as during replication (Irony-Tur Sinai et al. 2019). This can then lead to perturbed DNA-replication or mutation of the site (i.e. as observed in the expansion during FAME-mutation) (Kaushal and Freudenreich 2019).

Accordingly, during our study, we observed that the introduction of a C or G to TA-rich regions resulted in an amplification advantage in conventional PCR. In some individuals and families, we were able to validate our results with further testing, but some amplification difficulties remained unsolved. This leads to difficulties in first assessment of individuals, making the decision whether result are valid or flawed due to difficulties in technical procedures, challenging.

4.2.1 Repeat expansions and Essential Tremor

During the duration of this study, Sun et al. discovered an expansion of GGC repeats in *NOTCH2NLC*, associated with Essential tremor.

After identifying the expansion in 2 families with Long-read sequencing, they screened a total of 195 additional ET families using RP-PCR and GC-PCR essays and found that 9 other families carried expanded GGC repeats in *NOTCH2NLC*. The expansion size of affected family-members showed an average of 108,6 GGC-repeats, whereas unaffected family-members, as well as control-individuals showed repeat-sizes of 4 to 41 GGC repeats. Besides GGC, expansion motifs like GGA and AGC were also observed, but less frequent. (Sun et al. 2020)

This suggest that repeat expansions are likely to be found causative in ET and that, even though the GGAexpansion might be causative for the ET phenotype in these families, other ET-families may underly different mutations or mutation sites. The pathological mechanism could even be similar to the FAME mechanism.

Interestingly, neuronal intranuclear inclusion disease (NIID), a highly variable neurodegenerative disease, most commonly associated with dementia and peripheral neuropathy is also caused by GGC repeat expansions in *NOTCH2NLC*. One main difference is, that the repeat sizes in ET are significantly less than the repeat sizes in NIID. (Sun et al. 2020) This provides further evidence, that not only the repeat size seems to drive the phenotype, as suggested in the example of FAME and SCA37.

5 Conclusion

In this study, we provide evidence that FAME3 (FAME - familial autosomal dominant myoclonus and epilepsy) and FAME2 are caused by intronic T(Thymine)TTTA(Adenine)/TTTC(Cytosine)A expansions in *MARCHF6* (FAME3) and *STARD7* (FAME2).

FAME is an autosomal dominant, very slowly progressive condition characterized by cortical tremor and generalized myoclonic seizures. In the last years different chromosome loci, identified through linkage analysis have been reported, but the genetic variants underlying the disorder had remained elusive for 20 years despite extensive sequencing of genes contained in these intervals.

In 2018, intronic expansions composed of mixed TTTTA/TTTCA repeats in *SAMD12* on chromosome 8q24 have been identified as the main cause of FAME1 (BAFME1 (Benign adult familial myoclonic epilepsy)) in the Japanese and Chinese populations.

Subsequently in this study, we investigated the presence of TTTTA/TTTCA repeat-expansions in *MARCHF6* and *STARD7*. In confirming the presence of these expansions in affected members of 4 families in *MARCHF6* and in 3 families in *STARD7* we provide further evidence that the FAME-disease is monogenetically caused by TTTTA/TTTCA repeat-expansions. Moreover, we could observe considerable variability in expansion length and structure, supporting the existence of multiple expansion configurations in blood cells and fibroblasts of the same individual and show that the largest expansions were associated with micro-rearrangements occurring near the expansion in 20% of cells. The pathological mechanisms by which TTTCA-repeats cause the FAME-phenotype remain uncertain. Furthermore, investigated whether ET (Essential tremor) could have a similar genetic basis than FAME. Unfortunately, we failed to find any TTTTA/TTTCA repeat expansions in our ET-cohort. Nevertheless, we found that a considerable number of ET-individuals carried a motif other than TTTTA at the *STARD7* locus. Investigating this phenomenon further, we found that the TG(Guanine)TTA repeat-motif was the most frequent alternative motif found in our control-population and was even present in the orangutan.

6 Zusammenfassung

In dieser Studie liefern wir Beweise dafür, dass FAME3 (FAME - familial autosomal dominant myoclonus and epilepsy) und FAME2 monogenetisch durch intronische T(Thymine)TTTA(Adenine)/TTTC(Cytosine)A-Expansionen in *MARCHF6* (FAME3) und *STARD7* (FAME2) verursacht werden.

FAME ist eine autosomal dominante, sehr langsam fortschreitende Erkrankung, die durch kortikalen Tremor und generalisierte myoklonische Anfälle gekennzeichnet ist. In den letzten Jahren wurden verschiedene Loci, die durch Linkage Analysis identifiziert wurden beschrieben, jedoch ist die Mutation, welche die der Erkrankung zugrunde liegende Ursache ist trotz umfassender Sequenzierung der in diesen Loci enthaltenen Gene über 20 Jahre nicht identifiziert worden.

In 2018 wurde die Expansion eines Introns, welches aus gemischten TTTTA / TTTCA-repeats in *SAMD12* auf Chromosom 8q24 besteht, als Hauptursache für FAME1 (BAFME1) in der japanischen und chinesischen Bevölkerung identifiziert. Indem wir das Vorhandensein dieser Expansionen bei erkrankten Mitgliedern von 4 Familien in *MARCHF6* und von 3 Familien in *STARD7* bestätigen, liefern wir weitere Beweise dafür, dass die FAME-Krankheit monogenetisch durch TTTTA / TTTCA repeat-Expansionen verursacht wird. Darüber hinaus konnten wir eine beträchtliche Variabilität der Expansionslänge und -struktur beobachten, was die Theorie der Existenz mehrerer Expansionskonfigurationen in Blutzellen und Fibroblasten desselben Individuums bestätigt. Wir zeigten zudem, dass die größten Expansionen mit Mikrorearrangements in der Nähe der Expansion in 20% der Zellen assoziiert waren.

Darüber hinaus haben wir versucht nachzuweisen, dass in einer großen ET(Essentieller Tremor)-Kohorte zumindest eine beträchtliche Anzahl von Personen fehldiagnostiziert ist und in Wahrheit tatsächlich von FAME betroffen ist. Leider konnten wir in unserer ET-Kohorte keine TTTTA / TTTCA-Repeatexpansionen finden. Wir fanden jedoch heraus, dass eine beträchtliche Anzahl von ET-Individuen in *STARD7* ein anderes Motiv als TTTTA trug. Weitere Untersuchungen ergaben jedoch, dass das TG(Guanine)TTA-Motiv als häufigstes alternatives Motiv auch in unserer Kontrollpopulation zu finden war und sogar im Orang-Utan vorhanden war.

7 References

1. Applied Biosystems. 2009. DNA Sequencing by Capillary Electrophoresis: Applied Biosystems Chemistry Guide (Applied Biosystems: Online).
2. Banez-Coronel, M., F. Ayhan, A. D. Tarabochia, T. Zu, B. A. Perez, S. K. Tusi, O. Pletnikova, D. R. Borchelt, C. A. Ross, R. L. Margolis, A. T. Yachnis, J. C. Troncoso, and L. P. Ranum. 2015. 'RAN Translation in Huntington Disease', *Neuron*, 88: 667-77.
3. Banez-Coronel, M., and L. P. W. Ranum. 2019. 'Repeat-associated non-AUG (RAN) translation: insights from pathology', *Lab Invest*, 99: 929-42.
4. Bennett, M. F., K. L. Oliver, B. M. Regan, S. T. Bellows, A. L. Schneider, H. Rafehi, N. Sikta, D. E. Crompton, M. Coleman, M. S. Hildebrand, M. A. Corbett, T. Kroes, J. Gecez, I. E. Scheffer, S. F. Berkovic, and M. Bahlo. 2020. 'Familial adult myoclonic epilepsy type 1 SAMD12 TTTCA repeat expansion arose 17,000 years ago and is present in Sri Lankan and Indian families', *Eur J Hum Genet*, 28: 973-78.
5. Bourdain, F., E. Apartis, J. M. Trocello, J. S. Vidal, P. Masnou, L. Vercueil, and M. Vidailhet. 2006. 'Clinical analysis in familial cortical myoclonic tremor allows differential diagnosis with essential tremor', *Mov Disord*, 21: 599-608.
6. Buijsen, R. A., J. A. Visser, P. Kramer, E. A. Severijnen, M. Gearing, N. Charlet-Berguerand, S. L. Sherman, R. F. Berman, R. Willemsen, and R. K. Hukema. 2016. 'Presence of inclusions positive for polyglycine containing protein, FMRpolyG, indicates that repeat-associated non-AUG translation plays a role in fragile X-associated primary ovarian insufficiency', *Hum Reprod*, 31: 158-68.
7. Carr, J. A., P. E. van der Walt, J. Nakayama, Y. H. Fu, V. Corfield, P. Brink, and L. Ptacek. 2007. 'FAME 3: a novel form of progressive myoclonus and epilepsy', *Neurology*, 68: 1382-9.
8. Cen, Z., Y. Chen, D. Yang, Q. Zhu, S. Chen, X. Chen, B. Wang, F. Xie, Z. Ouyang, Z. Jiang, A. Fu, B. Hu, H. Yin, X. Qiu, F. Yu, X. Du, W. Hao, Y. Liu, H. Wang, L. Wang, X. Yu, Y. Xiao, C. Liu, J. Xiao, Y. Zhou, W. Yang, B. Zhang, and W. Luo. 2019. 'Intronic (TTTGA)_n insertion in SAMD12 also causes familial cortical myoclonic tremor with epilepsy', *Mov Disord*, 34: 1571-76.

9. Cen, Z. D., F. Xie, J. F. Xiao, and W. Luo. 2016. 'Rational search for genes in familial cortical myoclonic tremor with epilepsy, clues from recent advances', *Seizure*, 34: 83-9.
10. Cen, Z., Z. Jiang, Y. Chen, X. Zheng, F. Xie, X. Yang, X. Lu, Z. Ouyang, H. Wu, S. Chen, H. Yin, X. Qiu, S. Wang, M. Ding, Y. Tang, F. Yu, C. Li, T. Wang, H. Ishiura, S. Tsuji, C. Jiao, C. Liu, J. Xiao, and W. Luo. 2018. 'Intronic pentanucleotide TTTCA repeat insertion in the SAMD12 gene causes familial cortical myoclonic tremor with epilepsy type 1', *Brain*, 141: 2280-88.
11. Corbett, M. A., T. Kroes, L. Veneziano, M. F. Bennett, R. Florian, A. L. Schneider, A. Coppola, L. Licchetta, S. Franceschetti, A. Suppa, A. Wenger, D. Mei, M. Pendziwiat, S. Kaya, M. Delledonne, R. Straussberg, L. Xumerle, B. Regan, D. Crompton, A. F. van Rootselaar, A. Correll, R. Catford, F. Bisulli, S. Chakraborty, S. Baldassari, P. Tinuper, K. Barton, S. Carswell, M. Smith, A. Berardelli, R. Carroll, A. Gardner, K. L. Friend, I. Blatt, M. Iacomino, C. Di Bonaventura, S. Striano, J. Buratti, B. Keren, C. Nava, S. Forlani, G. Rudolf, E. Hirsch, E. Leguern, P. Labauge, S. Balestrini, J. W. Sander, Z. Afawi, I. Helbig, H. Ishiura, S. Tsuji, S. M. Sisodiya, G. Casari, L. G. Sadleir, R. van Coller, M. A. J. Tijssen, K. M. Klein, Amjm van den Maagdenberg, F. Zara, R. Guerrini, S. F. Berkovic, T. Pippucci, L. Canafoglia, M. Bahlo, P. Striano, I. E. Scheffer, F. Brancati, C. Depienne, and J. Gecz. 2019. 'Intronic ATTTC repeat expansions in STARD7 in familial adult myoclonic epilepsy linked to chromosome 2', *Nat Commun*, 10: 4920.
12. Cortese, A., R. Simone, R. Sullivan, J. Vandrovcova, H. Tariq, W. Y. Yau, J. Humphrey, Z. Jaunmuktane, P. Sivakumar, J. Polke, M. Ilyas, E. Tribollet, P. J. Tomaselli, G. Devigili, I. Callegari, M. Versino, V. Salpietro, S. Efthymiou, D. Kaski, N. W. Wood, N. S. Andrade, E. Buglo, A. Rebelo, A. M. Rossor, A. Bronstein, P. Fratta, W. J. Marques, S. Zuchner, M. M. Reilly, and H. Houlden. 2019. 'Biallelic expansion of an intronic repeat in RFC1 is a common cause of late-onset ataxia', *Nat Genet*, 51: 649-58.
13. De Fusco, M., R. Vago, P. Striano, C. Di Bonaventura, F. Zara, D. Mei, M. S. Kim, S. Muallem, Y. Chen, Q. Wang, R. Guerrini, and G. Casari. 2014. 'The alpha2B-

adrenergic receptor is mutant in cortical myoclonus and epilepsy', *Ann Neurol*, 75: 77-87.

14. Depienne, C., E. Magnin, D. Bouteiller, G. Stevanin, C. Saint-Martin, M. Vidailhet, E. Apartis, E. Hirsch, E. LeGuern, P. Labauge, and L. Rumbach. 2010. 'Familial cortical myoclonic tremor with epilepsy: the third locus (FCMTE3) maps to 5p', *Neurology*, 74: 2000-3.

15. Deuschl, Günter. 2012. 'Extrapyramidalmotorische Störungen: Tremor.' in Deutsche Gesellschaft für Neurologie (ed.), *Leitlinien für Diagnostik und Therapie in der Neurologie* (Deutsche Gesellschaft für Neurologie).

16. Diener, H. C., J. Dichgans, M. Bacher, and B. Guschlbauer. 1984. 'Characteristic alterations of long-loop "reflexes" in patients with Friedreich's disease and late atrophy of the cerebellar anterior lobe', *J Neurol Neurosurg Psychiatry*, 47: 679-85.

17. Florian, R. T., F. Kraft, E. Leitao, S. Kaya, S. Klebe, E. Magnin, A. F. van Rootselaar, J. Buratti, T. Kuhnel, C. Schroder, S. Giesselmann, N. Tschernoster, J. Altmueller, A. Lamiral, B. Keren, C. Nava, D. Bouteiller, S. Forlani, L. Jornea, R. Kubica, T. Ye, D. Plassard, B. Jost, V. Meyer, J. F. Deleuze, Y. Delpu, M. D. M. Avarello, L. S. Vijfhuizen, G. Rudolf, E. Hirsch, T. Kroes, P. S. Reif, F. Rosenow, C. Ganos, M. Vidailhet, L. Thivard, A. Mathieu, T. Bourgeron, I. Kurth, H. Rafehi, L. Steenpass, B. Horsthemke, Fame consortium, E. LeGuern, K. M. Klein, P. Labauge, M. F. Bennett, M. Bahlo, J. Gecz, M. A. Corbett, M. A. J. Tijssen, Amjrm van den Maagdenberg, and C. Depienne. 2019. 'Unstable TTTTA/TTTCA expansions in MARCH6 are associated with Familial Adult Myoclonic Epilepsy type 3', *Nat Commun*, 10: 4919.

18. Gendron, T. F., K. F. Bieniek, Y. J. Zhang, K. Jansen-West, P. E. Ash, T. Caulfield, L. Daugherty, J. H. Dunmore, M. Castanedes-Casey, J. Chew, D. M. Cosio, M. van Blitterswijk, W. C. Lee, R. Rademakers, K. B. Boylan, D. W. Dickson, and L. Petrucelli. 2013. 'Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS', *Acta Neuropathol*, 126: 829-44.

19. Green, K. M., A. E. Linsalata, and P. K. Todd. 2016. 'RAN translation-What makes it run?', *Brain Res*, 1647: 30-42.
20. Guerrini, R., P. Bonanni, A. Patrignani, P. Brown, L. Parmeggiani, P. Grosse, P. Brovedani, F. Moro, P. Aridon, R. Carrozzo, and G. Casari. 2001. 'Autosomal dominant cortical myoclonus and epilepsy (ADCME) with complex partial and generalized seizures: A newly recognized epilepsy syndrome with linkage to chromosome 2p11.1-q12.2', *Brain*, 124: 2459-75.
21. Gymrek, M., T. Willems, A. Guilmatre, H. Zeng, B. Markus, S. Georgiev, M. J. Daly, A. L. Price, J. K. Pritchard, A. J. Sharp, and Y. Erlich. 2016. 'Abundant contribution of short tandem repeats to gene expression variation in humans', *Nat Genet*, 48: 22-9.
22. Hannan, A. J. 2018. 'Tandem repeats mediating genetic plasticity in health and disease', *Nat Rev Genet*, 19: 286-98.
23. Ikeda, A., R. Kakigi, N. Funai, R. Neshige, Y. Kuroda, and H. Shibasaki. 1990. 'Cortical tremor: a variant of cortical reflex myoclonus', *Neurology*, 40: 1561-5.
24. Inazuki, G., H. Naito, E. Ohama, Y. Kawase, Y. Honma, S. Tokiguchi, S. Hasegawa, K. Tamura, K. Kawai, H. Nagai, and et al. 1990. '[A clinical study and neuropathological findings of a familial disease with myoclonus and epilepsy--the nosological place of familial essential myoclonus and epilepsy (FEME)]', *Seishin Shinkeigaku Zasshi*, 92: 1-21.
25. Invitrogen. 2009. User Manual: pcDNA3.1/V5-His TOPO TA Expression Kit (Invitrogen: Online).
26. Irony-Tur Sinai, M., A. Salamon, N. Stanleigh, T. Goldberg, A. Weiss, Y. H. Wang, and B. Kerem. 2019. 'AT-dinucleotide rich sequences drive fragile site formation', *Nucleic Acids Res*, 47: 9685-95.
27. Ishiguro, T., N. Sato, M. Ueyama, N. Fujikake, C. Sellier, A. Kanegami, E. Tokuda, B. Zamiri, T. Gall-Duncan, M. Mirceta, Y. Furukawa, T. Yokota, K. Wada, J. P. Taylor, C. E. Pearson, N. Charlet-Berguerand, H. Mizusawa, Y. Nagai, and K. Ishikawa. 2017.

'Regulatory Role of RNA Chaperone TDP-43 for RNA Misfolding and Repeat-Associated Translation in SCA31', *Neuron*, 94: 108-24 e7.

28. Ishiura, H., K. Doi, J. Mitsui, J. Yoshimura, M. K. Matsukawa, A. Fujiyama, Y. Toyoshima, A. Kakita, H. Takahashi, Y. Suzuki, S. Sugano, W. Qu, K. Ichikawa, H. Yurino, K. Higasa, S. Shibata, A. Mitsue, M. Tanaka, Y. Ichikawa, Y. Takahashi, H. Date, T. Matsukawa, J. Kanda, F. K. Nakamoto, M. Higashihara, K. Abe, R. Koike, M. Sasagawa, Y. Kuroha, N. Hasegawa, N. Kanosawa, T. Kondo, T. Hitomi, M. Tada, H. Takano, Y. Saito, K. Sanpei, O. Onodera, M. Nishizawa, M. Nakamura, T. Yasuda, Y. Sakiyama, M. Otsuka, A. Ueki, K. I. Kaida, J. Shimizu, R. Hanajima, T. Hayashi, Y. Terao, S. Inomata-Terada, M. Hamada, Y. Shirota, A. Kubota, Y. Ugawa, K. Koh, Y. Takiyama, N. Ohsawa-Yoshida, S. Ishiura, R. Yamasaki, A. Tamaoka, H. Akiyama, T. Otsuki, A. Sano, A. Ikeda, J. Goto, S. Morishita, and S. Tsuji. 2018. 'Expansions of intronic TTTCA and TTTTA repeats in benign adult familial myoclonic epilepsy', *Nat Genet*, 50: 581-90.

29. Ishiura, H., and S. Tsuji. 2020. 'Advances in repeat expansion diseases and a new concept of repeat motif-phenotype correlation', *Curr Opin Genet Dev*, 65: 176-85.

30. Jankovik, J., and E. Tolosa. 1988. *Parkinson's Disease and Movement Disorders* (Urban & Schwarzenberg: Baltimore, Maryland).

31. Kaushal, S., and C. H. Freudenreich. 2019. 'The role of fork stalling and DNA structures in causing chromosome fragility', *Genes Chromosomes Cancer*, 58: 270-83.

32. Krans, A., M. G. Kearse, and P. K. Todd. 2016. 'Repeat-associated non-AUG translation from antisense CCG repeats in fragile X tremor/ataxia syndrome', *Ann Neurol*, 80: 871-81.

33. Kudo, J., T. Kudo, and T. Yamauchi. 1984. '[Seven families with hereditary familial tremor and epilepsy]', *Rinsho Shinkeigaku*, 24: 1-8.

34. Kuo, S. H., C. Erickson-Davis, A. Gillman, P. L. Faust, J. P. Vonsattel, and E. D. Louis. 2011. 'Increased number of heterotopic Purkinje cells in essential tremor', *J Neurol Neurosurg Psychiatry*, 82: 1038-40.

35. Latorre, A., L. Rocchi, F. Magrinelli, E. Mulroy, A. Berardelli, J. C. Rothwell, and K. P. Bhatia. 2020. 'Unravelling the enigma of cortical tremor and other forms of cortical myoclonus', *Brain*, 143: 2653-63.
36. Leggett, R. M., and M. D. Clark. 2017. 'A world of opportunities with nanopore sequencing', *J Exp Bot*, 68: 5419-29.
37. Louis, E. D. 2011. 'Essential tremor', *Handb Clin Neurol*, 100: 433-48.
38. Louis, E. D., and J. J. Ferreira. 2010. 'How common is the most common adult movement disorder? Update on the worldwide prevalence of essential tremor', *Mov Disord*, 25: 534-41.
39. Matsuura, T., T. Yamagata, D. L. Burgess, A. Rasmussen, R. P. Grewal, K. Watase, M. Khajavi, A. E. McCall, C. F. Davis, L. Zu, M. Achari, S. M. Pulst, E. Alonso, J. L. Noebels, D. L. Nelson, H. Y. Zoghbi, and T. Ashizawa. 2000. 'Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10', *Nat Genet*, 26: 191-4.
40. Mikami, M., T. Yasuda, A. Terao, M. Nakamura, S. Ueno, H. Tanabe, T. Tanaka, T. Onuma, Y. Goto, S. Kaneko, and A. Sano. 1999. 'Localization of a gene for benign adult familial myoclonic epilepsy to chromosome 8q23.3-q24.1', *Am J Hum Genet*, 65: 745-51.
41. Mori, K., S. M. Weng, T. Arzberger, S. May, K. Rentzsch, E. Kremmer, B. Schmid, H. A. Kretschmar, M. Cruts, C. Van Broeckhoven, C. Haass, and D. Edbauer. 2013. 'The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS', *Science*, 339: 1335-8.
42. Rocchi, L., A. Latorre, J. Ibanez Pereda, D. Spampinato, K. E. Brown, J. Rothwell, and K. Bhatia. 2019. 'A case of congenital hypoplasia of the left cerebellar hemisphere and ipsilateral cortical myoclonus', *Mov Disord*, 34: 1745-47.
43. Sato, N., T. Amino, K. Kobayashi, S. Asakawa, T. Ishiguro, T. Tsunemi, M. Takahashi, T. Matsuura, K. M. Flanigan, S. Iwasaki, F. Ishino, Y. Saito, S. Murayama, M. Yoshida, Y. Hashizume, Y. Takahashi, S. Tsuji, N. Shimizu, T. Toda, K. Ishikawa, and H. Mizusawa. 2009. 'Spinocerebellar ataxia type 31 is associated with "inserted" penta-nucleotide repeats containing (TGGAA)n', *Am J Hum Genet*, 85: 544-57.

44. Seixas, A. I., J. R. Loureiro, C. Costa, A. Ordonez-Ugalde, H. Marcelino, C. L. Oliveira, J. L. Loureiro, A. Dhingra, E. Brandao, V. T. Cruz, A. Timoteo, B. Quintans, G. A. Rouleau, P. Rizzu, A. Carracedo, J. Bessa, P. Heutink, J. Sequeiros, M. J. Sobrido, P. Coutinho, and I. Silveira. 2017. 'A Pentanucleotide ATTTTC Repeat Insertion in the Non-coding Region of DAB1, Mapping to SCA37, Causes Spinocerebellar Ataxia', *Am J Hum Genet*, 101: 87-103.
45. Singh, S., A. Zhang, S. Dlouhy, and S. Bai. 2014. 'Detection of large expansions in myotonic dystrophy type 1 using triplet primed PCR', *Front Genet*, 5: 94.
46. Soragni, E., L. Petrosyan, T. A. Rinkoski, E. D. Wieben, K. H. Baratz, M. P. Fautsch, and J. M. Gottesfeld. 2018. 'Repeat-Associated Non-ATG (RAN) Translation in Fuchs' Endothelial Corneal Dystrophy', *Invest Ophthalmol Vis Sci*, 59: 1888-96.
47. Stogmann, E., E. Reinthaler, S. Eltawil, M. A. El Etribi, M. Hemedda, N. El Nahhas, A. M. Gaber, A. Fouad, S. Edris, A. Benet-Pages, S. H. Eck, E. Pataraiia, D. Mei, A. Brice, S. Lesage, R. Guerrini, F. Zimprich, T. M. Strom, and A. Zimprich. 2013. 'Autosomal recessive cortical myoclonic tremor and epilepsy: association with a mutation in the potassium channel associated gene CNTN2', *Brain*, 136: 1155-60.
48. Urbanek, M. O., and W. J. Krzyzosiak. 2016. 'RNA FISH for detecting expanded repeats in human diseases', *Methods*, 98: 115-23.
49. van den Ende, T., S. Sharifi, S. M. A. van der Salm, and A. F. van Rootselaar. 2018. 'Familial Cortical Myoclonic Tremor and Epilepsy, an Enigmatic Disorder: From Phenotypes to Pathophysiology and Genetics. A Systematic Review', *Tremor Other Hyperkinet Mov (N Y)*, 8: 503.
50. van Rootselaar, A. F., E. Aronica, E. N. Jansen Steur, J. M. Rozemuller-Kwakkel, R. A. de Vos, and M. A. Tijssen. 2004. 'Familial cortical tremor with epilepsy and cerebellar pathological findings', *Mov Disord*, 19: 213-7.
51. van Rootselaar, A. F., A. J. Groffen, B. de Vries, P. M. C. Callenbach, G. W. E. Santen, S. Koelewijn, L. S. Vijfhuizen, A. Buijink, M. A. J. Tijssen, and Amjm van den Maagdenberg. 2017. 'delta-Catenin (CTNND2) missense mutation in familial cortical myoclonic tremor and epilepsy', *Neurology*, 89: 2341-50.

52. van Rootselaar, A. F., N. M. Maurits, J. H. Koelman, J. H. van der Hoeven, L. J. Bour, K. L. Leenders, P. Brown, and M. A. Tijssen. 2006. 'Coherence analysis differentiates between cortical myoclonic tremor and essential tremor', *Mov Disord*, 21: 215-22.
53. van Rootselaar, A. F., S. M. van der Salm, L. J. Bour, M. J. Edwards, P. Brown, E. Aronica, J. M. Rozemuller-Kwakkel, P. J. Koehler, J. H. Koelman, J. C. Rothwell, and M. A. Tijssen. 2007. 'Decreased cortical inhibition and yet cerebellar pathology in 'familial cortical myoclonic tremor with epilepsy'', *Mov Disord*, 22: 2378-85.
54. van Rootselaar, A. F., I. N. van Schaik, A. M. van den Maagdenberg, J. H. Koelman, P. M. Callenbach, and M. A. Tijssen. 2005. 'Familial cortical myoclonic tremor with epilepsy: a single syndromic classification for a group of pedigrees bearing common features', *Mov Disord*, 20: 665-73.
55. Yasuda, Takeshi. 1991. 'Benign Adult Familial Myoclonic Epilepsie (BAFME)', *Kawasaki Med J.*, 17: 1 - 13.
56. Yeetong, P., S. Ausavarat, R. Bhidayasiri, K. Piravej, N. Pasutharnchat, T. Desudchit, C. Chunharas, J. Loplumlert, C. Limotai, K. Suphapeetiporn, and V. Shotelersuk. 2013. 'A newly identified locus for benign adult familial myoclonic epilepsy on chromosome 3q26.32-3q28', *Eur J Hum Genet*, 21: 225-8.
57. Yeetong, P., C. Chunharas, M. Pongpanich, M. F. Bennett, C. Srichomthong, N. Pasutharnchat, K. Suphapeetiporn, M. Bahlo, and V. Shotelersuk. 2020. 'Founder effect of the TTTCA repeat insertions in SAMD12 causing BAFME1', *Eur J Hum Genet*.
58. Yeetong, P., M. Pongpanich, C. Srichomthong, A. Assawapitaksakul, V. Shotelersuk, N. Tantirukdham, C. Chunharas, K. Suphapeetiporn, and V. Shotelersuk. 2019. 'TTTCA repeat insertions in an intron of YEATS2 in benign adult familial myoclonic epilepsy type 4', *Brain*, 142: 3360-66.
59. Yu, M., K. Ma, P. L. Faust, L. S. Honig, E. Cortes, J. P. Vonsattel, and E. D. Louis. 2012. 'Increased number of Purkinje cell dendritic swellings in essential tremor', *Eur J Neurol*, 19: 625-30.

60. Zu, T., J. D. Cleary, Y. Liu, M. Banez-Coronel, J. L. Bubenik, F. Ayhan, T. Ashizawa, G. Xia, H. B. Clark, A. T. Yachnis, M. S. Swanson, and L. P. W. Ranum. 2017. 'RAN Translation Regulated by Muscleblind Proteins in Myotonic Dystrophy Type 2', *Neuron*, 95: 1292-305 e5.
61. Zu, T., B. Gibbens, N. S. Doty, M. Gomes-Pereira, A. Huguet, M. D. Stone, J. Margolis, M. Peterson, T. W. Markowski, M. A. Ingram, Z. Nan, C. Forster, W. C. Low, B. Schoser, N. V. Somia, H. B. Clark, S. Schmechel, P. B. Bitterman, G. Gourdon, M. S. Swanson, M. Moseley, and L. P. Ranum. 2011. 'Non-ATG-initiated translation directed by microsatellite expansions', *Proc Natl Acad Sci U S A*, 108: 260-5.

8 Abbreviations and Units

8.1 Abbreviations

A	Adenine
ADCME	Autosomal dominant cortical myoclonus and epilepsy
AED	Anti-Epileptic drugs
ALS	Amyotrophe Lateralsklerose
ASD	Autism Spectrum Disorder
BAFME	Benign Adult Familial Myoclonic Epilepsy
BB	Betablocker
BZD	Benzodiazepine
C	Cytosine
CANVAS	Cerebellar ataxia, neuropathy, and vestibular areflexia syndrome
CBZ	Carbamazepine
CCG	Cologne Center for Genomics
cDNA	Complementary DNA
CNS	Central Nervous System
CO ₂	Carbon dioxide
CrtTr	Cortical Tremor
DMI	Myotonic dystrophy
DNA	Deoxyribonucleic Acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotidetriphosphate
Dr.	Doctor
dsDNA	double stranded DNA
EEG	Electroencephalography
EMG	Electromyography
ER	Endoplasmatisches Reticulum
ET	Essential Tremor
EtBr	Ethidium bromide
EtOH	Ethanol
Exp	Expanded

Ext	Extensors
F	Female
FA	Formaldehyde
6-FAM	6-Carboxyfluorescein
FAME	Familial adult myoclonic epilepsy
FCMT	Familial cortical myoclonic tremor
FCMTE	Familial cortical myoclonic tremor with epilepsy
FCTE	Familial cortical tremor with epilepsy
FDI	First dorsal interosseus
FECD	Fuchs' endothelial corneal dystrophy
FEME	familial essential myoclonus and epilepsy
Fig	Figure
FISH	Fluorescens in situ hybridization
FMEA	Familial benign myoclonus epilepsy of adult onset
FTD	Frontotemporal demtia
FTLD	Familial frontotemporal lobar degeneration
FTM-Scale	Fahn-Tolosa-Marin-Scale
FXPOI	Fragile X-associated primary ovarian insufficiency
FXTAS	Fragile X tremor/ataxia syndrome
G	Guanine
GABA	Beta-aminobutyric acid
gDNA	Genomic DNA
GOF	Gain of function
GTCS	Generalized tonic-clonic seizures
GV	Genomic Vision
HD	Huntington Disease
HTE	Heredofamilial tremor and epilepsy
H ₂ O	Water
ICM	Institut du cerveau et de la moelle (Paris, France)
ID	Interlectual Disability
LAM	Lamotrigin
LLR	Long-Latency-Reflexes

LLSR	Long-Latency Sensory Reflexes
LOF	Loss of function
LVT	Levetiracetam
M	Male
MgCl ₂	Magnesium Chloride
MRI	Magnetic Resonance Imaging
mRNA	Messenger RNA
NaCl	Sodium Chloride
NIID	Neuronal intranuclear inclusion disease
OD	Optic density
ONT	Oxford Nanopore technologies
p	petit; short arm of a chromosome
PCR	Polymerase Chain Reaction
pH	Phot
Prof.	Professor
q	queue; long arm of a chromosome
RAN	Repeat Associated Non-AUG
RNA	Ribonucleic Acid
RNase	Ribonuclease
ROI	Region of interest
RP-PCR	Repeat Primed PCR
RT	Room Temperature
SCA	Spinocerebellar ataxia
SEP	Sensory-Evoked-Potentials
SNP	Single Nucleotide Polymorphism
STR	Short Tandem Repeat
T	Thymine
TAE	Tris-acetate-EDTA Buffer
TE	Tris-EDTA Buffer
TR	Tandem Repeat
TRE	Tandem Repeat Expansion
Tris	Tris(hydroxymethyl)aminomethane

U	Uracil
UV	Ultraviolet
VE-Water	Vollentsalzenes Wasser
VPA	Sodium Valproate
WNGME	Tryptophan-asparagineglycine-methionine-glutamic acid

8.2 Units

°C	degrees Celsius
aa	amino acid(s)
bp	base pair(s)
g	gram
g	earth's acceleration
h	hour(s)
Hz	Hertz
kb	kilobase(s)
l	litre
M	molar (mol/l)
μ	micro (10 ⁶)
m	milli (10 ³)
Mb	megabase
min	minute
ml	millilitre
n	nano (10 ⁻⁹)
rpm	rotations per minute
s	second(s)
sec	second(s)
U	unit(s)
V	Volt
v/v	volume per volume
w/v	weight per volume

9 Figures

- Figure 1: Bipolar electromyogram of FAME patient
- Figure 2: Expansions in FAME Ishiura et al. 2018
- Figure 3: Expansion motifs and localisation in FAME Ishiura et al 2018
- Figure 4: Bipolar electromyogram of FAME and ET patients
- Figure 5: Pedigrees of all included FAME-families
- Figure 6: Schematic on repeat-primed PCR
- Figure 7: Schematic on sanger sequencing
- Figure 8: Schematic on nanopore sequencing
- Figure 9: Example of nanopore sequencing in intron 1 of *MARCHF6*
- Figure 10: Example of grading of drawings of an archimedis spiral
- Figure 11: Identification of TTTCA/TTTTA expansions in *MARCHF6* and *STARD7*
- Figure 12: Conventional PCR of the region of interest
- Figure 13: Somatic mosaicism in individual 1-III-5
- Figure 14: Characterization of *MARCHF6* expansions by nanopore sequencing
- Figure 15: Somatic mosaicism of *MARCHF6* expansions detected by molecular combing
- Figure 16: Distribution of expansion length and genotype-phenotype correlation
- Figure 17: Apparent non-mendelian inheritance observed in family 24
- Figure 18: Example of less efficiently quantified second allele
- Figure 19: Negative results for RP-PCR targeting TTTCA-repeats in *MARCHF6*
- Figure 20: Examples of individuals with one less efficiently quantified allele that carry an alternative motif
- Figure 21: Sanger sequencing of five clones of *STARD7*
- Figure 22: sanger sequencing of *STARD7* in monkeys

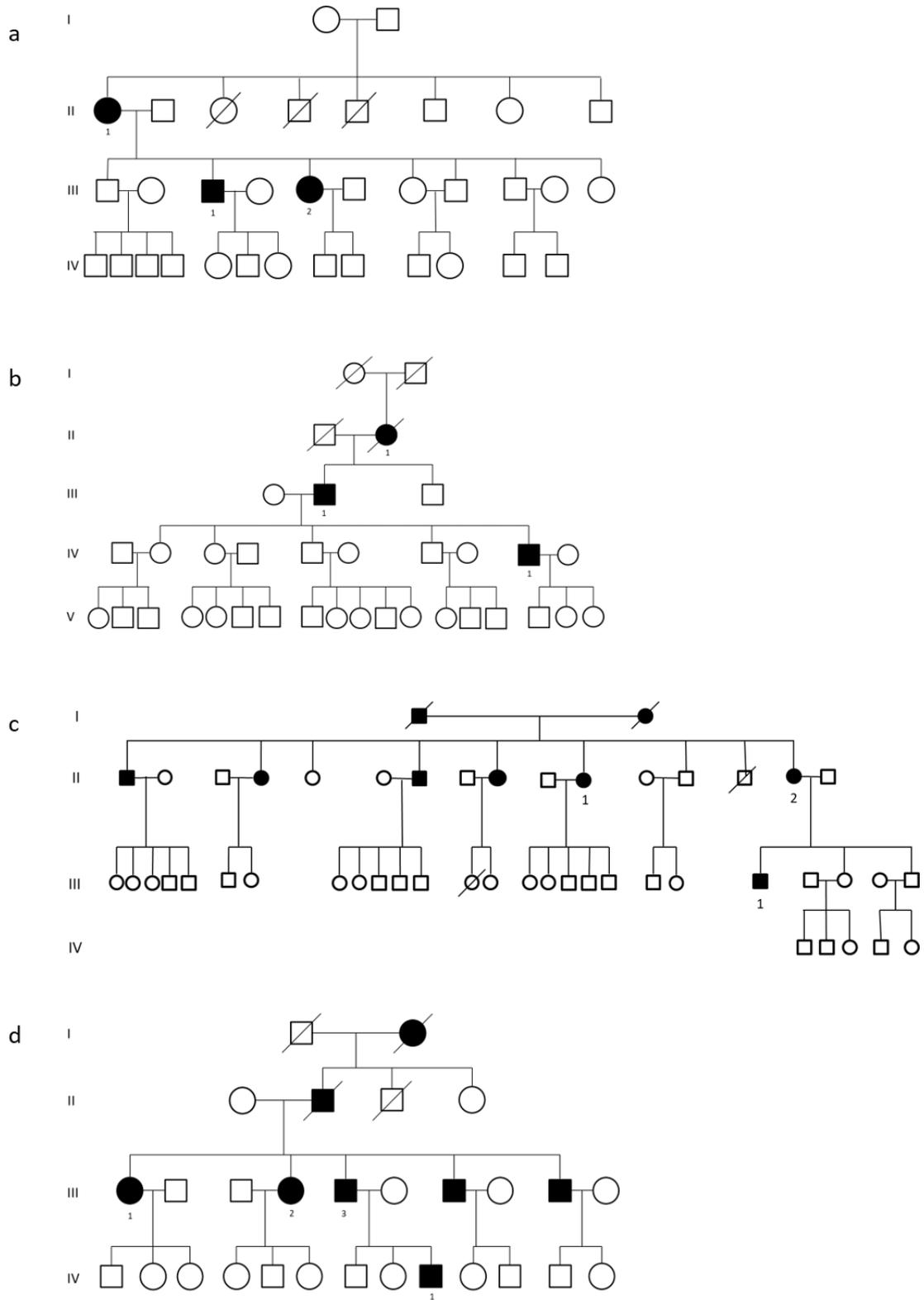
- Supplementary figure 1: Pedigrees of all ET-Families
- Supplementary figure 2: Results of positive RP-PCR results in *MARCHF6* and *STARD7*
- Supplementary figure 3: Visualisation of extracted nanopore reads covering the expansion
- Supplementary figure 4: Distribution, analysis and interpretation of signals detected by molecular combing
- Supplementary figure 5: Summary of all rearrangements observed in individual 2-IV-8 and 2-V-3

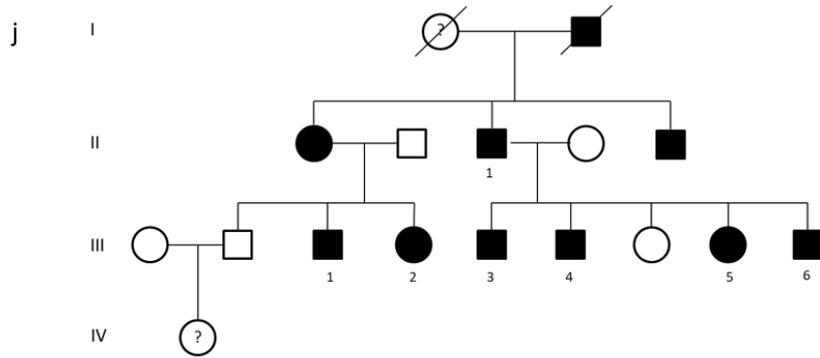
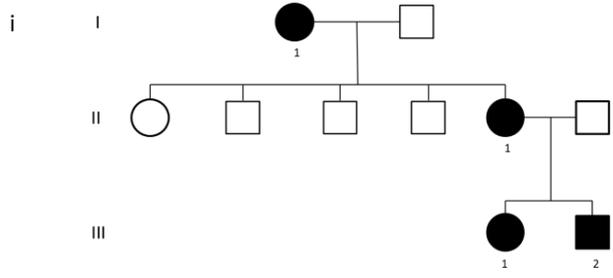
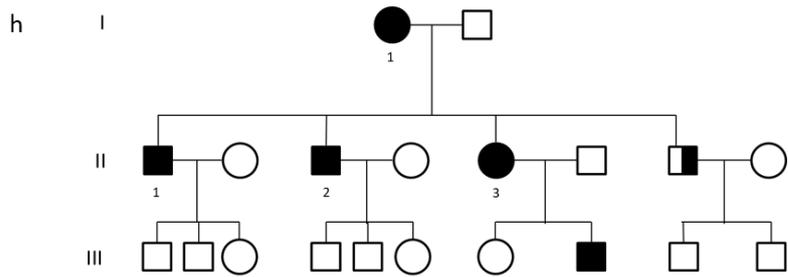
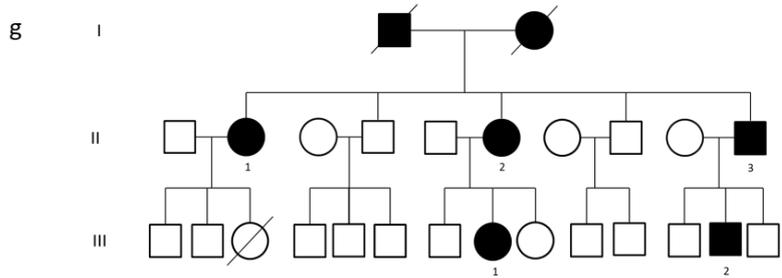
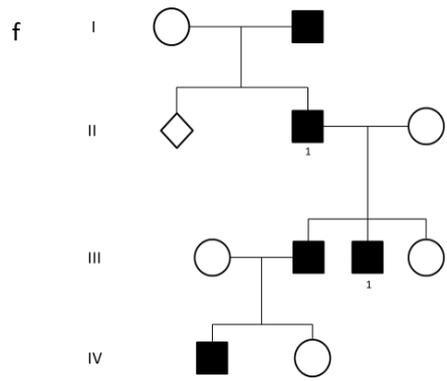
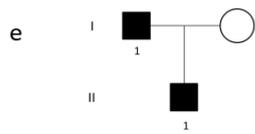
10 Tables

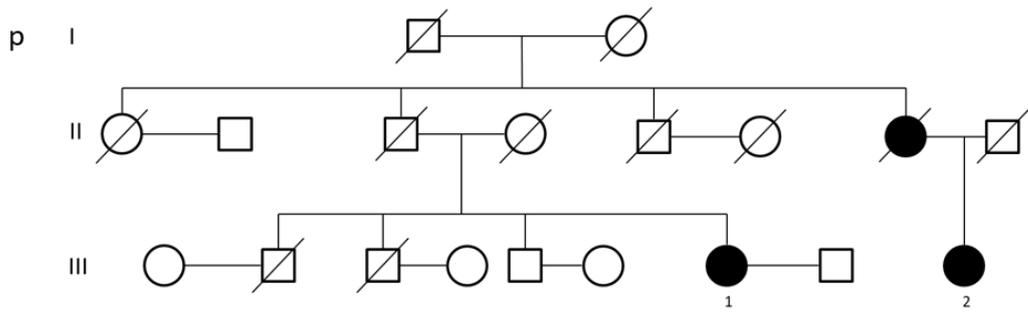
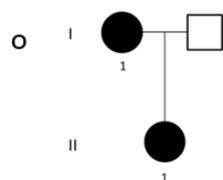
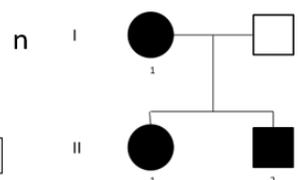
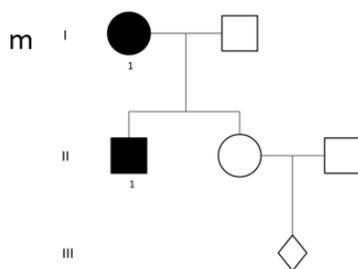
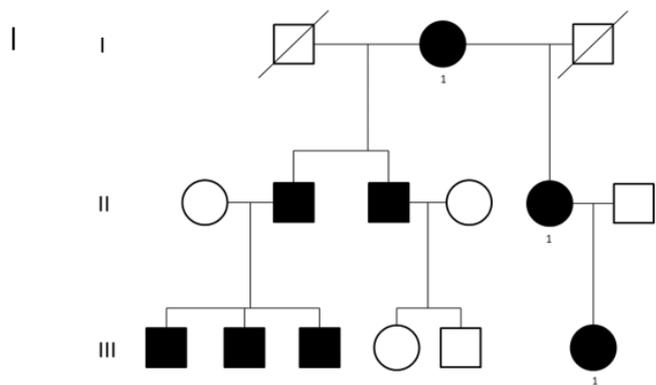
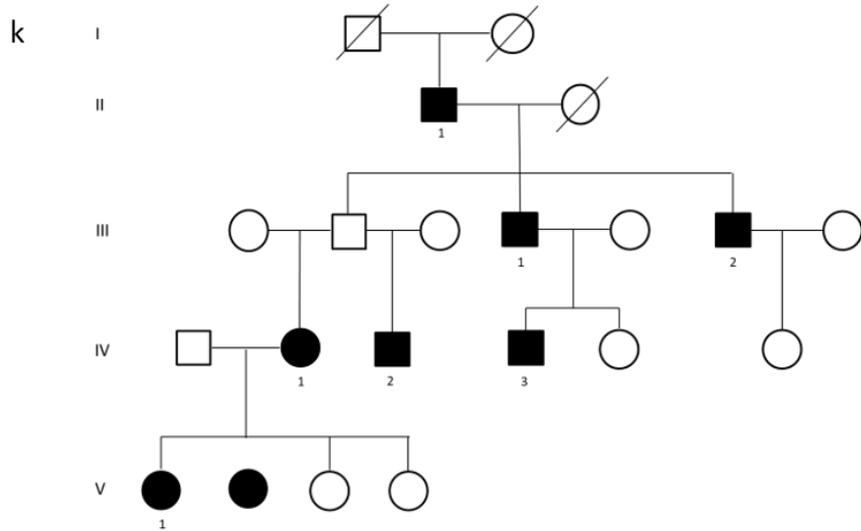
Table 1:	List of all abbreviations and names used for FAME
Table 2:	List of less common clinical features associated with FAME
Table 3:	List of FAME loci
Table 4:	Comparing factors between FAME and ET
Table 5:	Summarized clinical data for family 2
Table 6:	Summarized clinical data for family 1
Table 7:	Summarized clinical features and expansion characteristics for the 10 resamples individuals
Table 8:	Summary of all ET-individuals tested with sanger sequencing in <i>STARD7</i>
Table 9:	Summary of FAME-related genes carrying a TTTCA-expansion with function and location of the expansion
Table 10:	Neurological disorders with intronic pentanucleotide repeat expansions
Supplementary table 1:	List of all affected individuals of family 1 amplified at <i>MARCHF6</i> and family 5 amplified at <i>STARD7</i>
Supplementary table 2:	Individuals with FAME related expansions
Supplementary table 3:	Summary of all ET individuals tested with sanger sequencing in <i>MARCHF6</i>

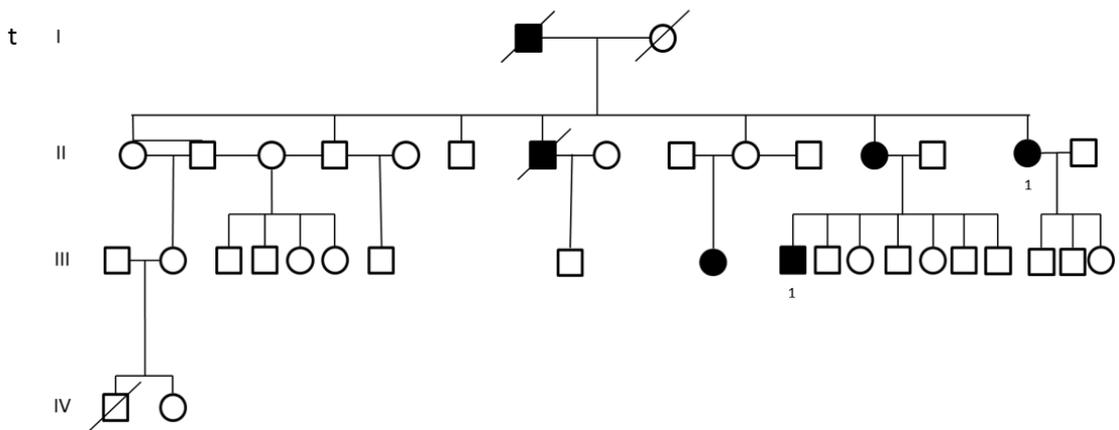
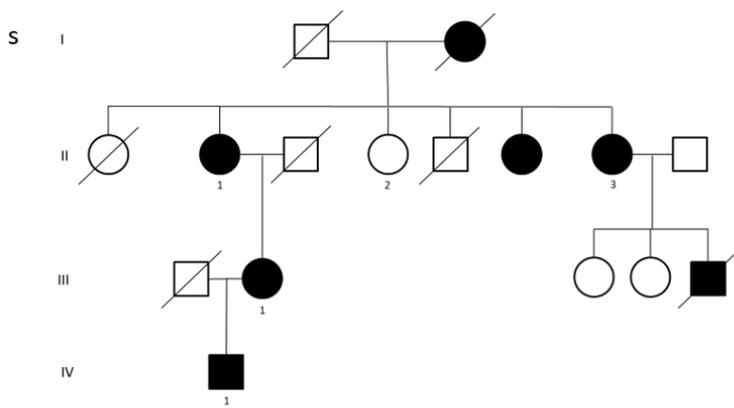
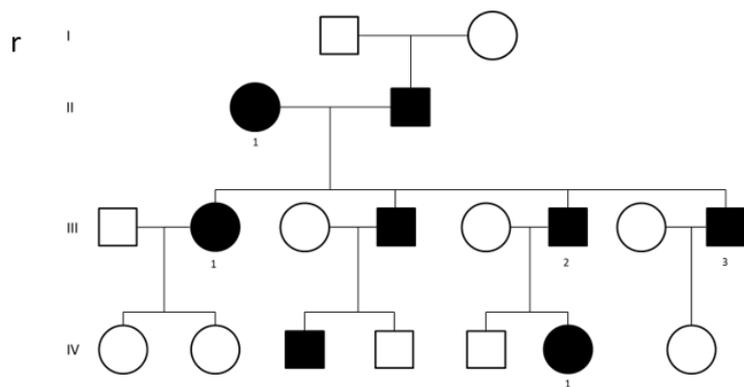
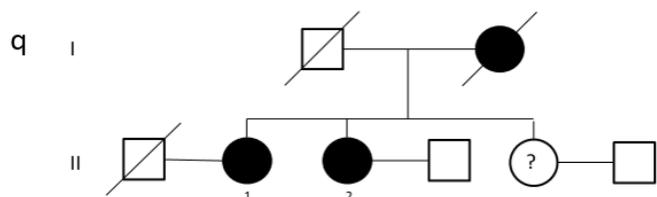
11 Supplementary Data

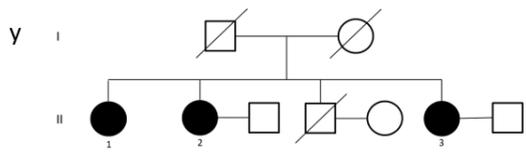
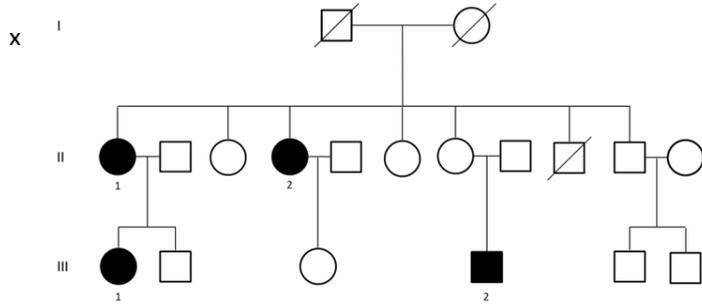
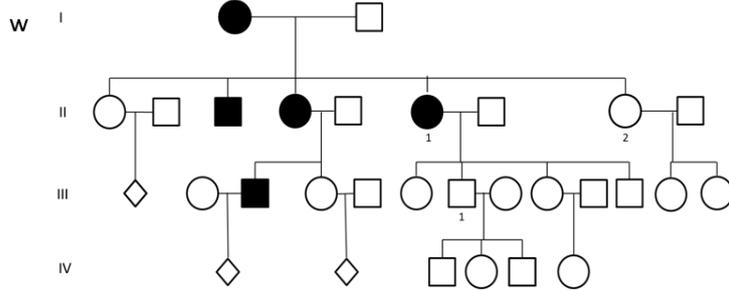
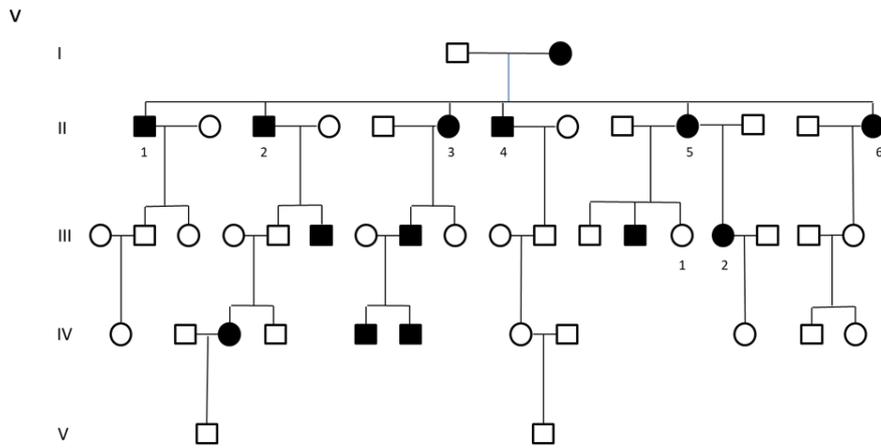
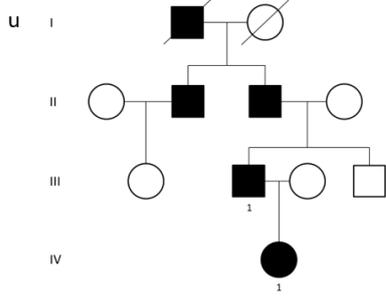
11.1 Supplementary Figures

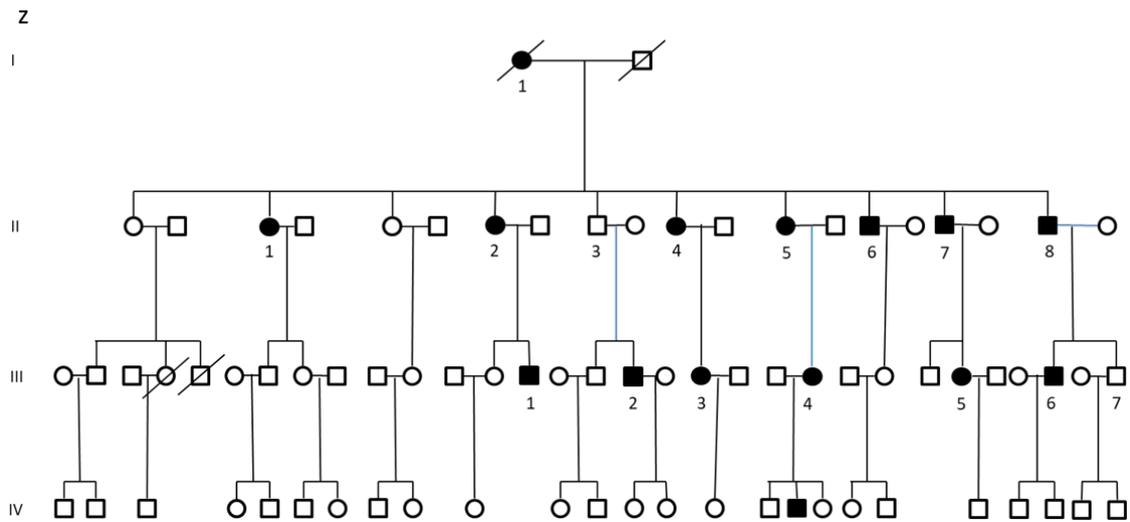




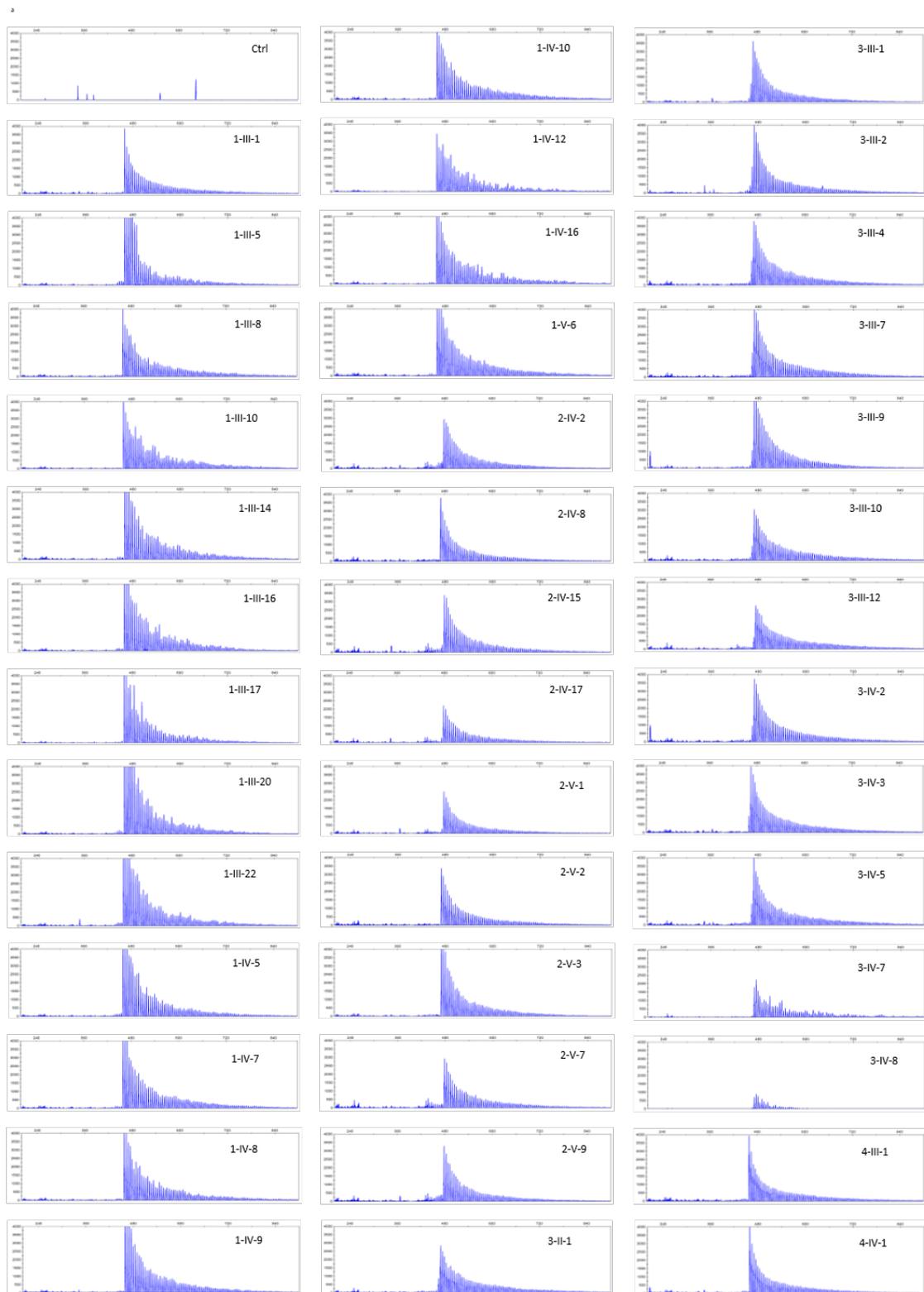




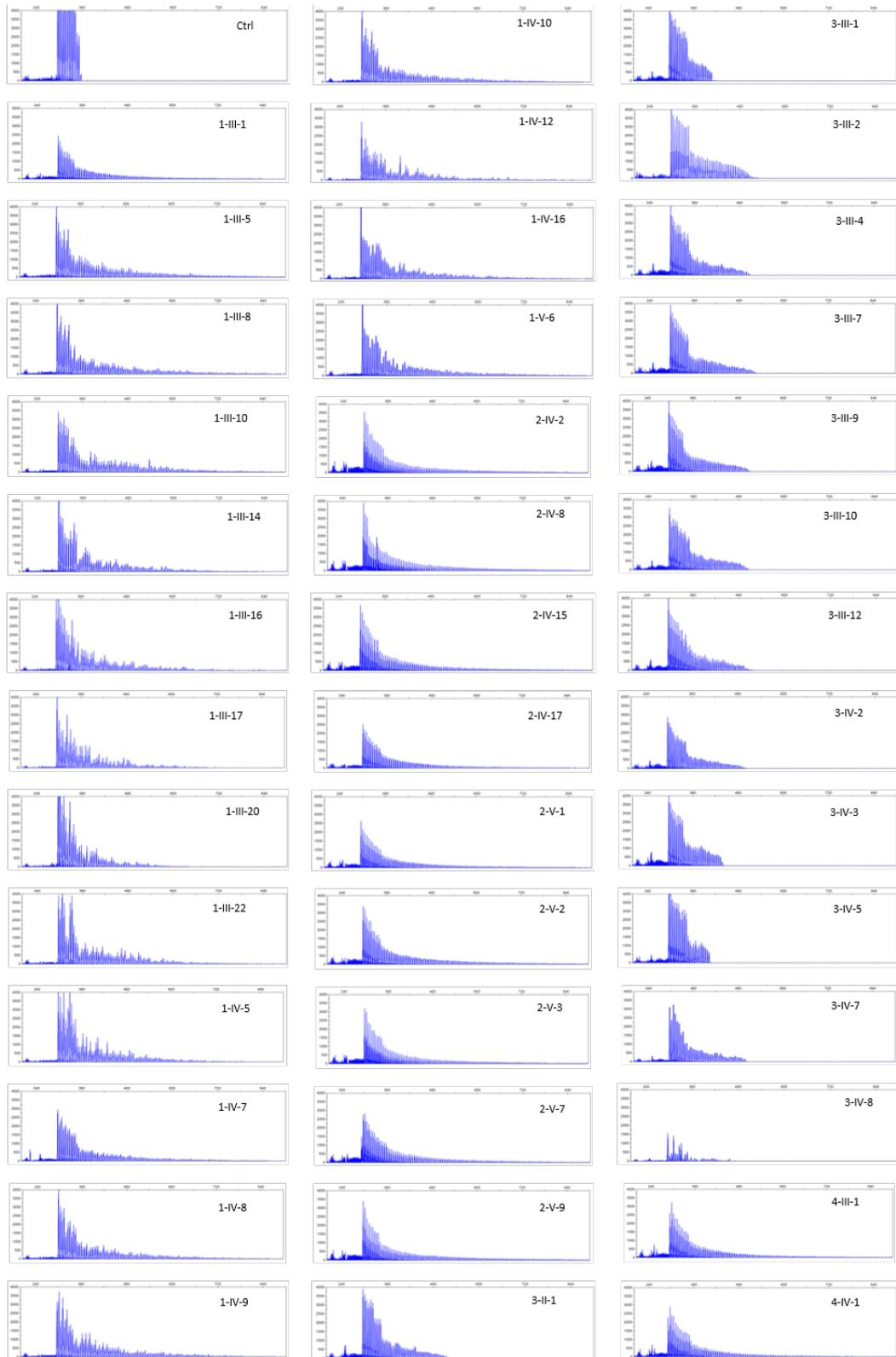




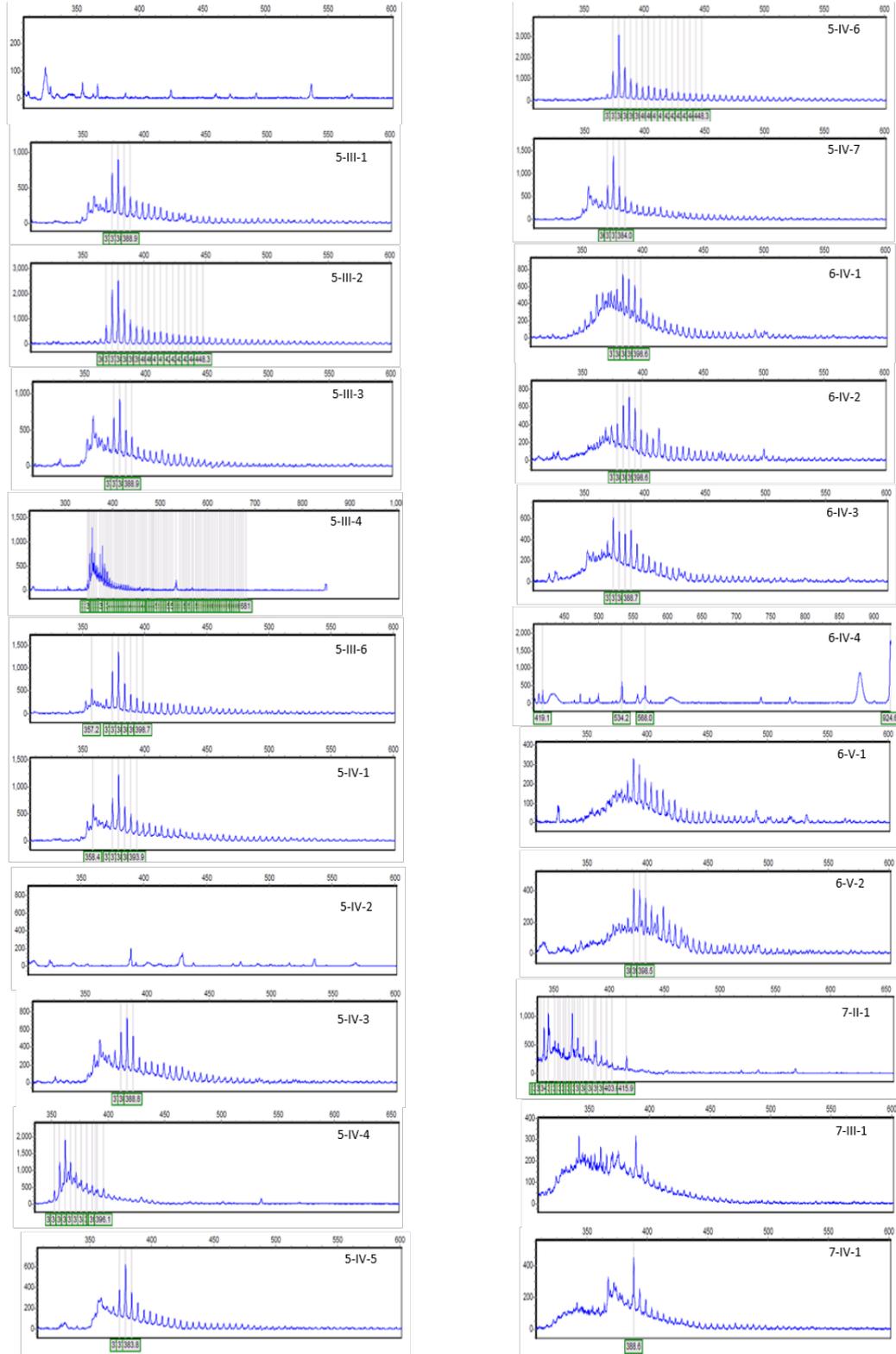
Supplementary Fig. 1. Pedigrees of all ET-families. a) Family 14, b) Family 16, c) Family 17, d) Family 18, e) Family 19, f) Family 20, g) Family 21, h) Family 22, i) Family 24, j) Family 25, k) Family 26, l) Family 27, m) Family 28, n) Family 29, o) Family 30, p) Family 31, q) Family 32, r) Family 33, s) Family 34, t) Family 35, u) Family 36, v) Family 37, w) Family 38, x) Family 39, y) Family 40, z) Family 41
 Family 13, 14 and 23 did not have available pedigrees. Affected individuals are marked black.

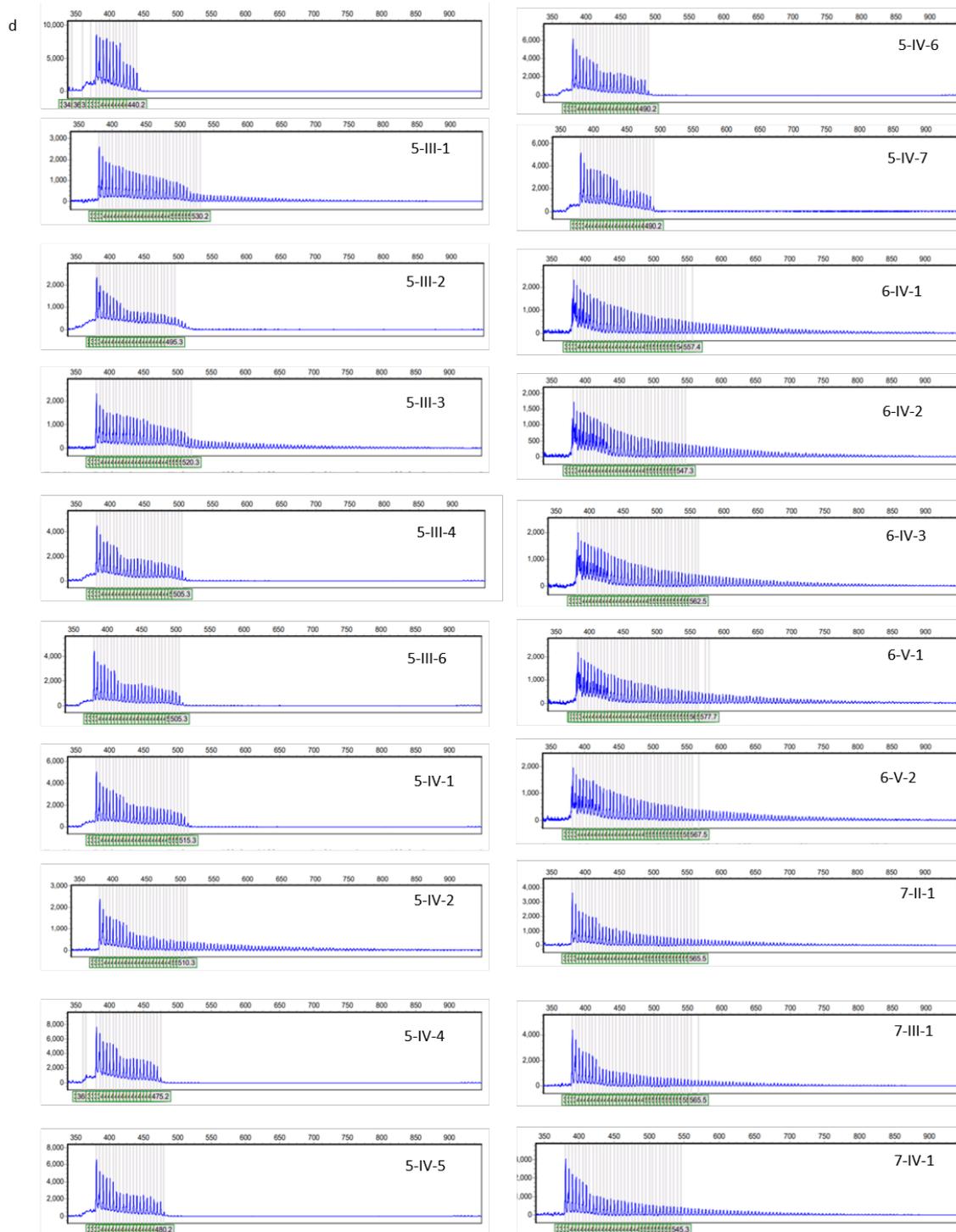


b

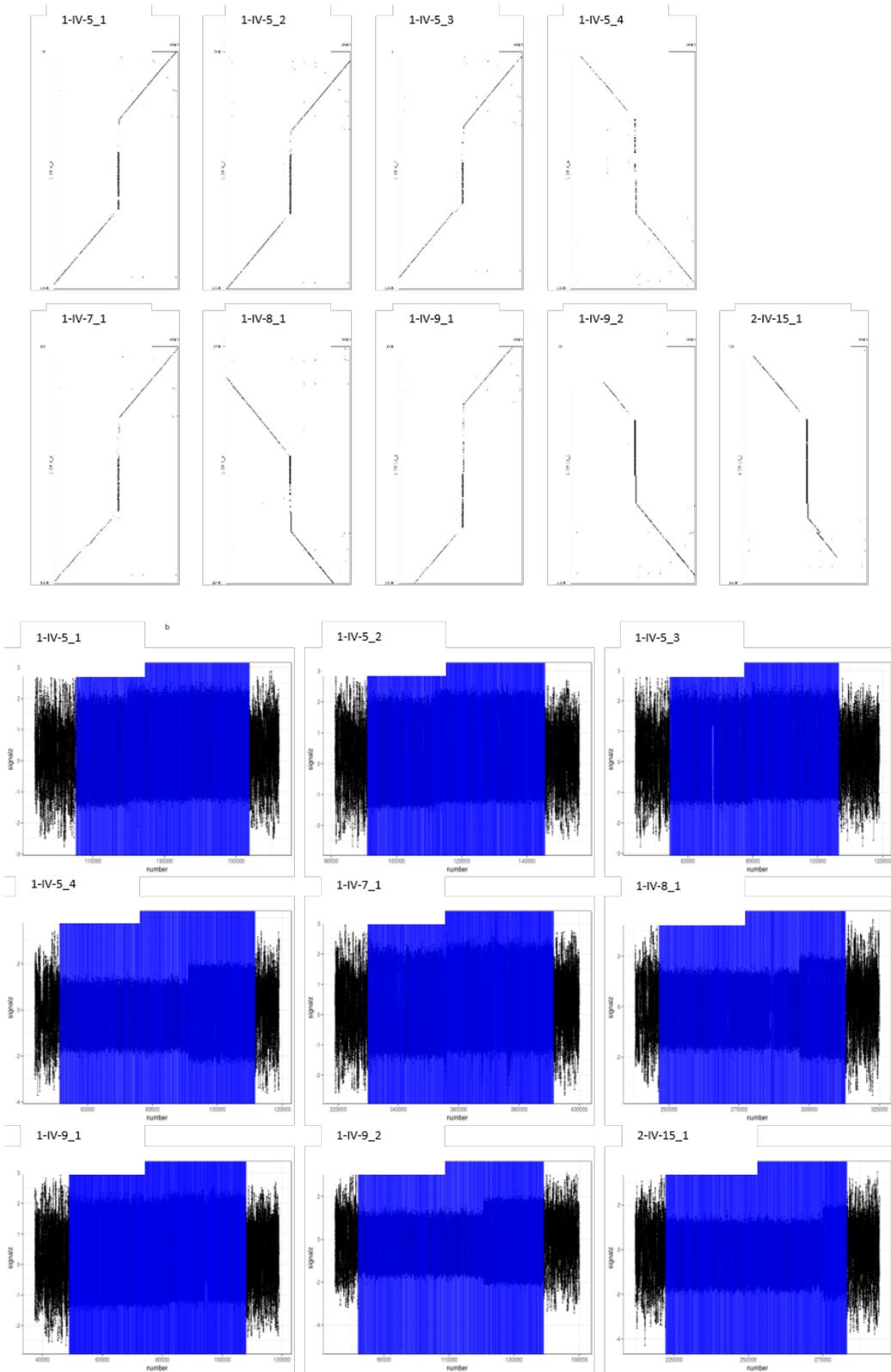


C

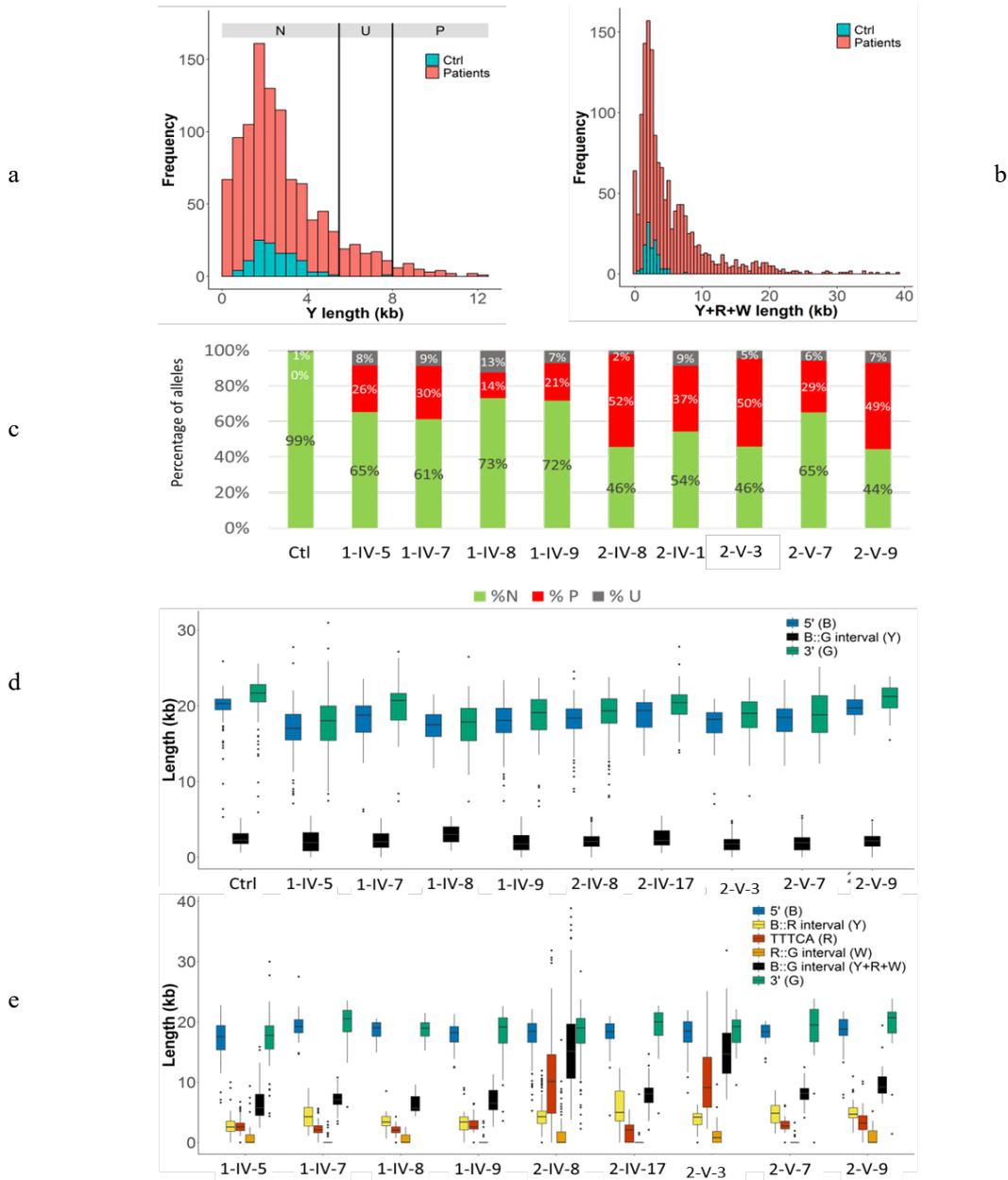




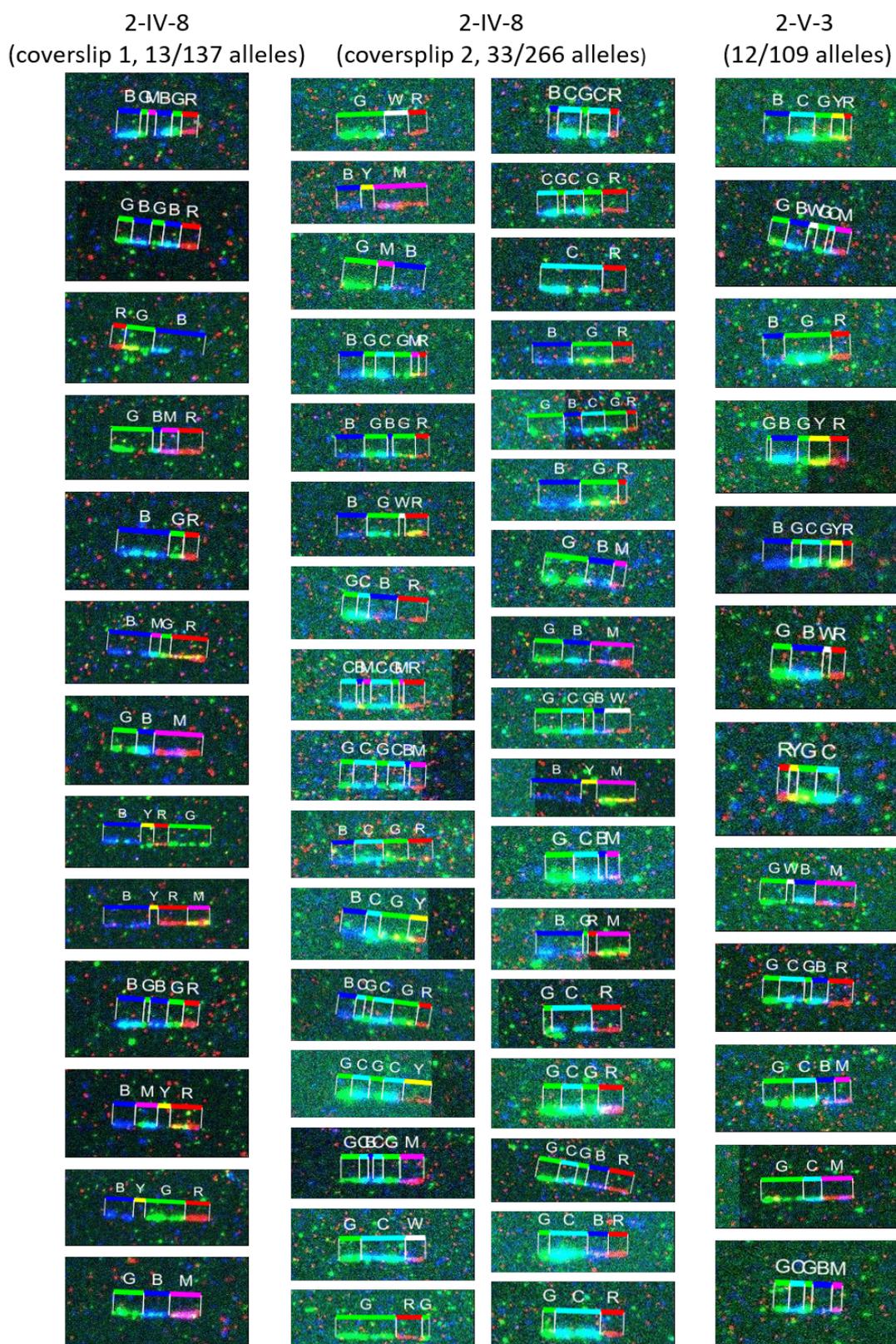
Supplementary Fig. 2. Results of positive 3'-TTTCA (a) and 5'-AAAAT (b) RP-PCR assays in *MARCHF6* corresponding to Families 1-4 and Results positive 3'-TTTCA (c) and 5'-AAAAT (d) RP-PCR assays in *STARD7* corresponding to Families 5-7.



Supplementary Fig. 3. Visualization of extracted nanopore reads covering the expansion. **a)** Dot plots comparing nanopore reads displaying the expansion (Y-axis, scale: 13 kb) with the corresponding Hg19 reference region (X-axis, scale: kb). The expansion appears as a vertical line. **b)** Analysis of raw nanopore reads using NanoSatellite. The signals corresponding to the expanded repeats appear in blue.



Supplementary Fig. 4. Distribution, analysis and interpretation of signals detected by molecular combing. **a)** Distribution of the length of the unstained part between the blue (B) and green (G) signals for alleles (Y) without red (R) staining in a control individual and in patients. Alleles with $Y < 5.5$ were interpreted as normal (N); alleles with $5.5 \leq Y < 8.5$ were interpreted as undefined (U); alleles with $Y \geq 8.5$ were interpreted as likely pathogenic (P) based on the distribution of Y observed in the control individuals. **b)** Distribution of the length of the unstained part between the blue and green signals (Y+R+W) for all alleles in control individuals and in patients. **c)** Percentage of normal (N), undefined (U) and pathogenic (P) alleles in the control individual (Ctrl) and the nine patients analyzed by molecular combing. P includes both likely pathogenic alleles (no red staining but $Y \geq 8.5$) and definite pathogenic (with red staining) alleles. These graphs show that for the individuals with the smallest expansions, the percentage of pathogenic alleles is lower than expected by chance, which suggests a higher overlap between normal and pathogenic alleles in these individuals. **d)** Box plots showing the size distributions of each part of the signals (B: blue, G: green) and the interval between them (Y) for alleles interpreted as normal (N). **e)** Box plots showing the distribution of each part of the signals for alleles interpreted as pathogenic (P): B (blue): 5' flanking region; R (red): TTTCA; G (green): 3' flanking region; Y (yellow): distance from B to R, interpreted as 5'-TTTTA; W (orange): distance from R to G, interpreted as 3'-TTTTA; Y+R+W (black): distance from B to G.



Supplementary Fig. 5. Summary of all rearrangements observed in individuals 2-IV-8 (46/403 1409 alleles counted on two different coverslips) and 2-V-3 (12/109 alleles, 1 coverslip).

11.2 Supplementary Tables

Patient ID	Clinically affected	Alleles <i>MARCHF6</i>	Alleles <i>STARD 7</i>
1-III 1	+	na	
1-III 5	+	721, exp?	
1-III 8	+	721, exp?	
1-III 10	+	721, exp?	
1-III 14	+	731, exp?	
1-III 16	+	736, exp?	
1-III 17	+	736, exp?	
1-III 20	+	736, exp?	
1-III 22	+	731, exp?	
1-IV 5	+	731, exp?	
1-IV 7	+	736, exp?	
1-IV 8	+	736, exp?	
1-IV 9	+	731, exp?	
1-IV 10	+	731, exp?	
1-IV 12	+	741, exp?	
1-IV 16	+	746, exp?	
5-III 1	+		na
5-III 2	+		709, exp?
5-III 3	+		na
5-III 4	+		704, exp?
5-III 6	+		704, exp?
5-IV 1	+		714, exp?
5-IV 3	+		704, exp?
5-IV 4	+		704, exp?
5-IV 5	+		na
5-IV 7	+		729, exp?

Supplementary Table 1. List of affected individuals from family 1 amplified at *MARCHF6* and family 5 amplified at *STARD7*. All individuals tested appear to be homozygous with the unaffected “normal” allele.

exp? = expanded allele?, na = not available

a

Pedigree	Clinically Affected	3'-End	3'-End	5'-End
		TTTCA	TTTTA	TTTTA
Family 1				
1-III 1	Affected	+		+
1-III 2	Not-Affected	-		-
1-III 3	Not-Affected	-	-	-
1-III 4	Not-Affected	-	-	-
1-III 5	Affected	+	+	+
1-III 6	Not-Affected	-	-	
1-III 7	Not-Affected	-	-	-
1-III 8	Affected	+	-	+
1-III 9	Not-Affected	-	-	-
1-III 10	Affected	+	-	+
1-III 11	Not-Affected	-	-	-
1-III 12	Not-Affected	-	-	-
1-III 13	Not-Affected	-	-	-
1-III 14	Affected	+	-	+
1-III 15	Not-Affected	-	-	
1-III 16	Affected	+	-	+
1-III 17	Affected	+	-	+
1-III 18	Not-Affected	-	-	
1-III 20	Affected	+	-	+
1-III 21	Not-Affected	-	-	-
1-III 22	Affected	+	-	+
1-IV 1	Not-Affected	-	-	
1-IV 2	Not-Affected	-	-	
1-IV 3	Not-Affected	-	-	-
1-IV 4	Not-Affected	-	-	
1-IV 5	Affected	+	-	+
1-IV 6	Not-Affected	-	-	-
1-IV 7	Affected	+	-	+
1-IV 8	Affected	+	-	+

1-IV 9	Affected	+	-	+
1-IV 10	Affected	+	-	+
1-IV 11	Not-Affected	-	-	-
1-IV 12	Affected	+	-	+
1-IV 13	Not-Affected	-	-	-
1-IV 14	Not-Affected	-	-	-
1-IV 15	Not-Affected	-	-	-
1-IV 16	Affected	+	-	+
1-V 1	Not-Affected	-	-	-
1-V 2	Not-Affected	-	-	-
1-V 3	Not-Affected	-	-	-
1-V 4	Not-Affected	-	-	-
1-V 5	Not-Affected	-	-	-
1-V 6	Not-Affected	+	-	+
Family 2				
2-IV 1	Not-Affected	-	-	-
2-IV 2	Affected	+	-	+
2-IV 3	Not-Affected	-	-	
2-IV 4	Not-Affected		-	-
2-IV 5	Not-Affected		-	
2-IV 6	Not-Affected	-	-	-
2-IV 7	Not-Affected	-	-	-
2-IV 8	Affected	+	-	+
2-IV 9	Not-Affected		-	-
2-IV 10	Not-Affected	-	-	-
2-IV 11	Not-Affected	-	-	-
2-IV 12	Unclear	-	-	-
2-IV 13	Unclear	-	-	-
2-IV 14	Not-Affected	-	-	-
2-IV 15	Affected	+	-	+
2-IV 16	Not-Affected		-	
2-IV 17	Affected	+	-	+
2-IV 18	Not-Affected	-	-	
2-V 1	Affected	+	-	+

2-V 2	Affected	+	-	+
2-V 3	Affected	+	-	+
2-V 4	Not-Affected		-	
2-V 5	Not-Affected		-	-
2-V 6	Not-Affected	-	-	-
2-V 7	Affected	+	-	+
2-V 8	Not-Affected	-	-	-
2-V 9	Affected	+	-	+
Family 3				
3-II 1	Affected	+		+
3-II 2	Not-Affected	-	-	-
3-II 3	Not-Affected	-	-	-
3-II 4	Not-Affected	-	-	
3-II 5	Not-Affected	-	-	-
3-III 1	Affected	+		+
3-III 2	Affected	+		+
3-III 3	Not-Affected	-	-	-
3-III 4	Affected	+		+
3-III 5	Not-Affected	-	-	-
3-III 6	Not-Affected	-	-	
3-III 7	Affected	+		+
3-III 8	Not-Affected	-	-	-
3-III 9	Affected	+		+
3-III 10	Affected	+		+
3-III 11	Not-Affected	-	-	
3-III 12	Affected	+		+
3-IV 1	Not-Affected	-	-	-
3-IV 2	Affected	+		+
3-IV 3	Affected	+	-	+
3-IV 4	Not-Affected	-	-	-
3-IV 5	Affected	+		+
3-IV 6	Not-Affected	-	-	
3-IV 7	Affected	+	-	+
3-IV 8	Affected	+		+

Family 4				
4-III 1	Affected	+	-	+
4-IV 1	Affected	+	-	+

B

Pedigree	Clinically Affected	3'-End TTTCA	3'-End TTTGA	5'-End TTTGA
Family 5				
5-III 1	Affected	+	+	+
5-III 2	Affected	+	-	+
5-III 3	Affected	+	+	+
5-III 4	Affected	+	-	+
5-III 5	Not-Affected	-	-	
5-III 6	Affected	+	-	+
5-III 7	Not-Affected	-	-	-
5-III 8	Not-Affected	-	-	-
5-IV 1	Affected	+	-	+
5-IV 2	Not-Affected	-	+	+
5-IV 3	Affected	+	-	-
5-IV 4	Affected	+	-	+
5-IV 5	Affected	+	-	+
5-IV 6	Not-Affected	+	-	+
5-IV 7	Affected	+	-	+
Family 6				
6-IV 1	Affected	+	+	+
6-IV 2	Affected	+	+	+
6-IV 3	Affected	+	+	+
6-IV 4	Affected	-	-	-
6-V 1	Affected	+	+	+
6-V 2	Affected	+	+	+
Family 7				
7-II 1	Affected	+	-	+
7-III 1	Affected	+	-	+
7-IV 1	Affected	+	-	+

Supplementary Table 2. Table of individuals with FAME-related expansion. Family 1-4 positive for expansions in intron 1 of *MARCHF6* (a) and family 5-7 positive for expansions in intron 1 of *STARD7*. Individuals marked in green were tested positive for 3'-TTTTA_{exp} and 3'TTTC_{exp}. Individuals marked in red were not (yet) symptomatic but carried the expansions. Individuals marked in yellow were reportedly affected with a tremor-like disorder but did not carry the expansions and individuals marked in blue showed tremor in higher age during the clinical trial but were previously reported negative and did not carry the expansion.

Individual ID	Alleles	Sanger Sequencing
13-2	736	(TTTTA) ₁₄
13-3	741	(TTTTA) ₁₅
13-4	736	(TTTTA) ₁₄
13-5	736	(TTTTA) ₁₄
22-II-2	(707); 735	(TTTTA) ₁₄
22-II-3	(707); 735	(TTTTA) ₁₄
24-II-1	(705); 735	(TTTTA) ₁₄
24-III-1	(705); 735	(TTTTA) ₁₄
24-III-2	(705); 735	(TTTTA) ₁₄
27-I-1	(705); 735	(TTTTA) ₁₄
27-II-1	(705); 735	(TTTTA) ₁₄
27-III-1	(705); 735	(TTTTA) ₁₄
29-I-1	(701); 735	(TTTTA) ₁₄
31-III-1	(712); 740	(TTTTA) ₁₅

Supplementary Table 3. Summary of all ET-individuals tested with sanger sequencing in *MARCHF6*. All carried an unexpanded TTTTA-repeat

12 Acknowledgements

My special thanks belong to Prof. Dr. Christel Depienne. She has been the best guidance and overall the kindest and most patient support I could have wished for through this journey. I would like to thank you for sharing some of your knowledge with me, for always creating space for discussion and for making me grow intellectually and as a person. Your enthusiasm and energy are contagious. I would also like to thank you for giving me the opportunity to travel to France myself to gather clinical data and get a deeper insight into the people behind the FAME-disease.

For teaching me everything I needed to know in and around the lab and for being a constant support even on “bad-lab-days” I would like to thank Sabine Kaya. You have been an uncomplicated source for all things needed and I am very thankful for having you.

The Institute of human genetics has been a great working-environment for me and I would like to thank everyone I was able to get to know for inviting me so friendly and for the constant support present in the lab. I want to thank Theresa and Julia especially for moral support and the laughter shared.

I would also like to thank our collaborators and the FAME-consortium for supporting the project and for taking part in the process of this thesis. I especially want to thank Dr. Florian Kraft, Prof. Dr. Stephan Klebe, Dr. Mark Corbett, Dr. Eloi Magnin, Dr. Giovanni Stevanin and the Cologne Center for Genomics.

For his patients, love and determination to keep me moving I thank Charly. You have always given me the strength to achieve my dreams and supported me relentlessly without any second guessing. I am forever grateful to have you by my side.

At last, I want to thank my parents for the unquestioned support. Thank you for the advice on the way, the understanding of struggles, the absence of pressure and the freedom of my development. Thank you for having someone to look up to and to still meet eye to eye, I don't think many people can achieve that.

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.