Supplemental Data

Identification of candidate genes for familial early-onset essential tremor

Xinmin Liu, Nora Hernandez, Sergey Kisselev, Aris Floratos, Ashley Sawle, Iuliana Ionita-Laza, Ruth Ottman, Elan D. Louis, and Lorraine N. Clark

Supplemental Material and Methods

Linkage Analysis

SNP Genotyping For Linkage Analysis in 37 ET Families:

Single nucleotide polymorphism (SNP) genotyping was performed using the Illumina Omni Express Array on the Illumina iScan platform in all available family members in 37 ET families. The Illumina Omni Express array contains >715,000 SNPs with 392,197 SNPs located within, 10Kb of Refseq genes, and 15,062 nonsynonymous SNPs. The mean and median spacing of SNPs across the genome is, 4.0Kb and 2.1Kb, respectively. Extensive quality control (QC) steps of SNP data was performed prior to linkage analysis, and included sample call rate (>97%), SNP assay performance including call rate (>98%), Gentrain score (0.25) and Gencall score (>0.15), and concordance for duplicate samples (n=20; 100%). A total of 6 samples had call rates <97% and were removed. A total of, 41,490 SNPs, had call rates <98% and were removed from further analysis. The final number of SNPs used in the analysis was 678,176.

Nonparametric linkage analysis:

MERLIN (version 1.1.2) was used to evaluate individual SNPs for Mendelian inconsistencies based on the pedigree structure, and families showing Mendelian

inconsistencies at a particular SNP were excluded from analysis. SNP marker alleles frequencies were tested for departures from Hardy-Weinberg equilibrium (HWE) using MERLIN. SNPs that showed departure from HWE and were excluded from further analysis. Nonparametric linkage analysis was performed using MERLIN version 1.1.2. In order to minimize the risk of false positives due to LD among SNPs, we pruned the original list of SNPs such that the remaining set of SNPs contained only SNPs with (r^2<0.05). The final number of SNPs used for linkage analysis was 64,262. We calculated two nonparametric linkage statistics, namely the Whittemore and Halpern non-parametric linkage (NPL) pairs (--pairs) and NPL all (--npl) statistics. We performed two sets of analyses: 1) a stringent analysis (phenotype 1, P1) in which only individuals with a diagnosis of definite or probable ET were classified as affected, and individuals with a diagnosis of possible ET were classified as unknown, and 2) a secondary analysis with a broader phenotype (phenotype 2, P2) in which individuals with a diagnosis of definite, probable or possible ET were classified as affected.

Exome Sequencing

Gene and Variant Level Annotation:

The web URLs for all software tools and programs used for gene and variant level annotation are provided in Web Resources. Gene level annotation was performed using the residual variant intolerance score (RVIS). Variant level annotation was performed using the variant annotation tool (VAT), SnpEff and Annovar. Non-synonymous single nucleotide variants were also annotated with *in silico* prediction programs including PROVEAN, Mutation Taster, Polyphen and SIFT. Annotation was also performed using the 'combined annotation dependent depletion score' (CADD). Evolutionary conservation

of variants was assessed with Genomic Evolutionary Rate Profiling (GERP) scores. We examined the function of genes that carry 'candidate' variants predicted to be damaging or deleterious with documented neurodevelopmental, neurobehavioral or neurodegenerative phenotype in humans or in animal models (mouse, C.elegans, Drosophila, Zebrafish). For human annotation, we used published literature in Pubmed together with Genecards, the database for annotation, visualization and integrated discovery (DAVID) v6.7, the human phenotype ontology database, Phenotips, OMIM, HGMD, Orphanet and DECIPHER. For animal models we used published literature, Genecards, mammalian phenotype ontology flybase, wormatlas, and the zebrafish model organism database.

pVAAST Analyses:

We used pVAAST (http://www.yandell-lab.org/software/vaast.html), a disease-gene identification tool designed for high throughput sequence data in pedigrees, to analyze WES data from affected ET cases with a definite, probable or possible diagnosis of ET in 37 early-onset ET families. pVAAST is a tool that combines linkage analysis, case-control association and functional variant prediction in a unified statistical framework that offers much higher power relative to each of the individual methods. pVAAST analysis was performed assuming an autosomal dominant mode of inheritance and using the maximum number of permutations of 1,000,000 in the test. Genes were retained with p-value<0.05 and there were a total of 146 genes with p-value<0.05. Gene candidates are ranked by p-value (based on combined linkage and association evidence, smallest to largest). The p-value and significance value was obtained from permutation and gene-drop simulations. When two genes have the same p-value, the pVAAST score of each gene determines the rank. The pVAAST score (CLRT score from VAAST plus 2ln(10)xLOD) combines variant frequency data with amino acid score and phylo-genetic conservation information using a composite likelihood ratio test (CLRT) and in addition integrates linkage information (quantified by a LOD score) as a separate likelihood ratio in the pVAAST CLRT (CLRTp).

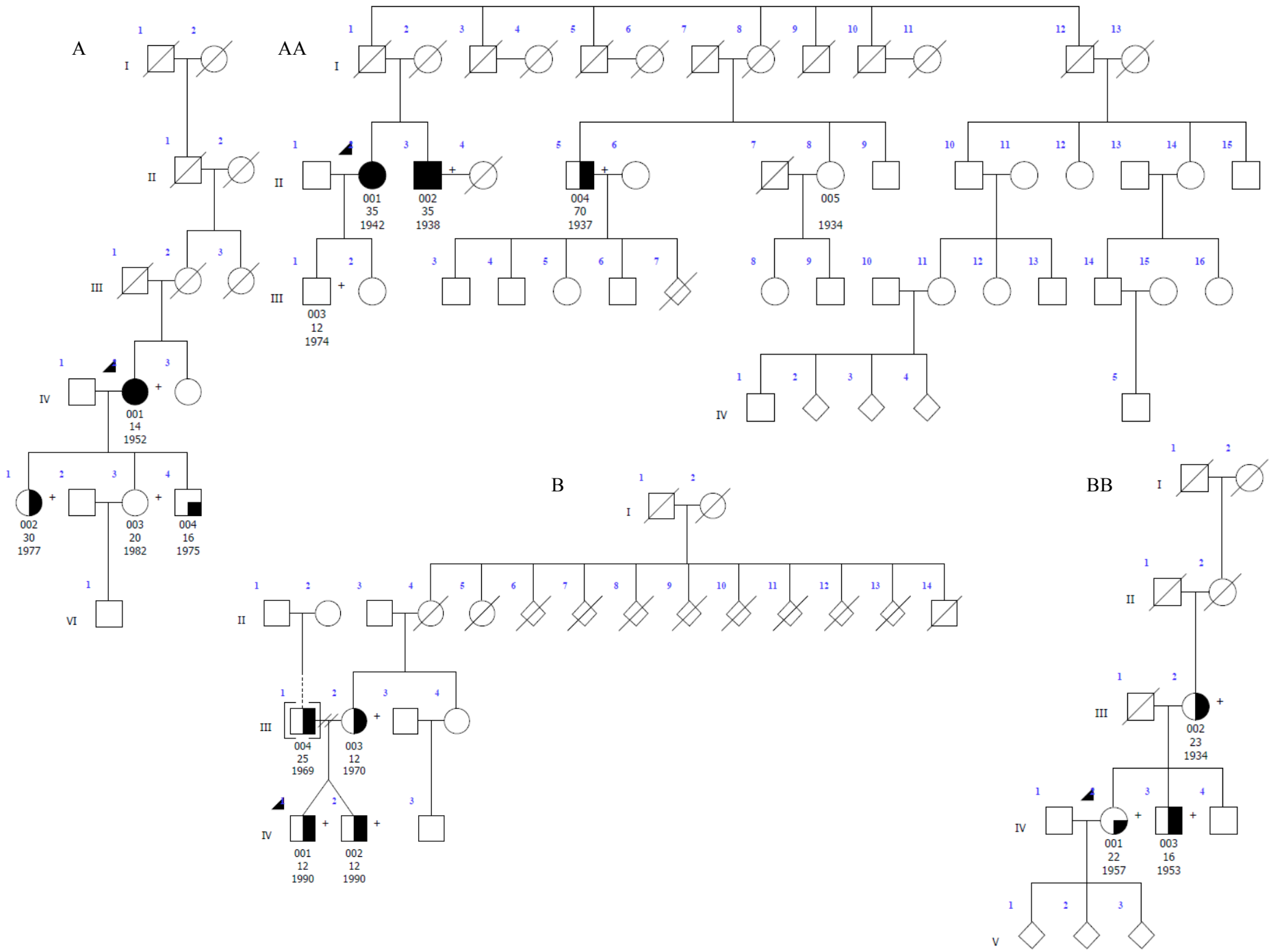
The background ‘control’ genome set used for association analysis consisted of 1,057 exomes from the 1000 genomes project phase I data.

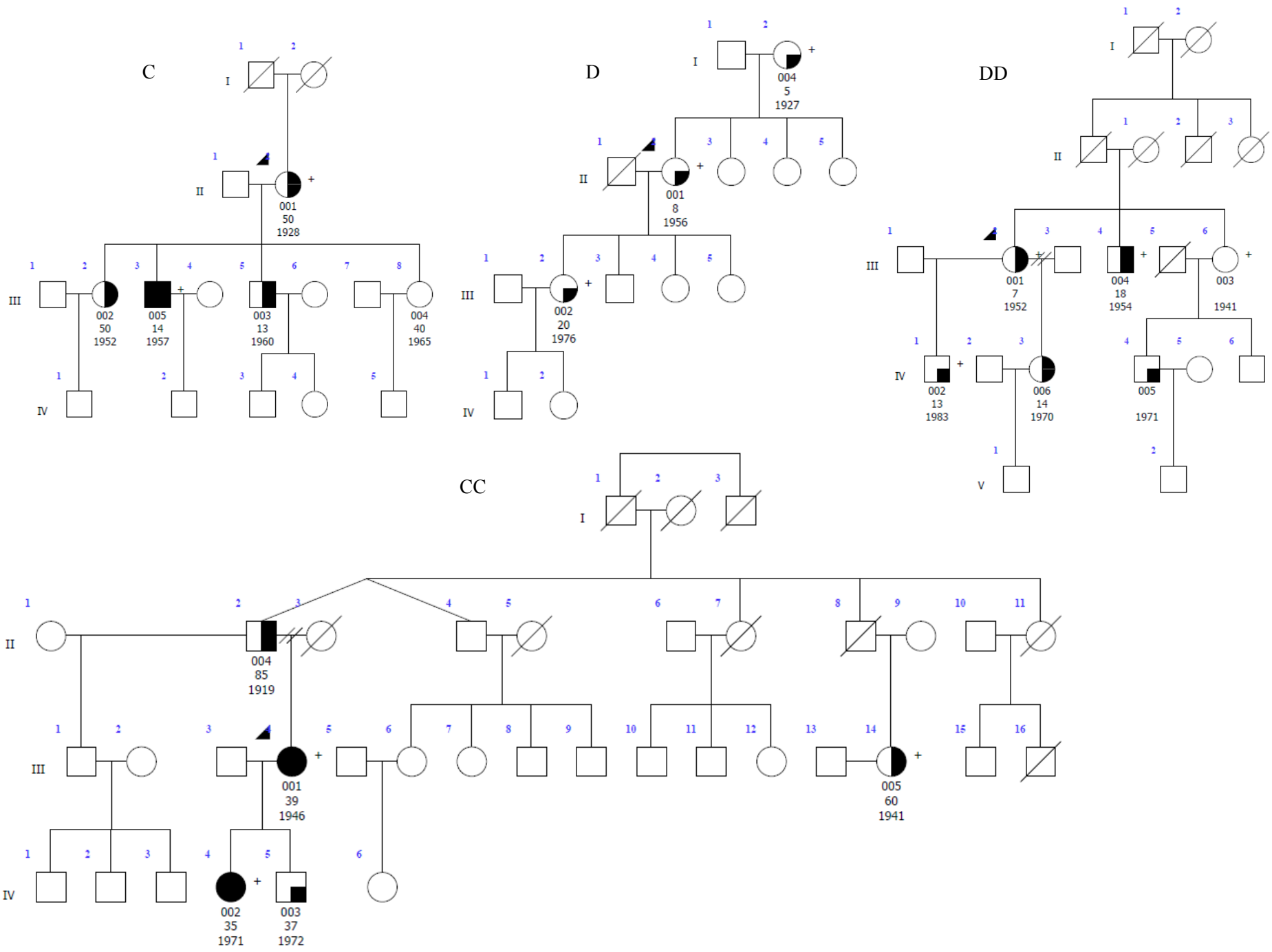
Variant Genotyping and Segregation Analysis:

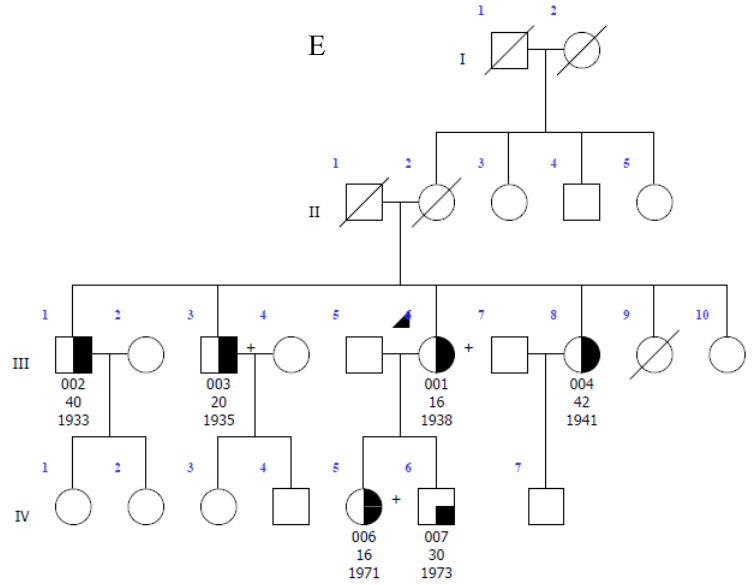
Genotyping of candidate variants in additional family members for segregation analysis was performed using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (Sequenom, San Diego, CA, USA) with Sequenom iPlex Gold custom assays designed using MassARRAY assay design software version 4.0 (Sequenom). Amplification was carried out on Applied Biosystems GeneAmp 9700 thermocyclers (Life Technologies, Carlsbad, CA, USA) using standard recommended cycling conditions for iPlex Gold assays. The mass of extension products was measured and recorded using a mass array compact mass spectrometer (Bruker Daltonik, Billerica, MA, USA). Genotypes were clustered and called using SpectroTYPER software version 2.0 (Sequenom). DNA samples were analyzed in duplicate, and genotype calls were assigned without prior knowledge of the diagnostic status of the samples analyzed. Genotypes that were discordant in replicates (N = 2) were manually called as no calls.

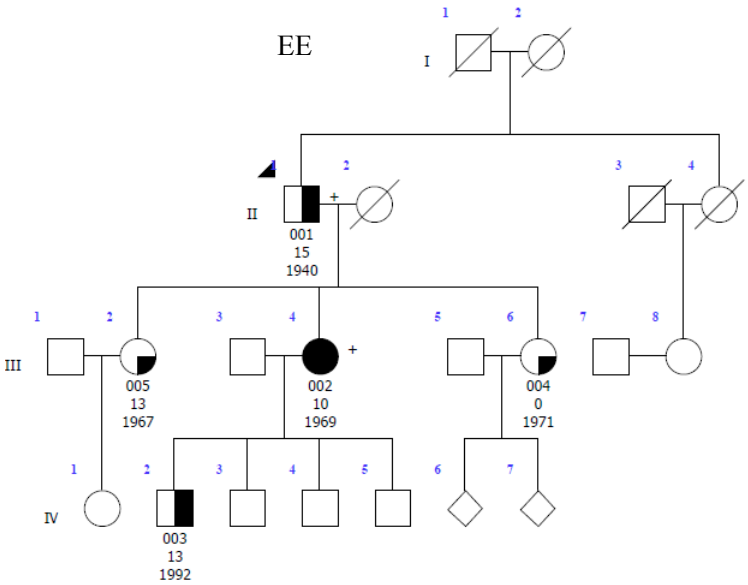
Figure S1 Pedigree Structures of the 37 ET Families included in the linkage and exome analysis

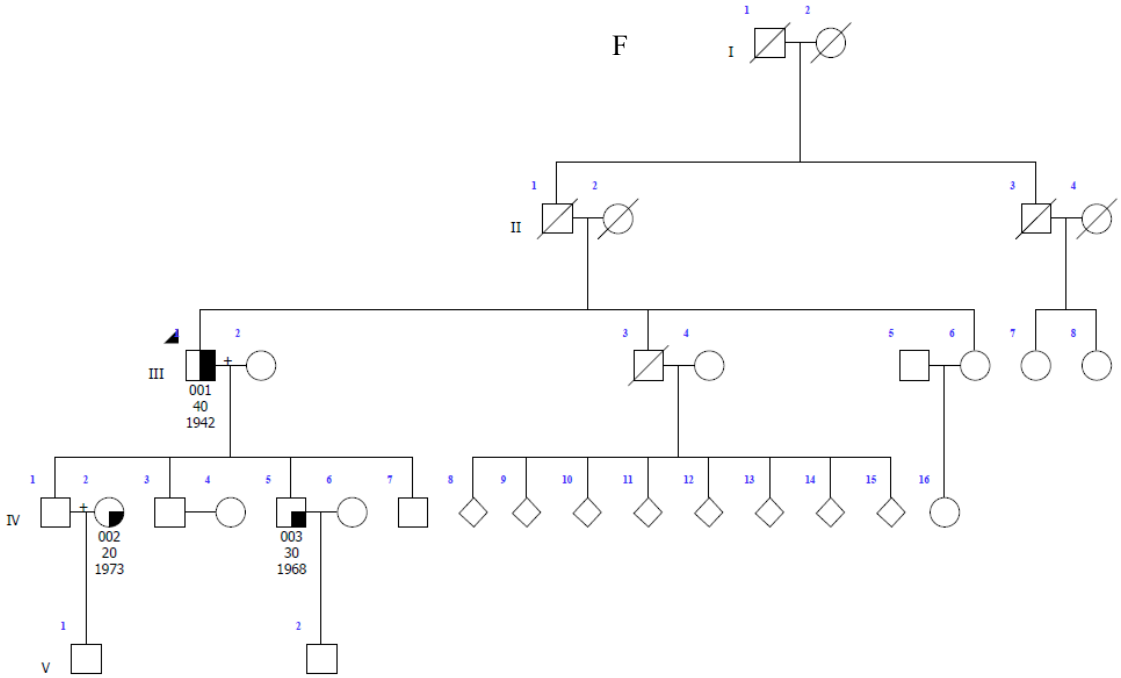
Genetic pedigrees for families with likely pathogenic mutations are shown. The generation in each pedigree is indicated by roman numerals. The proband is indicated by an arrowhead. A '+' symbol indicates subjects that were exome sequenced. Below each subject with DNA available for genetic analysis the subject id (00X), age at tremor onset and date of birth is indicated. Symbol shading is as follows: Definite ET, symbols completely black; Probable ET: symbols half vertical black fill; Possible ET, symbols with a quadrant in black; Unaffected, clear symbol.

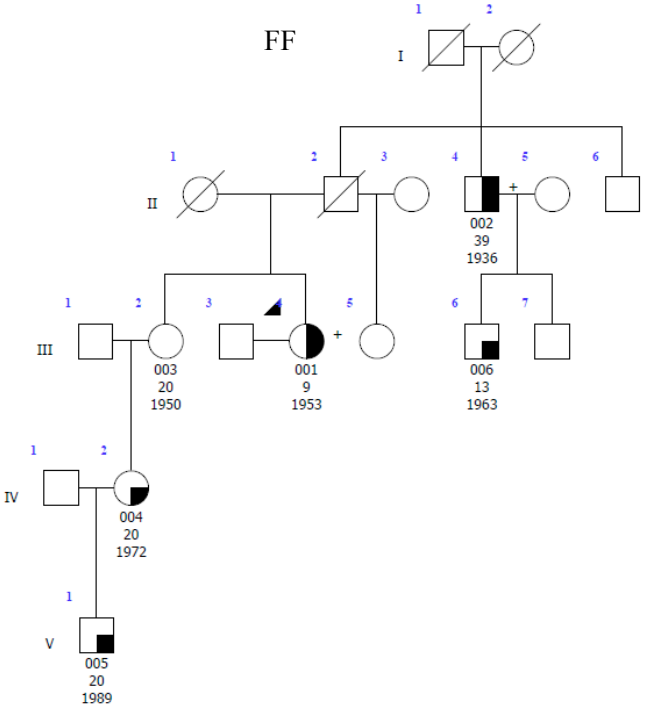

A
I
II
III
IV
001
14
1952
002
30
1977
003
20
1982
004
16
1975
VI
AA
I
II
001
35
1942
002
35
1938
004
70
1937
005
1934
III
003
12
1974
IV
B
I
II
III
004
25
1969
003
12
1970
IV
001
12
1990
002
12
1990
BB
I
II
III
002
23
1934
IV
001
22
1957
003
16
1953
V


C
I
II
001
50
1928
III
002
50
1952
005
14
1957
003
13
1960
004
40
1965
IV
D
I
004
5
1927
II
001
8
1956
III
002
20
1976
IV
DD
I
II
III
001
7
1952
004
18
1954
003
1941
IV
002
13
1983
006
14
1970
005
1971
V
CC
I
II
004
85
1919
III
001
39
1946
005
60
1941
IV
002
35
1971
003
37
1972

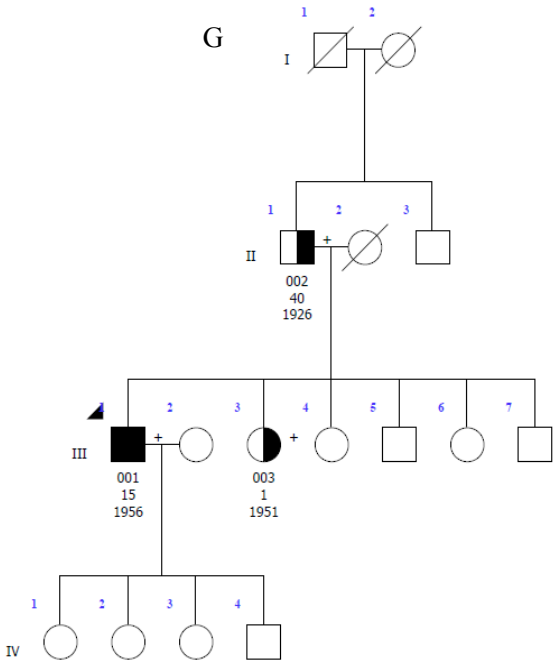

E
I
II
III
002
40
1933
003
20
1935
001
16
1938
004
42
1941
IV
006
16
1971
007
30
1973


EE
I
II
001
15
1940
III
005
13
1967
002
10
1969
004
0
1971
IV
003
13
1992

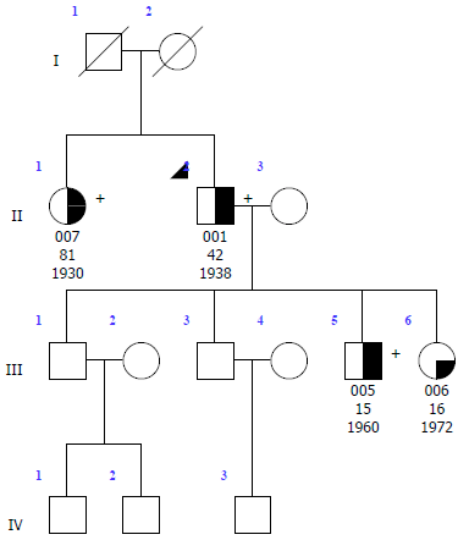

F
I
II
III
001
40
1942
IV
002
20
1973
003
30
1968
V


FF
I
II
002
39
1936
III
003
20
1950
001
9
1953
006
13
1963
IV
004
20
1972
V
005
20
1989

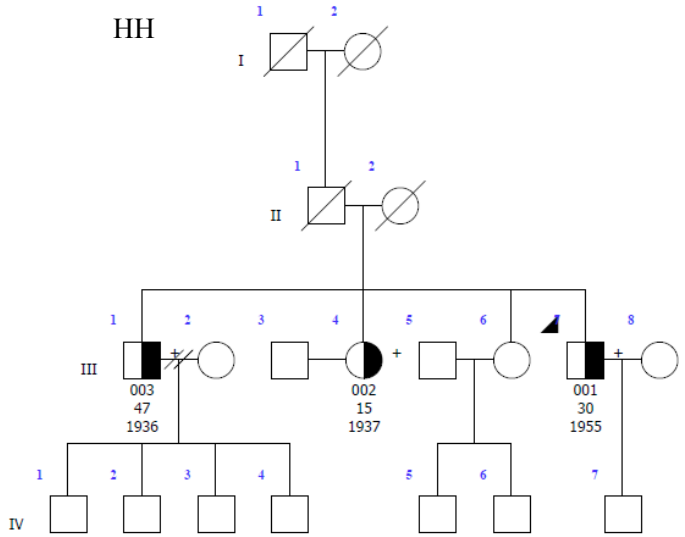
G


I
II
002
40
1926
III
001
15
1956
003
1
1951
IV

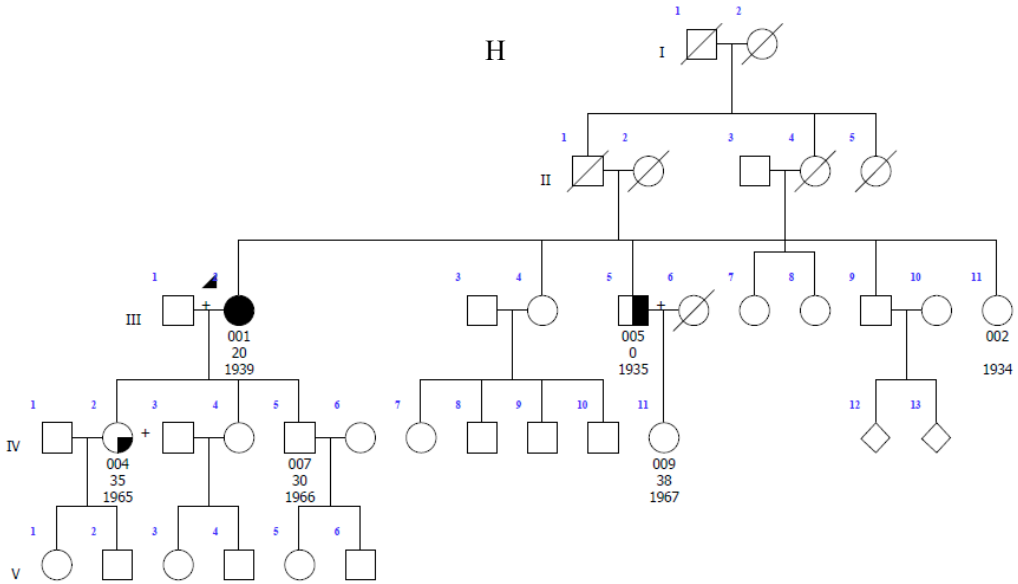
GG


I
II
007
81
1930
001
42
1938
III
005
15
1960
006
16
1972
IV

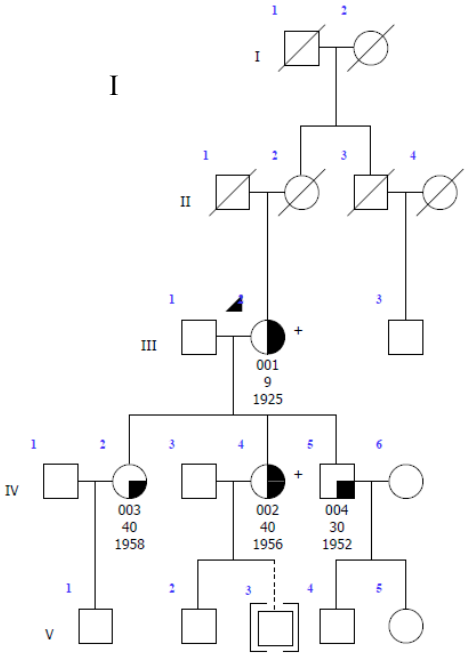
HH

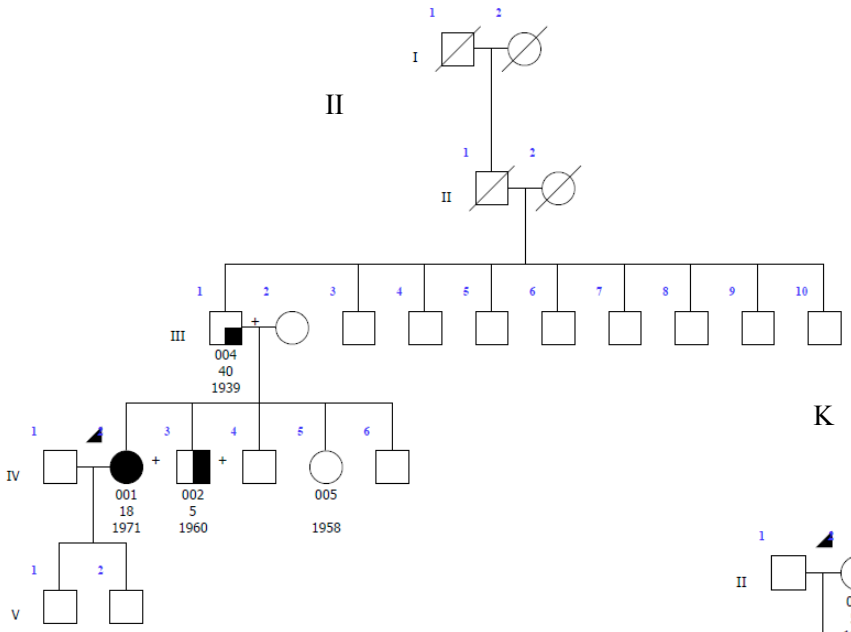

I
II
III
003
47
1936
002
15
1937
001
30
1955
IV

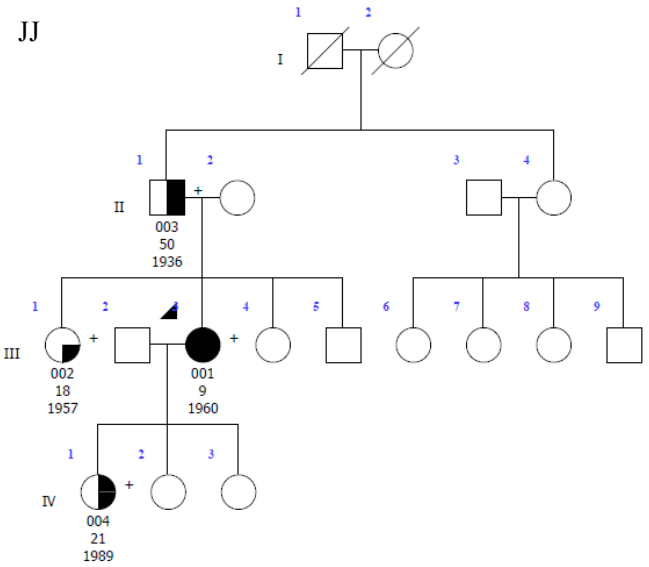
H

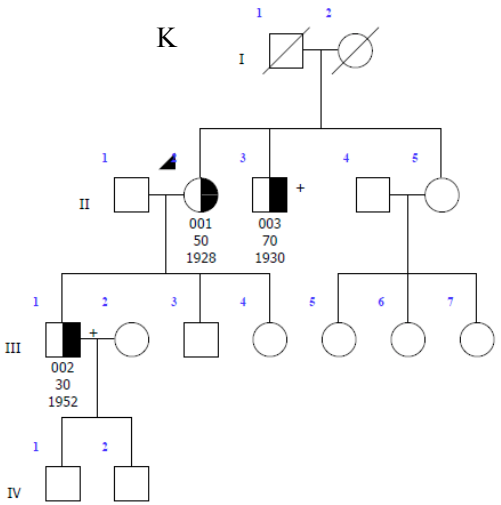

I
II
III
001
20
1939
005
0
1935
002
1934
IV
004
35
1965
007
30
1966
009
38
1967
V

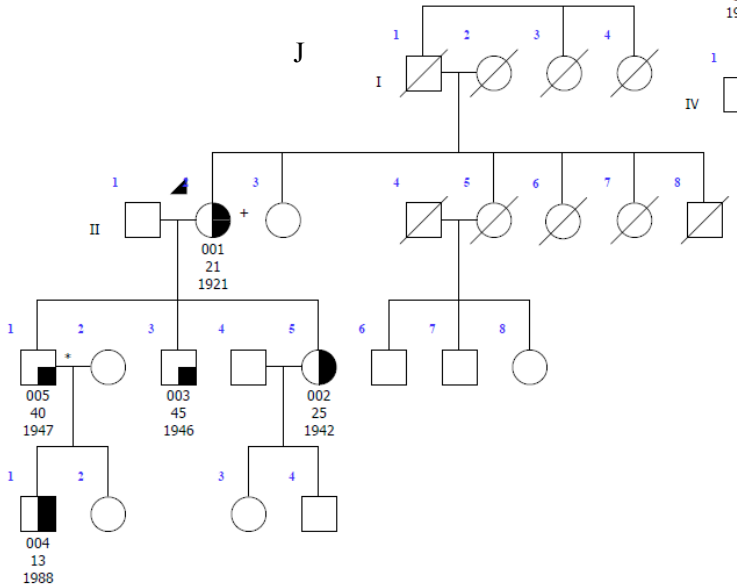
I

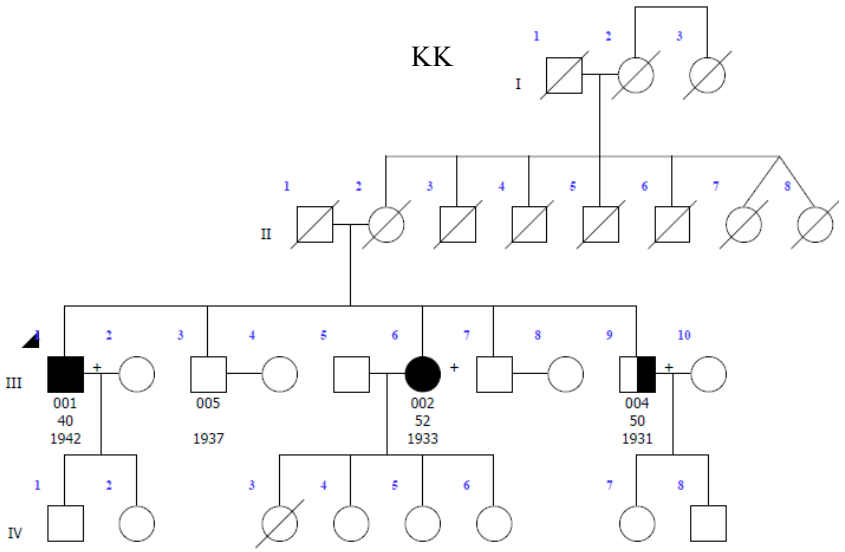

I
II
III
001
9
1925
IV
003
40
1958
002
40
1956
004
30
1952
V

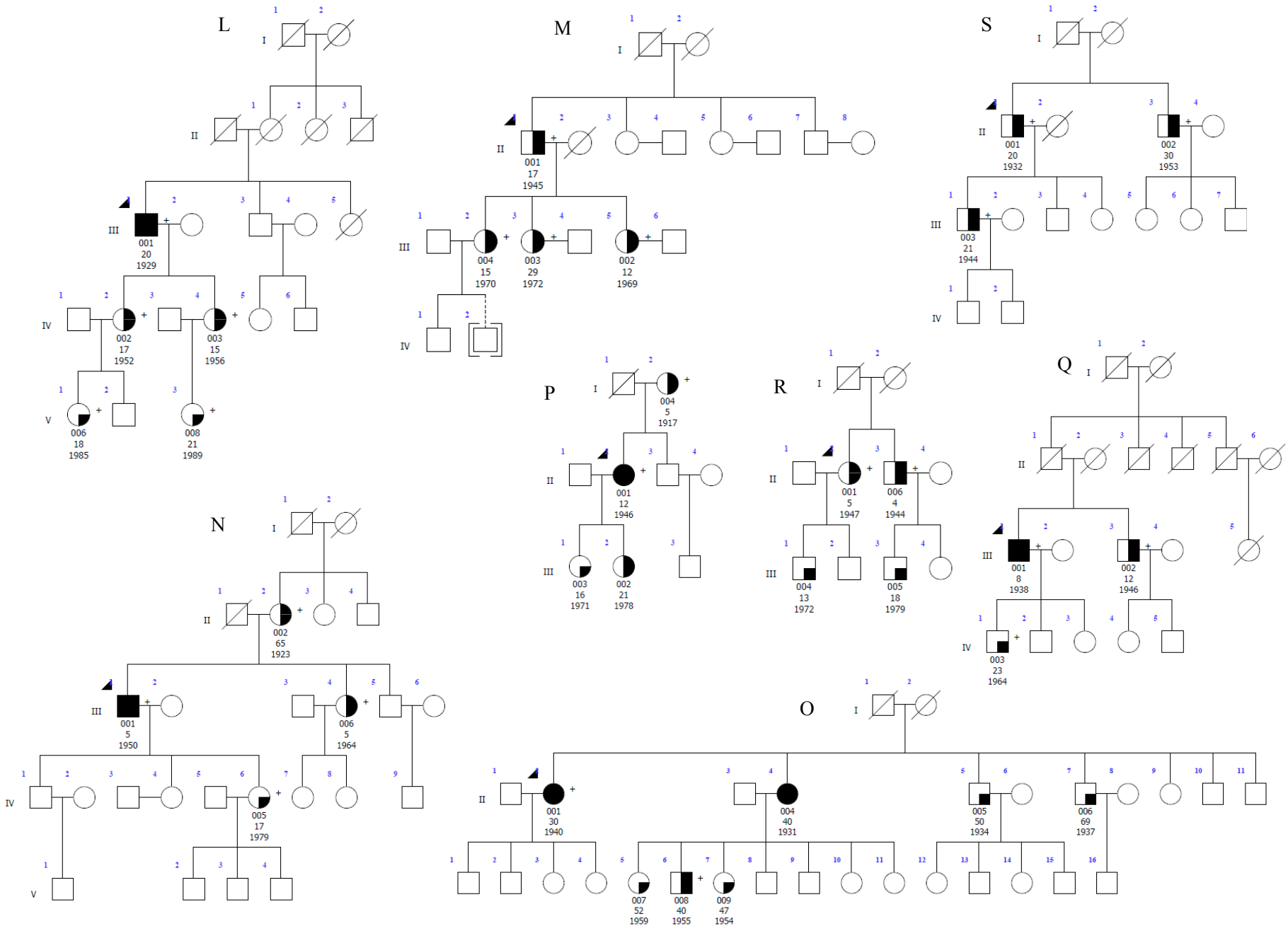

II
I
II
III
004
40
1939
IV
001
18
1971
002
5
1960
005
1958
V

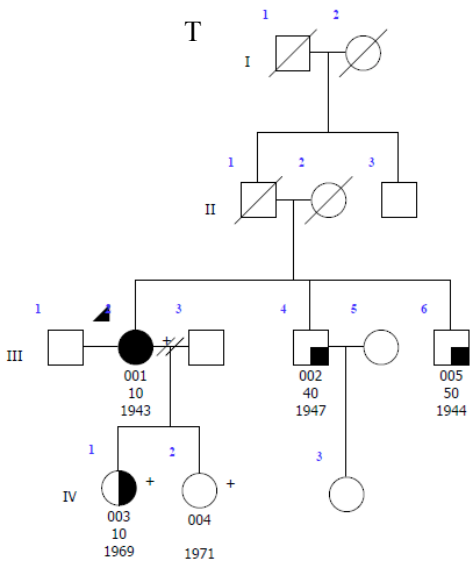

JJ
I
II
003
50
1936
III
002
18
1957
001
9
1960
IV
004
21
1989

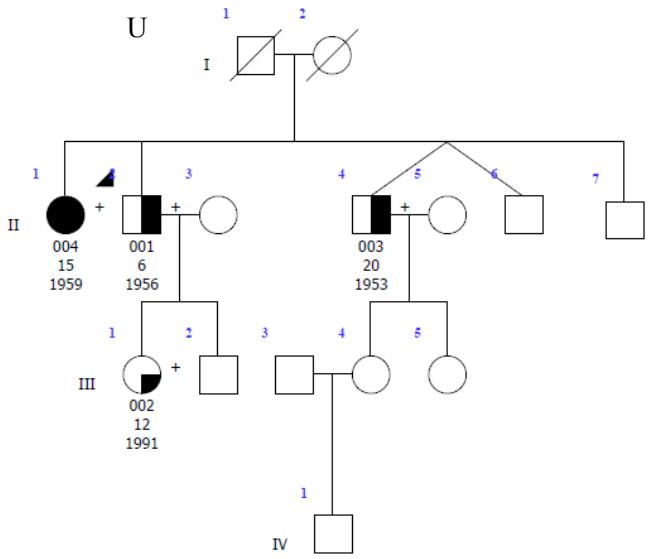

K
I
II
001
50
1928
003
70
1930
III
002
30
1952
IV

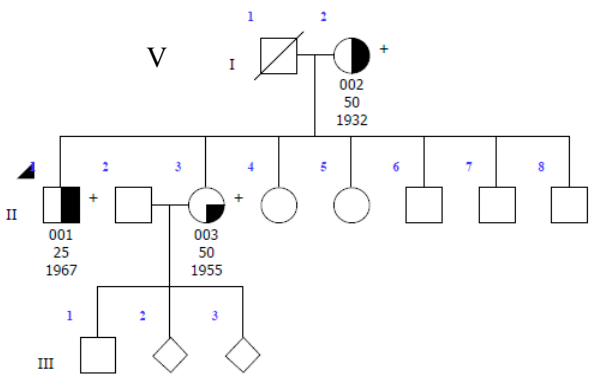

J
I
II
001
21
1921
005
40
1947
003
45
1946
002
25
1942
004
13
1988

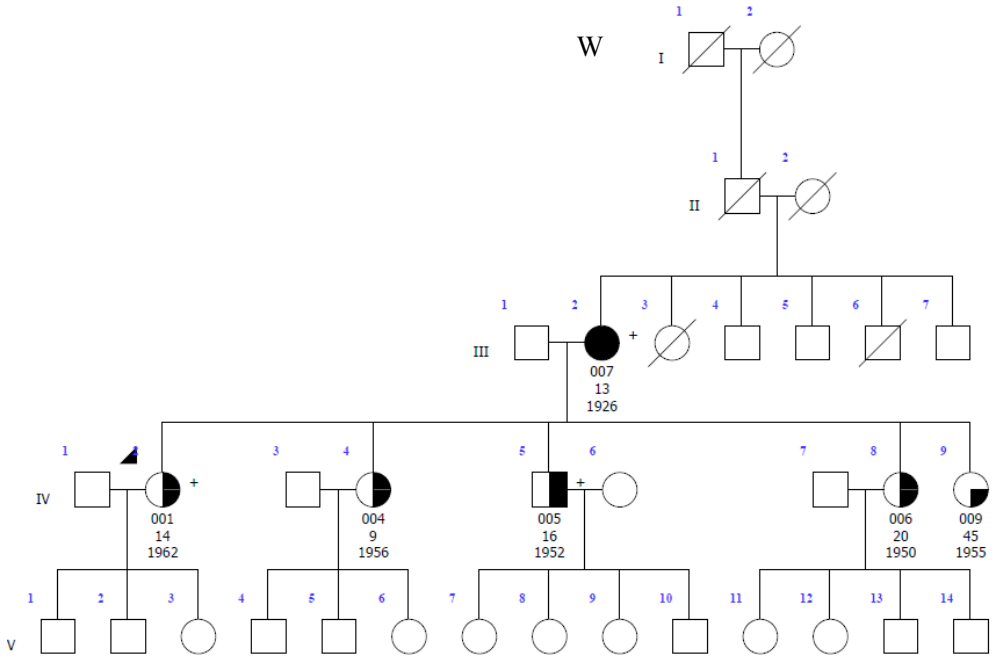

KK
I
II
III
001
40
1942
005
1937
002
52
1933
004
50
1931
IV

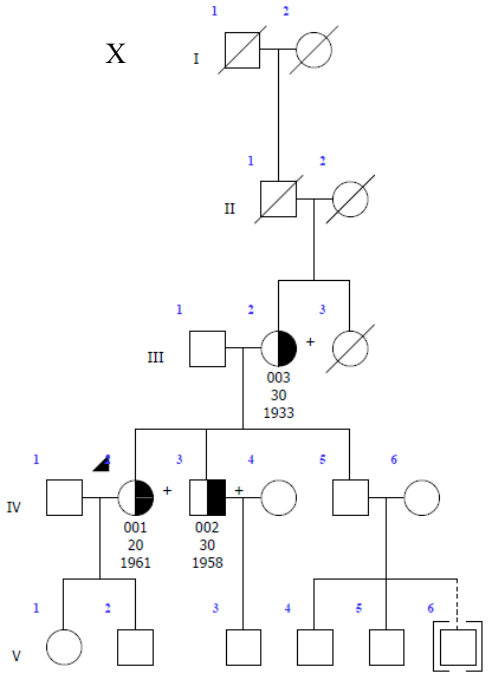

L
I
II
III
001
20
1929
IV
002
17
1952
003
15
1956
V
006
18
1985
008
21
1989
M
I
II
001
17
1945
III
004
15
1970
003
29
1972
002
12
1969
IV
S
I
II
001
20
1932
002
30
1953
III
003
21
1944
IV
P
I
004
5
1917
II
001
12
1946
III
003
16
1971
002
21
1978
R
I
II
001
5
1947
006
4
1944
III
004
13
1972
005
18
1979
Q
I
II
III
001
8
1938
002
12
1946
IV
003
23
1964
N
I
II
002
65
1923
III
001
5
1950
006
5
1964
IV
005
17
1979
V
O
I
II
001
30
1940
004
40
1931
005
50
1934
006
69
1937
007
52
1959
008
40
1955
009
47
1954


T
I
II
III
001
10
1943
002
40
1947
005
50
1944
IV
003
10
1969
004
1971

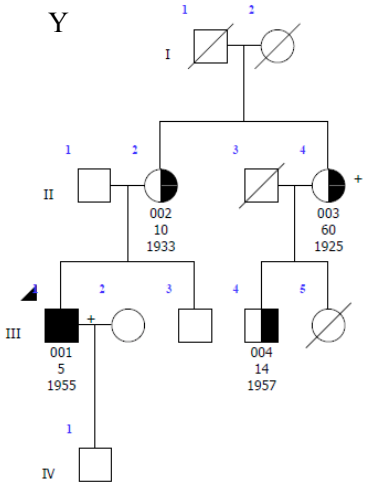

U
I
II
004
15
1959
001
6
1956
003
20
1953
III
002
12
1991
IV


V
I
002
50
1932
II
001
25
1967
003
50
1955
III


W
I
II
III
007
13
1926
IV
001
14
1962
004
9
1956
005
16
1952
006
20
1950
009
45
1955
V


X
I
II
III
003
30
1933
IV
001
20
1961
002
30
1958
V

Y


I
II
002
10
1933
003
60
1925
III
001
5
1955
004
14
1957
IV

Z

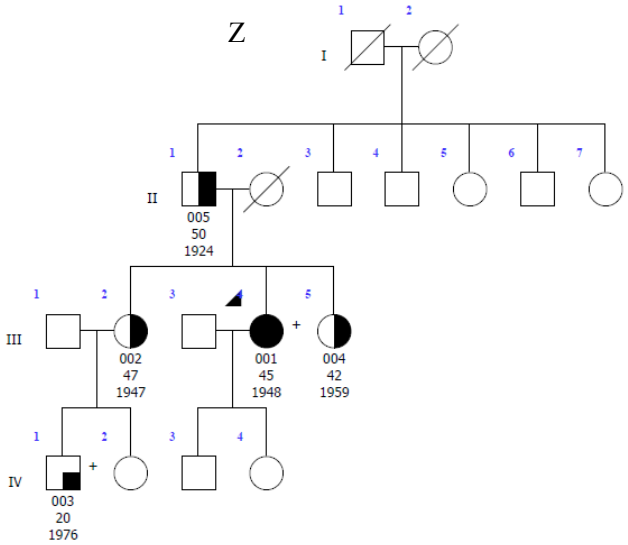

I
II
005
50
1924
III
002
47
1947
001
45
1948
004
42
1959
IV
003
20
1976

Table S1 List of candidate variants and oligo primers included in the Sequenom genotyping assay for co-segregation analysis

CHR	POSITION	1st-PCRP	2nd-PCRP	LENGTH (BP)	Tm (℃)	DIRECTION
1	3692043	ACGTTGGATGTGTGTCCAGCACAGAAGAGG	ACGTTGGATGGAGACACCAGCCTGCTATG	103	57.4	R
1	12023663	ACGTTGGATGCTACTACTTCAGCGTGGATG	ACGTTGGATGAGCCTCACTTGTTCTGTTGG	103	53.9	F
1	25893419	ACGTTGGATGAAGCTGAAGAGGTCATCCTG	ACGTTGGATGAGGACATGCATTACGCCCAG	108	55.5	R
1	31425169	ACGTTGGATGCAGCCTGGAGGATTTCATTG	ACGTTGGATGATTCAGCTGAAACTGGAGCG	90	49.8	R
1	31465406	ACGTTGGATGGTTGGAGAAATCCTCCACAG	ACGTTGGATGAACCAGAATGGTTCTGAGGG	95	61.8	R
1	98205947	ACGTTGGATGTGTCAGAAGAGCTGTCCAAC	ACGTTGGATGCTGTTATTTTCATTTGCAGAG	120	46.8	F
1	155224465	ACGTTGGATGAACAACTGTGCACCCAACTC	ACGTTGGATGACCACCGAGAGCATCAAGAG	90	55.3	F
1	156876581	ACGTTGGATGATGTTCTTGCCCTTCTGGTC	ACGTTGGATGATGGCACTGGCAGCGGAACT	117	48.7	F
1	158585056	ACGTTGGATGTACAGCACCATTGGATTGGC	ACGTTGGATGTGTTGCTCCAGGTTGTGTTG	100	49	R
1	161495025	ACGTTGGATGGTTGCGGATCATCAATGAGC	ACGTTGGATGCCAGGTCAAAAATGAGCACG	115	53.1	R
1	161495098	ACGTTGGATGTCATTTTTGACCTGGGTGGG	ACGTTGGATGTATCTCCAGCAGTGGCTTTC	110	52.6	R
1	206239580	ACGTTGGATGGTTCTCCAGGGTCAGAAAAG	ACGTTGGATGCCCACATGTCTCTGCTTTTC	99	49.3	R
1	207013255	ACGTTGGATGAAGAGTTGGCAATGCTGCTG	ACGTTGGATGAGAACCTCCTGGCGTTCTAC	116	51.7	F
2	110332225	ACGTTGGATGGCTGTCAAGGTTCTTCATGG	ACGTTGGATGTTCTCGCATCCATGTGTGTC	111	47.9	F
2	179401890	ACGTTGGATGGCTCTATTGAAGAACTCCGC	ACGTTGGATGCCACATAGTTGGTGATCCAG	99	48.7	R
2	179482562	ACGTTGGATGTGTGGCTGCACTTGGTTTTC	ACGTTGGATGATGTGGTAGAAGGACAGGAG	100	47.7	F
2	203500098	ACGTTGGATGGTTCCGTTCCAGCTGAAGCA	ACGTTGGATGACCACAGCAGCCACCGTTG	119	49.1	R
2	203560633	ACGTTGGATGTACGGTGCAGCAAGAGTATC	ACGTTGGATGGACAAAACACGACAGCCTTC	100	51.2	F
2	231865095	ACGTTGGATGCTCACCTCTCTCCTTCTGTC	ACGTTGGATGAAGTGGCACAGGAGCAGGAA	92	53	F
3	57557991	ACGTTGGATGGCCAGTCAAGTCCTTCATAC	ACGTTGGATGTACAGTGGTATGTTCAAGCC	86	59	R
3	78667061	ACGTTGGATGCCCACGTAACAAAAGCCTTC	ACGTTGGATGGGTCTCAGATATGGATACGG	117	46.6	F
4	15627106	ACGTTGGATGTGTCTGTTGGCAGCAGCATT	ACGTTGGATGTGGGAGTGATGACATAGTTC	120	49.3	F
4	15638171	ACGTTGGATGGCCCAATGAACAGGGTAAAG	ACGTTGGATGGTCGATGCAGTCAAGTAAGC	105	45.2	R
4	24801667	ACGTTGGATGTCCTCGCCAGCGTGGACGAC	ACGTTGGATGACGGCAGCCTCTGGAGGTA	117	49.8	F
4	37962105	ACGTTGGATGAGCTTCAGCACATCCTTGGC	ACGTTGGATGGGCTACTCTGATCTACGTTG	95	51.3	R
4	53492348	ACGTTGGATGTCCTGTTTCTTCTCCTCCTG	ACGTTGGATGCTTTGATAACTACATGCAGC	113	46.1	R
4	138450988	ACGTTGGATGGTAGGCACCAATGTGATGTC	ACGTTGGATGTGGCCGAATCAACTTACCAG	99	50.1	F
5	60834687	ACGTTGGATGTGTGGCTTAGAGTGAACCAG	ACGTTGGATGGAATCTACTGCTCCATCAGG	114	48	R
5	70308515	ACGTTGGATGACAGCAGAAGCACTGAATCC	ACGTTGGATGGTGACTTATGAGCCGTACAG	119	54.4	F
5	89949082	ACGTTGGATGTAGCAGAAACCAACCTCCAG	ACGTTGGATGCTAAACACTCTGGATCCAAG	107	50.1	F
5	114598521	ACGTTGGATGATGGCGGCCACTGAGGATGA	ACGTTGGATGAAAATCGCACGTGCCGATCC	105	51.6	F
5	122425849	ACGTTGGATGGCGCCGCGGGTCTCCTGA	ACGTTGGATGGCTGTCCGGCGGAGGCTCA	109	53.3	R
5	170863253	ACGTTGGATGGGACCAGTGGGAAACACATC	ACGTTGGATGGCACATACCATACTTGTCCC	100	55.8	R
5	175923567	ACGTTGGATGGGTCATCAGGTTGAATGAGG	ACGTTGGATGTCCTGGCCATGATTATGCTG	98	46.6	F
6	10903147	ACGTTGGATGTCCAAAAAGGGCTTGAAGAC	ACGTTGGATGACATGCCTTCTGACAAAGGG	108	51.5	F
6	70410688	ACGTTGGATGGGTTTTCCCTCTTGATTATA	ACGTTGGATGCCAAAAGAACCATATGCCA	120	46.1	F

6	70451681	ACGTTGGATGGTCCCTATTTGAAGAACTTGG	ACGTTGGATGCAAAAGTCCAAAGTATCAGG	110	45.2	F
6	71571555	ACGTTGGATGACTCTTCATGGTGACACTCG	ACGTTGGATGGATAACAACAGTCGAAGAAC	118	50.5	R
6	71665630	ACGTTGGATGACTGAGCAGCGCAACGCGG	ACGTTGGATGTCGTCAGCGAAGAAGAGCAC	109	51.7	F
6	71665759	ACGTTGGATGACTGGATCCTGGTGGAGGAC	ACGTTGGATGACGTGCAGGTGAGTGCTGG	111	59	F
6	90482411	ACGTTGGATGTTGCTGACACACACTGCAAG	ACGTTGGATGCAGGGAGAAGTTCACTTTCG	95	46.3	F
7	23871820	ACGTTGGATGCTGATTGCTCCTTTGGCATC	ACGTTGGATGAATCCCTCCTCTGTAGCTTG	116	48.1	R
7	73083825	ACGTTGGATGCACGAAGCTCCTGGTATTTG	ACGTTGGATGTACGCGCTGGCCAAGGAGAA	108	52.8	F
7	106851022	ACGTTGGATGAGGAGGAGCCTGTCCTTTTC	ACGTTGGATGTGTGTTTCAACCCTGCAGAG	112	55.2	R
7	106871106	ACGTTGGATGCTTACCTGAATGATCTCAGC	ACGTTGGATGTTCAGATGGAGTTGGCTGTG	107	48.4	R
7	107204365	ACGTTGGATGGATTCTGCGTCACCACCTG	ACGTTGGATGACCTTCCATGTTGGCAGGTG	95	49.4	F
7	150690937	ACGTTGGATGTGCTTGCCGCACAGCCCAA	ACGTTGGATGACATGGGCAACTTGAAGAGC	108	56	F
7	150692296	ACGTTGGATGTCTTCACACGAGGGAACTTG	ACGTTGGATGACAAACCCTTCCTGATGACC	119	58.9	F
7	150695652	ACGTTGGATGGTTCCAGATTCGGAAGTCTC	ACGTTGGATGACCAGCTCTTTCCCCATGC	107	52.8	F
8	23177415	ACGTTGGATGATGGTGTGTGGCCAAAACTG	ACGTTGGATGCTTCTCACCTGGAAGGCGT	115	55.9	R
8	86127236	ACGTTGGATGGATAAACATTATGAATGAGG	ACGTTGGATGTTGCGAAATGTGCAAGTTCC	103	47.8	F
8	99441344	ACGTTGGATGAGGCAGAGGCAGTCAGCTTT	ACGTTGGATGTACCGTCAGTATGACCACAG	97	52	F
8	99441520	ACGTTGGATGAGCTGTGACTTTGGAGATGG	ACGTTGGATGTTCGTCAGGCTTGCCATAAG	118	49.8	R
9	90535706	ACGTTGGATGATGATCTTGCTGGACTGACG	ACGTTGGATGACAGTCAAGTTTCTGCCCTC	103	45.2	R
9	91090065	ACGTTGGATGAACCAGGAGAAGTTGTGGAC	ACGTTGGATGTGACCATGCCAGTCCTTTTC	104	47.5	R
9	114337013	ACGTTGGATGATATAACAGAACCGGCCCAC	ACGTTGGATGTTTCTCCTTGTCCAGCCTTG	102	51.2	F
9	114345767	ACGTTGGATGATGGTTAATGCAGCAGCTGG	ACGTTGGATGACTGGAGAAAGGCACTTACC	99	47.2	F
9	127618232	ACGTTGGATGTGGAACTGGTCCTTGAATCC	ACGTTGGATGATCAGGTGGCTGGGACAAG	113	48.3	F
9	127619818	ACGTTGGATGTCTAGTCTGATCGGCAGCTC	ACGTTGGATGTCTACTCTTAGCAGGGAGTC	115	64.9	F
9	130941045	ACGTTGGATGATGAGCAGCCTCACACCCAG	ACGTTGGATGACTGCATCAGGTGGCATCTC	120	53.1	R
10	49667877	ACGTTGGATGCCTAGAGGAAACAGTCCACC	ACGTTGGATGATGAAGTCCACACACTGCTC	101	55.8	R
10	82298130	ACGTTGGATGTGCACGACATGTACATCGAC	ACGTTGGATGCCCGCATTTTGTAGAAGAGG	101	52.4	F
11	556092	ACGTTGGATGTCCAGGTCTGGGGCGAGG	ACGTTGGATGAGCCGCGGATGGAGAGGAG	113	67.3	R
11	10014554	ACGTTGGATGGTACTCTATCTTACCTTGTGG	ACGTTGGATGCCAAGGATATAGATCCTGCC	101	46.1	F
11	18723197	ACGTTGGATGATGGGTAGAGGAGGGCATAG	ACGTTGGATGATGCAGCCACTGGCTCTTC	96	54.3	F
11	64374967	ACGTTGGATGACTCCCGGATCACCTCCACT	ACGTTGGATGCTTCCCCCATCTGCCCACA	100	52.7	R
11	85436267	ACGTTGGATGTTCCCTTCCAGCTGGATAAC	ACGTTGGATGAGTCATTCTTCCACCCAGAC	117	48.4	F
11	126143283	ACGTTGGATGTGTCTGCACACAGGAGTACC	ACGTTGGATGTTTCTGAAGCCAGCAAGAGG	112	50.1	F
12	56079002	ACGTTGGATGTTCTCCTCCTTGAACTGCTG	ACGTTGGATGTCTTCAAACGGGCGAAGCAC	114	51.8	F
12	57442091	ACGTTGGATGAGGAGGACAAGATCCTCCAC	ACGTTGGATGGGACATCAGCAAACTGTGCC	119	49.7	F
12	58111943	ACGTTGGATGAACTGTCTTTCTCCCCACTC	ACGTTGGATGTTCTGCAGCATCATCCACAG	101	52.6	F
12	58131363	ACGTTGGATGAGAGACCGAAGCTCCAGTCC	ACGTTGGATGTGTCATGGCCCGAAAGCGAG	97	48.3	R
12	58131868	ACGTTGGATGAAGAGACGTTCGTGCCGCTT	ACGTTGGATGTGCCAGAGGCTCCGAGACTG	101	51.9	R
12	62997109	ACGTTGGATGCCGCTCGAAAGCAGGAAATG	ACGTTGGATGCAGTCCAGAACCAGTTGGTG	81	55.6	F
12	77424029	ACGTTGGATGACAGAAAATACTGGGTGGGC	ACGTTGGATGCTGTTGCTCCTTTCCCTGTC	96	46.6	F
12	78388607	ACGTTGGATGTTCTCCTCATACAGTCTGGC	ACGTTGGATGGACTGGTTGCAATTCCACTG	87	57.3	F
12	112037104	ACGTTGGATGTCCCTCCCGGCAGAGCTCG	ACGTTGGATGCCGTTGCCGTTGCTACCAAA	80	67.1	R
12	133297389	ACGTTGGATGATCACCCACCTGGTGATCC	ACGTTGGATGCCTCAGGATCGAGTGATCTT	118	56.3	F
13	28155579	ACGTTGGATGGCTACAAGTGCCTCAGAAAC	ACGTTGGATGTACTAGACTTCTTGCACAAC	113	49.5	F

13	28155645	ACGTTGGATGGAGGCACTTGTAGCAGAATG	ACGTTGGATGCCTAGTCTGCCATATTTGCC	102	54.5	R
13	43935483	ACGTTGGATGGAGTACAGCTTTTGCAGTGG	ACGTTGGATGCCCATTAACCCAATGATACC	120	54.4	R
13	115047277	ACGTTGGATGAGCAGCAGGAGGCTGAGAC	ACGTTGGATGTCTTCTCCTCGCGAGGTTTG	98	61.6	R
13	115047281	ACGTTGGATGAGGCTGAGACGCCGCCAAC	ACGTTGGATGCCGTCCCCACCTTGCTCAG	116	52.7	R
14	21876549	ACGTTGGATGAATTACTAACTGGGAGCGAG	ACGTTGGATGGTTGAATCATCTGCCTGCTG	112	50.1	R
14	77580363	ACGTTGGATGAAGCGCCTATGTTTGCTCTG	ACGTTGGATGTAGCTCACCTTTATGGGCTG	104	49.1	F
15	40627799	ACGTTGGATGCTGTGGAAGCCACTGACTC	ACGTTGGATGATGGAGCAGGGCCGACTGG	111	69.3	R
15	40629956	ACGTTGGATGAGCGGGAGCAGATCGACCTA	ACGTTGGATGTGGACTTGGCCTTGTCCAG	113	55.5	F
15	65621867	ACGTTGGATGGATGTGGAAAACCAGCTGTC	ACGTTGGATGTCCCCGTCTCGCCCCATTTA	107	57.3	R
15	71197007	ACGTTGGATGCTCGGATACAAAATATTTC	ACGTTGGATGACATCTCAGAGTCGAAAGCC	118	48.5	R
15	71203904	ACGTTGGATGCCTGATTTCCATGAAGATCC	ACGTTGGATGTCTGGTTTTCAGAATCAAG	100	45.3	F
15	71300911	ACGTTGGATGAATCTTCAGTTTAGGAAGC	ACGTTGGATGTGGAGTGTTCAAACAGCAGG	113	48.5	F
16	824264	ACGTTGGATGACTGGTTTGGTCTTGGACTC	ACGTTGGATGTTCCGCAGCGTGGTGGCTT	118	52	F
16	11370067	ACGTTGGATGAGCCTCCTTCGAGAGCAGT	ACGTTGGATGACGTCGAGGTCTACGAGAG	90	51.4	R
16	11370190	ACGTTGGATGAACACCATGGTCCGATACCG	ACGTTGGATGTCTTGCCCATGCAACTGCTG	100	58.5	F
16	23569478	ACGTTGGATGACTTCACGTCATGCTTGGTC	ACGTTGGATGTCGGTCAGCAACGGCGAAG	114	55.6	F
16	30762527	ACGTTGGATGAAGGATGTGTGTCTCTCGCC	ACGTTGGATGCGGTGAAGATTATGGAAGTG	112	54.1	F
16	30764591	ACGTTGGATGGGGAGAGCTGTTTGACTATC	ACGTTGGATGCTCAACCCTTACCTGGTTTC	89	50.1	R
16	88495665	ACGTTGGATGACATCGGCGACAGGGAAGAG	ACGTTGGATGTAGTGGGAGCCTCCCCCAG	107	55.7	F
16	88497041	ACGTTGGATGAAGAGGAAGGCTCGGGGCGG	ACGTTGGATGCGGTGGCGCCTGTTCTTCT	97	53.4	R
16	88497170	ACGTTGGATGTCCGCGAGTACGACTTCGCC	ACGTTGGATGTTCCTCTTCTCGCCTCGGC	120	57.5	R
16	88497219	ACGTTGGATGTTCCTTCCTCTTCTCGCCTC	ACGTTGGATGAGTACGACTTCGCCTCGGA	118	52.5	F
16	88497245	ACGTTGGATGTCTCTGGGACCCTGGGTCA	ACGTTGGATGAGGAGGACGAGCAGCCTCCG	120	48.9	F
16	88497318	ACGTTGGATGAGACCCGCCCGGCTCCTCG	ACGTTGGATGAGAGGATGAGCCACAGAAAC	111	59	F
16	88497828	ACGTTGGATGCACGGGTGCCAGAAATTCTC	ACGTTGGATGACACAGGCCCCAGTCTCCCA	92	49	F
16	88498257	ACGTTGGATGAGAGGGGTGGGTCGAATCTG	ACGTTGGATGGACCTTGGAGTCCTCATC	82	47.4	F
16	88502673	ACGTTGGATGTCTGCCTCTGCCATGAGGAC	ACGTTGGATGCTATTGAGGAACCCATCCAG	105	50.4	R
17	3118915	ACGTTGGATGCCCAGGAGGATGAATTCCAG	ACGTTGGATGTGATATTCCTCTCCCCTTTC	101	52.4	R
17	4836266	ACGTTGGATGAATCAGCTGCAAAGCCTGCC	ACGTTGGATGACCAAGAGGCAGCGAGGTCA	116	48.2	R
17	4837740	ACGTTGGATGCTCCTTCCTGCCACTAGAG	ACGTTGGATGGAGGTTCGCTTCCCACTTTC	105	48.6	F
17	8025206	ACGTTGGATGTTGGGGCGCAGATAGCCGT	ACGTTGGATGGAGTGCCTGCTTCGCTTGG	115	48.8	R
17	27041879	ACGTTGGATGGATTTCTCCAGCTCAGCCAG	ACGTTGGATGTGTGTCAGGTGAGAATGTCC	107	53.2	R
17	40950556	ACGTTGGATGGTCACGATGTTCCGGGTTC	ACGTTGGATGGAGCAACTGCATTCCATGCG	103	53	R
17	59949696	ACGTTGGATGACAGAGCAAGATAGGCCTTC	ACGTTGGATGTTCTTCCCTGGTAACTTCCG	101	46.1	F
17	60002456	ACGTTGGATGGCAATTATTTCCATATCAGTG	ACGTTGGATGTGATACCAGGATGCTCTCTC	119	48.4	F
17	60003873	ACGTTGGATGTTTCTTATCCTGAGCCCAGC	ACGTTGGATGGACTTCTTCTGCCCTGTTTG	100	56.1	R
17	65344710	ACGTTGGATGAAATGGAGGCTGCCTCTTTC	ACGTTGGATGAAATTGAGCGTGCGCGACTG	101	45.4	R
17	80397532	ACGTTGGATGAGGTACGACCACTACTCTG	ACGTTGGATGATACCGCGTAGAAGCAACTG	109	45.2	F
19	7974729	ACGTTGGATGCTCGCCATCCTCAGAGAGC	ACGTTGGATGTGAACAGGGTTGACGGGAG	106	54.1	R
19	9801082	ACGTTGGATGGGAGAGAAGCCTTATGGATG	ACGTTGGATGAAGGGCCTTTCTTCATGGTG	118	50.9	R
19	10073550	ACGTTGGATGAGCCTCCAGGCTTTTCCTTG	ACGTTGGATGAGATCTAGATGGCACCAAGG	116	52.8	F
19	12911622	ACGTTGGATGCATTGCTCAACCTCTGGAAC	ACGTTGGATGCATCCTTAAAGACTTGGGCG	97	45.5	F
19	19368787	ACGTTGGATGGGTAGCAGTAGACGCCGAA	ACGTTGGATGTGAACCCGCGAGCGCGCTG	116	61.9	R

19	39294173	ACGTTGGATGTGTTTCCCCAAGCCATACCG	ACGTTGGATGAGATATCTGTCGACTCTTGG	119	46.5	F
20	1433202	ACGTTGGATGAGTGAAGGCTTTGAAGGCTC	ACGTTGGATGGGTGAACTTGGATATGGAGG	93	48	R
20	62319887	ACGTTGGATGAGAAGCTGCCTTTGCTCCTG	ACGTTGGATGCTGCCTCTTCCTCCCACAG	112	47	R
22	28501710	ACGTTGGATGAATTGGCTGTGCAGACTTCC	ACGTTGGATGCGTGGTTGCTCATGAACTTG	98	47.1	F
X	1475174	ACGTTGGATGCCCAACATGACTGCAAAGTG	ACGTTGGATGTGTATCTGAAGCTCATAGCG	118	47.4	F
X	1508583	ACGTTGGATGACAATGCAGTCCACGATGCC	ACGTTGGATGTCTGGTCTGAACACCCTCTG	115	52.2	F
X	1537953	ACGTTGGATGAGACGCGGCTGAGGTTCAT	ACGTTGGATGAGAAGCGGGACAGATTGAAG	110	60	R
X	1540706	ACGTTGGATGCCCATTTCTCCAGGTTACAG	ACGTTGGATGAAGCCGTGCAGAGAGTATTC	100	46.4	R
X	1719897	ACGTTGGATGACCACGCTGCACCCCCTCG	ACGTTGGATGCCGTTCACGCTTTTGGGAG	118	54.6	F
X	2139200	ACGTTGGATGAGGGACTTGGTCTCTTTCTC	ACGTTGGATGTCCATCTACGCAGCAGTCAC	100	47.3	F
X	12937720	ACGTTGGATGCTCTATTTGGCCTGGAACTG	ACGTTGGATGGTCAGCGTTTCAAATACTCC	100	46.9	R
X	100177974	ACGTTGGATGCCTCACACTGTGGAAGAAAG	ACGTTGGATGCCTCGCCATCTATTAGCATC	100	46.6	F
X	140984599	ACGTTGGATGGAGGATGCCTCCTCCACTT	ACGTTGGATGAGGGTGAGGATGAGGACAAG	99	50	F
X	152990762	ACGTTGGATGGACTTTGTGGGCTCCATAGG	ACGTTGGATGTGACATGCCGGTGCTCTCCA	120	57.8	F

Table S2 Genomic regions providing evidence of linkage in ET families

Phenotype P1 or P2*	LOD score	*P* Value	Chr	1 LOD unit support interval (bp)	No. of Genes	Candidate Genes
P1 P2	2.288 3.707	$1.8x10^{-5}$	5	1166067-91384315	426	*SEMA5A* *NDUFS4*
P1	3.013	$1.1x10^{-3}$	6	7613899-8992572	11	*BLOC1S5* *TXNDC5*
P1 P2	3.114 3.774	$7.6x10^{-5}$	7	8939150-16780405	18	*NDUFA4* *PHF14* *TMEM106B* *ETV1*
P1 P2	2.729 3.312	$4.7x10^{-5}$	8	94960934-101969832	27	*NDUFAF6* *TP53INP1* ***KCNS2***
P1 P2	2.483 3.071	$8.4x10^{-5}$	11	99051583-105083703	45	*CNTN5*
P1 P2	2.698 3.568	$2.5x10^{-5}$	16	51954866-53935253	3	*IRX3*

*A stringent analysis (phenotype 1, P1) in which only individuals with a diagnosis of definite or probable ET were classified as affected, and individuals with a diagnosis of possible ET were classified as unknown, and a secondary analysis with a broader phenotype (phenotype 2, P2) in which individuals with a diagnosis of definite, probable or possible ET were classified as affected.

Abbreviations/definitions: Chr, Chromosome; LOD, The 'LOD' (logarithm (base 10) of odds) score corresponds to the LOD score of the gene; p-value, significance level for LOD score; 1 LOD unit support interval was determiend by finding the maximum lod score, Zmax and determining those points for theta for which the LOD score is at least Zmax-1. Candidate genes are positional candidates of interest based on ET phenotype.