MAYO CLINIC

Insights from whole genome sequencing of a family with Venous Thromboembolism

Mariza de Andrade, PhD
Professor of Biostatistics
Division of Biomedical Statistics and Informatics
Mayo Clinic, Rochester, MN

MAYO CLINIC

Acknowledgments

John A. Heit, MD
Sebastian M. Armasu, MS
Bryan M. McCauley, MS
Tanya M. Petterson, MS
Daniel O'Brien, MS

MAYO CLINIC

Venous Thromboembolism (VTE)

- It is a common, lethal disorder that affects hospitalized and non-hospitalized patients with over 350,000 incident cases per year in the US.
- Survival after PE is worse than expected. One quarter of PE patients present as sudden death.
- VTE incidence is about 1 per 1000 person-years between 1981 and 2000 and increased over 5-year period 2001-2005.

Silverstein et al. Arch Int Med 1998
Heit et al. Mayo Clin Proc 2001
Heit J Throm Thrombolysis 2006

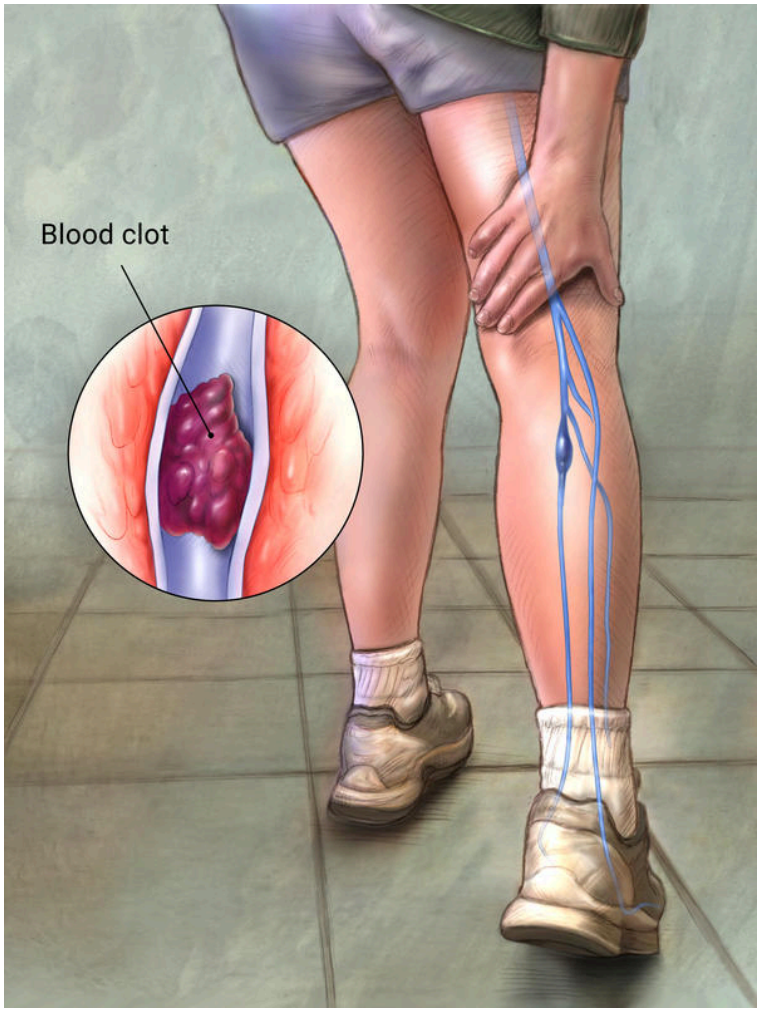
MAYO CLINIC

Venous Thromboembolism (VTE)

- **VTE** includes both **D**eep **V**ein **T**hrombosis (**DVT**) and **P**ulmonary **E**mbolism (**PE**).
- **D**eep **V**ein **T**hrombosis (**DVT**) is a blood clot that forms in a vein deep in the body. Blood clots occur when blood thickens and clumps together. Most of the deep vein blood clots occur in he lower leg or thigh, and they can also occur in other parts of the body.
- Pulmonary Embolism (**PE**) is a blockage in one of the pulmonary arteries in the lungs. In the majority of the cases, PE is caused by blood clots that travel to the lungs from the legs, or rarely from other parts of body (**DVT**).
- VTE is usually defined as DVT and PE±DVT.

MAYO CLINIC

Deep Venous Thrombosis (DVT)



Symptoms

Pain areas: calf, foot or leg

Also common: swelling in extremities, tenderness, or warm skin

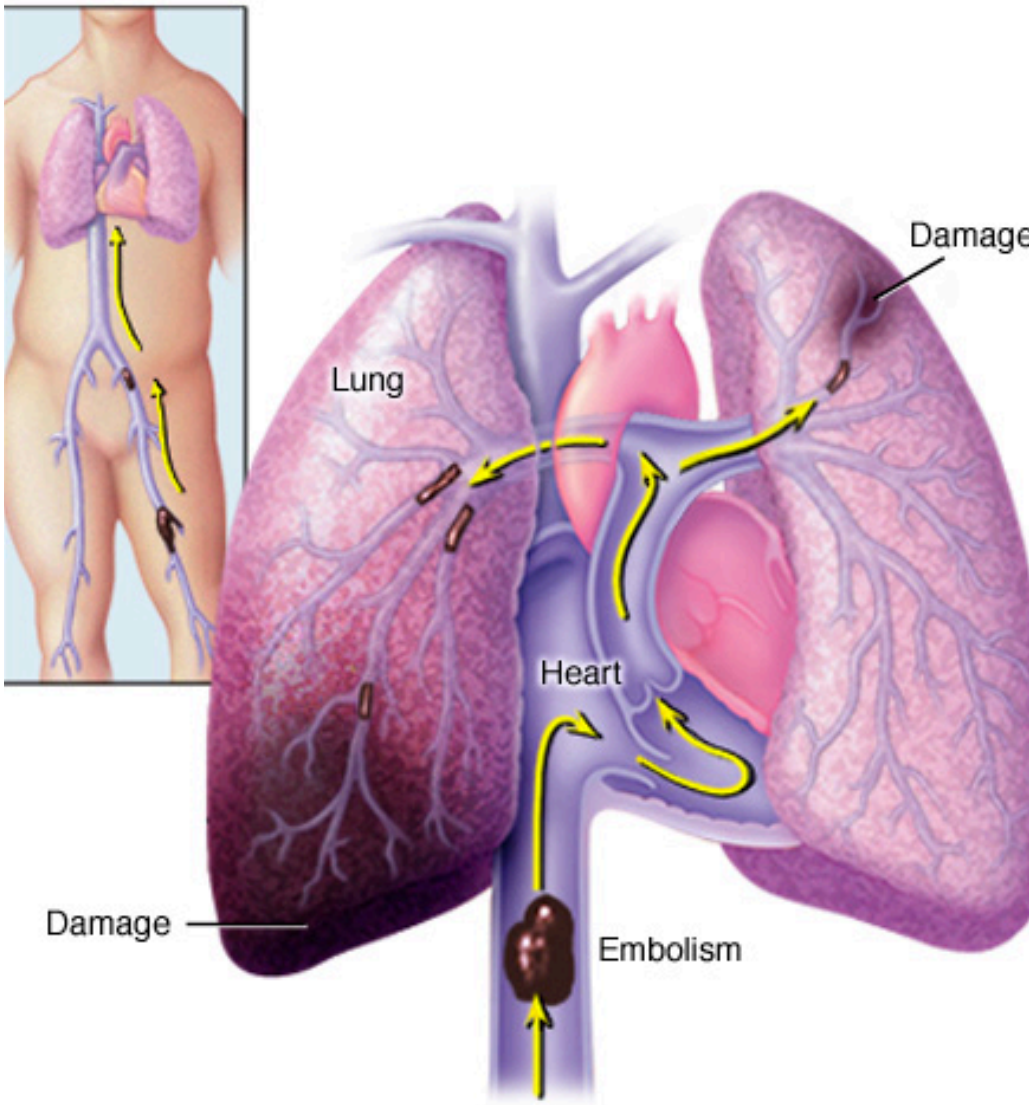
Treatments

Blood thinners

Lifestyle: Compression stockings

MAYO CLINIC

Pulmonary Embolism (PE)



Symptoms:

Shortness of breath, chest pain, and cough.

Treatment:

Blood thinner.s

MAYO CLINIC

Risk Factors

- Idiopathic VTE is an unprovoked VTE and it is not associated with any risk risk factors. About 25% to 50% of patients with first-time VTE have an idiopathic VTE.
- Un-idiopathic VTE is a provoked VTE due to cancer, hospitalization, pregnancy, post-partum, hormone replacement therapy, surgery among others.
- The known genetic risk factors are factor V (F5) Leiden (rs6025) on 1q23, prothrombin gene (F2) G20210A on 11p11, and non-O blood type (8176719) on 9q34.1-q34.2.

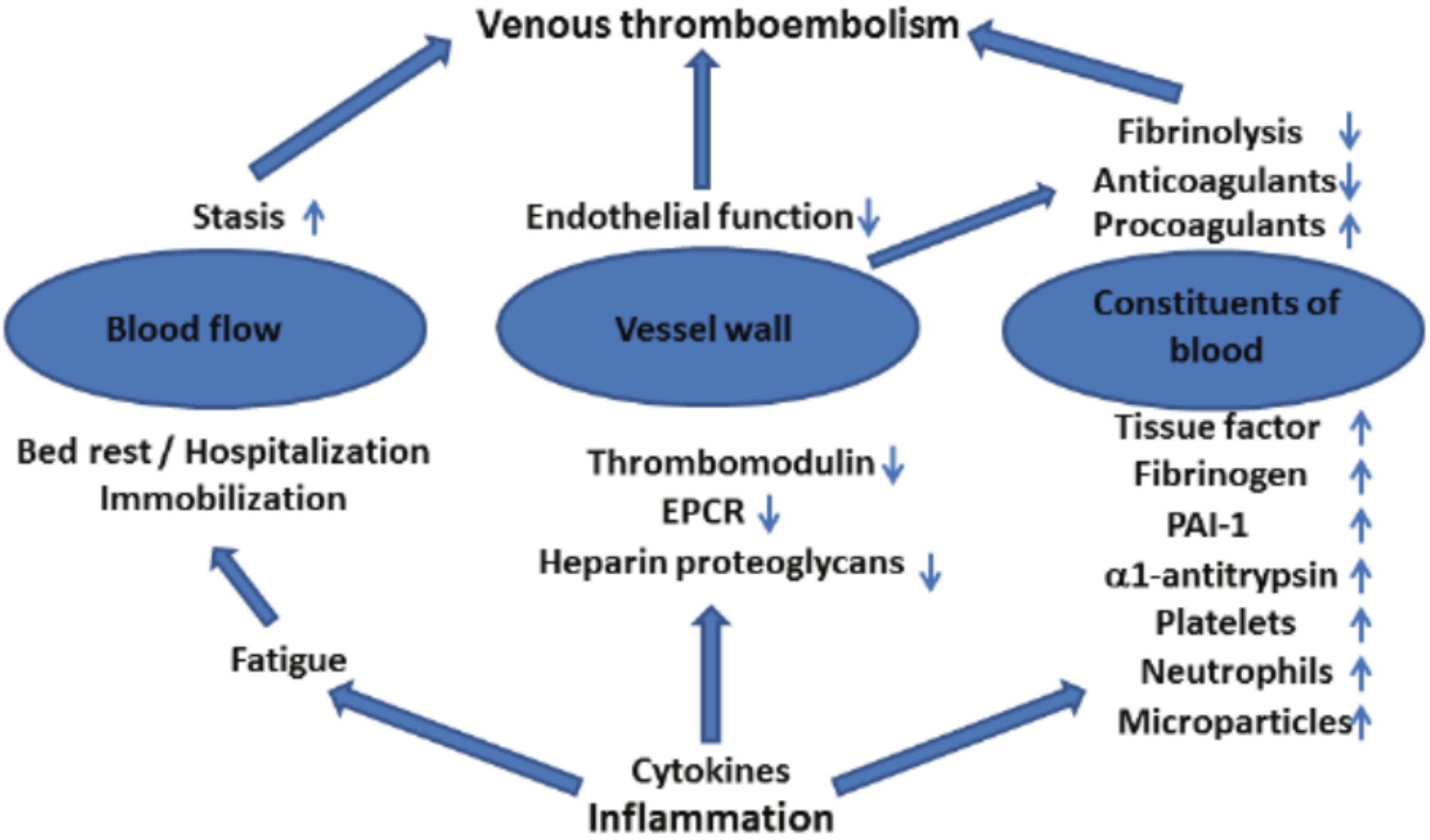
Heit et al. 2015

- The two main mechanisms are loss-of-function of anticoagulant proteins and gain-of-functions of procoagulants.
- The latter is due to increased synthesis or impaired downregulation of a normal protein or to synthesis of a functionally hyperactive molecule.
- The risk of recurrent VTE is increased in anticoagulant-deficient patients and in homozygotes for gain-of-function mutations.
- In families with thrombophilia and VTE, primary antithrombotic prophylaxis during risk situations lowers the rate of incident VTE.

Martinelli et al. 2014

MAYO CLINIC

Virchow's triad



MAYO CLINIC

Box 2 | Mechanisms associated with thrombophilia

Known mechanisms

Loss-of-function

- Antithrombin deficiency
- Protein C deficiency
- Protein S deficiency

Gain-of-function

- Factor V Leiden
- Prothrombin G20210A
- High factor VIII level
- Non-O blood group
- Dysfibrinogenaemia

Postulated mechanisms

- Low tissue factor pathway inhibitor level
- High fibrinogen level
- High factor IX level
- High factor X level
- High factor XI level
- Resistance to antithrombin
- Global hypofibrinolysis
- High thrombin activatable fibrinolysis inhibitor level
- Hyperhomocysteinaemia

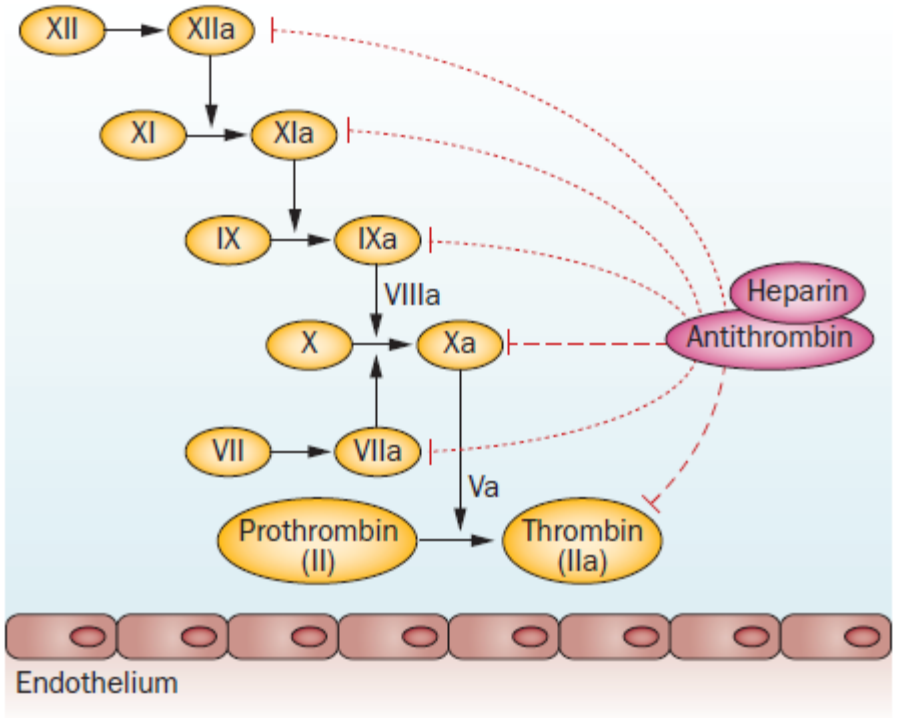


Figure 1 | Anticoagulant mechanisms of antithrombin, which mainly inhibits factor IIa and factor Xa, but also factors VIIa, IXa, XIa, and XIIa. The rate of interaction with target proteases is accelerated by heparin. Solid lines denote activation and broken lines inhibition (dashed lines, strong inhibition; dotted lines, weak inhibition).

Martinelli et al. 2014

Table 3 | Incidence of VTE in studies of families with thrombophilia

Reference (year)	Study information			Incidence of VTE per 100 person-years (where available, the relative risk vs noncarriers is in brackets)				
	Relatives/ carriers	Observation years	Antithrombin deficiency	Protein C deficiency	Protein S deficiency	FVL	PT20210A	Multiple defects
Prospective studies								
Pabinger (1994)[157]	93/44	211	NA	2.5	2.2	NA	NA	NA
Sanson (1999)[158]	735/208	611	4.0	1.0	0.7	NA	NA	NA
Middeldorp (2001)[159]	855/470	1,564	NA	NA	NA	0.58§	NA	NA
Simioni (2002)[160]	561/313	1,255	NA	NA	NA	0.67	NA	NA
Vossen (2005)[161]	575‡	3,283	1.7	0.7	0.8	0.1	NA	NA
Coppens (2006)[162]*	464/236	1,816	NA	NA	NA	NA	0.37§ (3.1)	NA
Mahmoodi (2010)[163]	382/149	3,472	2.29§ (10.2)	0.95§ (4.1)	1.55§ (9.6)	NA	NA	NA
Retrospective studies								
Lijfering (2009)[63]*	2,479/1,528	NA	1.77§ (28.2)	1.52§ (24.1)	1.90§ (30.6)	0.49§ (7.5)	0.34§ (5.2)	NA
Martinelli (1998)[164]*	723/396	NA	1.0 (8.1)	0.72 (7.4)	0.78 (10.4)	0.25 (4.6)	NA	NA
Middeldorp (1998)[165]	437/236	12,240	NA	NA	NA	0.45§ (4.2)	NA	NA
Bucciarelli (1999)[166]	513‡	19,542	1.07	0.54	0.50	0.30	NA	0.67
Simioni (1999)[167]	793/405	19,685	0.87 (10.6)‖	0.43 (10.6)‖	1.65 (10.6)‖	0.28 (2.8)	NA	NA
Lensen (2000)[168]	197/108	8,760	NA	NA	NA	0.34 (2.9)	NA	NA
Martinelli (2000)[169]*	1,093/640	43,208	NA	NA	NA	0.19 (2.9)	0.13 (2.0)	0.42 (6.4)
Tirado (2001)[170]	722/435	NA	2.94 (10.6)	0.36 (6.4)	1.04 (7.6)	0.31 (6.2)	0.23 (4.2)	NA
Bank (2004)[171]*	407/209	12,085	NA	NA	NA	NA	0.35 (1.9)	NA
Tormene (2004)[172]	294/152	8,347	NA	NA	NA	NA	0.11 (1.7)	NA
Brouwer (2006)[173]	468/224	4,174	1.94 (18.3)	1.58 (16.2)	1.50 (16.2)	NA	NA	NA
Couturaud (2006)[174]	553/322	17,532	NA	NA	NA	0.43 (2.5)	NA	NA
Rossi (2011)[175]	1,088/625	40,405	0.92 (12.8)	0.12 (5.1)‖	0.12 (5.1)‖	0.14 (2.3)¶	0.05 (0.6)¶	0.24/0.19/0.58#
Holzhauer (2012)[176]	533/146	18,278	2.82 (25.7)**	2.82 (25.7)**	2.82 (25.7)**	0.25 (1.7)	0.42 (2.6)	2.33 (19.6)

*Including probands with a history of VTE or arterial thrombosis. ‡Recruitment limited to carrier relatives of patients with VTE diagnosed with thrombophilia. §Observation period starting from 14–15 years of age. ‖Relative risk calculated on the total carriers of antithrombin, protein C, and protein S deficiencies. ¶Homozygotes excluded. #Incidence estimated in double heterozygotes for FVL and PT20210A, homozygotes for FVL or PT20210A, and in carriers of other multiple abnormalities. **Incidence for antithrombin, protein C, and protein S deficiencies combined.
Abbreviations: FVL, factor V Leiden; NA, data not available; PT, prothrombin; VTE, venous thromboembolism.

Martinelli et al. 2014

MAYO CLINIC

Table 4 | Risk of VTE among users of OC or HRT with or without inherited thrombophilia

Thrombophilia	OC users*	RR in case–control studies	Absolute risk per person-years (estimated‡)	Absolute risk per person-years (observed in family studies‖)	HRT users§	RR in case–control studies	Absolute risk per person-years (estimated‡)	Absolute risk per person-years (observed in family studies‖)
None	No	1.0	1.8/10,000	0.2/100	No	1.0	1.0/1,000	NA
None	Yes	2.0–9.0	0.4–1.6/1,000	0–0.5/100	Yes	2.1–3.2	2.1–3.2/1,000	0
AT deficiency	Yes	12.6	2.3/1,000	5.1–10.0/100	Yes	NA	NA	NA
PC deficiency	Yes	6.3	1.2/1,000	0–7.1/100	Yes	NA	NA	NA
PS deficiency	Yes	4.9	0.9/1,000	2.4–4.2/100	Yes	NA	NA	NA
Heterozygous FVL	Yes	11.3–34.7	2.0–6.2/1,000	0.5–2.0/100	Yes	6.7–15.5	0.7–1.5/100	2.9/100
Heterozygous PT20210A	Yes	5.1–16.0	0.9–2.9/1,000	0–0.2/100	Yes	2.9	0.3/100	0
Homozygous FVL	Yes	NA	NA	NA	Yes	NA	NA	NA
Homozygous PT20210A	Yes	NA	NA	NA	Yes	NA	NA	NA

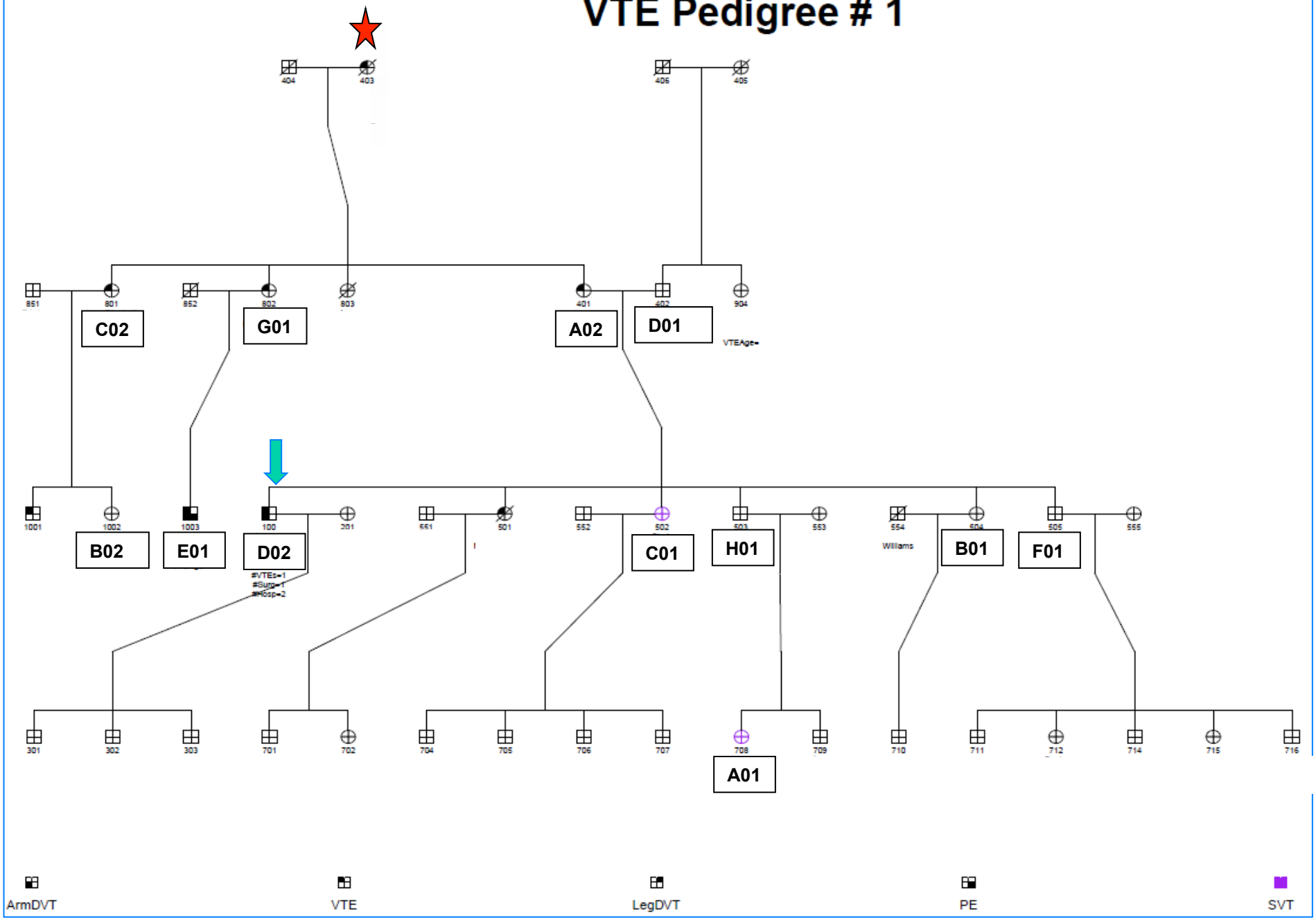
*Age <40 years. ‡Estimated as relative risk × VTE incidence in the general population (that is, 1.8 per 10,000 in women of fertile age, and 1.0 per 1,000 in women aged 40–60 years). §Age >40 years. ‖Families with the index case diagnosed with thrombophilia after VTE. Abbreviations: AT, antithrombin; FVL, factor V Leiden; HRT, hormone replacement therapy; NA, data not available; OC, oral contraceptives (combined); PC, protein C; PS, protein S; PT, prothrombin; RR, relative risk; VTE, venous thromboembolism.

SVT = Superficial Vein Thrombosis
DVT = Deep Venous Thrombosis
PE = Pulmonary Embolism
VTE = Venous Thromboembolism

Martinelli et al. 2014

MAYO CLINIC

VTE Pedigree # 1



ID	gender	age_vte	# vte	leg_dvt	pe	svt	# surg.	# hosp.	vital	Comments/event
403*	female	72	2	2	-	-	-	-	deceased	car trip
C02	female		2	1	1	-	-	-	alive	Factor V Leiden heterozyous
G01	female	59	1	1	-	-	-	-	alive	Breast cancer
A02	female	40/52	2	1	1(pe&dvt)	-	-	-	alive	
D01[&]	male	-	-	-	-	-	-	-	alive	no VTE
B02	female	-	-	-	-	-	-	-	alive	no VTE
E01	male	33/35/37	3	1	1 (pe&dvt)	1	-	-	alive	calf injury due to running
D02[P]	male	43	1	1	1	-	1	2	alive	calf tear-running
C01	female	54	1	1?	-	1?	-	-	alive	diagnosis unclear
H01	male	-	-	-	-	-	-	-	alive	-
A01	female	-	-	-	-	-	-	-	alive	-
B01	male	-	-	-	-	-	-	-	alive	-
F01	male	-	-	-	-	-	-	-	alive	-

* possible carrier
& control (married with A02)

SVT= Superficial Vein Thrombosis; DVT= Deep Vein Thrombosis; PE= Pulmonary Embolism
VTE = Venous Tromboembolism (DVT and/or PE)

MAYO CLINIC

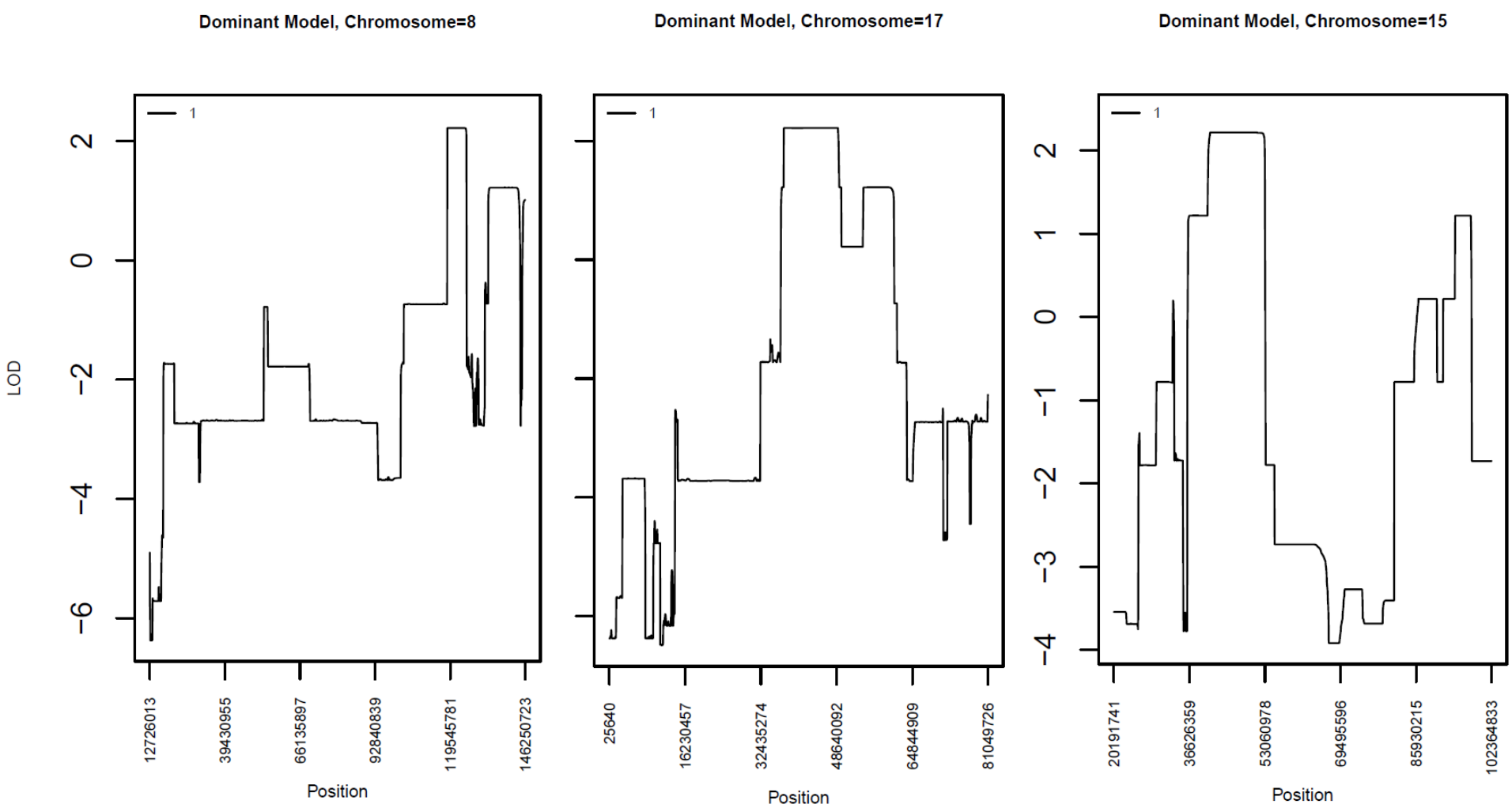
Linkage Analysis results using PVAAST

Chromosome)	Max Par. HLOD-Dominant	Max Par. Info-Dominant	Position-Dominant	Marker-Dominant	Max Par. HLOD-Recessive	Max Par. Info-Recessive	Position-Recessive	Marker-Recessive	Max NPL	BP-NPL	Marker-NPL	genes
1	0	0.92126	798959	1.798959	0	0.92126	798959	1.798959	0.126	5358384	1.5358384	
2	0.6994	0.87941	1.08E+08	2.108019522	0	0.93095	29350	2.2935	0.209	1.08E+08	2.1080195	
3	0	0.85782	105793	3.105793	0	0.85782	105793	3.105793	0.511	9218727	3.9218727	
4	1.2072	0.92698	1.08E+08	4.107603453	0.7524	0.92698	1.08E+08	4.107603453	0.564	1.08E+08	4.1076035	
5	0.8605	0.87386	1.5E+08	5.149689401	0	0.84242	75167	5.75167	0.378	164193	5.164193	
6	0	0.90022	206599	6.206599	0.7614	0.93741	3958190	6.395819	0.571	2885930	6.288593	
7	0	0.93155	62338	7.62338	0	0.93155	62338	7.62338	0.113	1.06E+08	7.1059957	
8	**2.2172**	**0.88001**	**1.23E+08**	**8.122516874**	**0.7614**	**0.93749**	**77640595**	**8.77640595**	**0.571**	**56886184**	**8.5688618**	**SNTB1, HAS2**
9	0.2181	0.93729	428706	9.428706	0	0.93178	46587	9.46587	0.512	880777	9.880777	
10	0	0.89858	135708	10.135708	0.5468	0.87511	37973308	10.37973308	0.259	37812532	10.378125	
11	0	0.87168	199256	11.199256	0.0181	0.91981	1.16E+08	11.11618783	0.127	1.16E+08	11.116323	
12	0	0.93749	218588	12.218588	0.7579	0.9375	59465715	12.59465715	0.571	59046845	12.590468	
13	0	0.92786	19168981	13.19168981	0.7916	0.9375	31904694	13.31904694	0.571	37385882	13.373859	
14	0	0.85258	20403909	14.20403909	0	0.85258	20403909	14.20403909	0.113	83263031	14.83263	
15	**2.2172**	**0.93727**	**47002318**	**15.47002318**	**0.7614**	**0.93747**	**99033438**	**15.99033438**	**0.511**	**27200942**	**15.272009**	**GATM,SQRDL,BLOC1S6,SEMA6D**
16	0.5014	0.80307	54691273	16.54691273	0.6719	0.89086	11333849	16.11333849	0.559	11333849	16.113338	
17	**2.2172**	**0.93742**	**37355066**	**17.37355066**	**0.7515**	**0.92441**	**60142971**	**17.60142971**	**0.567**	**59820068**	**17.598201**	**KANSL1 (KIAA1267) , CRHR1**
18	0	0.88571	159382	18.159382	0.7579	0.93723	12768846	18.12768846	0.35	12643193	18.126432	
19	**2.0924**	**0.88304**	**35643940**	**19.3564394**	**0.7579**	**0.9373**	**47536166**	**19.47536166**	**0.571**	**47323384**	**19.473234**	
20	1.2178	0.93691	7016222	20.7016222	0.1274	0.85401	7992466	20.7992466	0.518	7992466	20.799247	
21	0	0.85334	14669931	21.14669931	0	0.85334	14669931	21.14669931	-0.014	47036019	21.47036	
22	0	0.85701	16855618	22.16855618	0	0.85701	16855618	22.16855618	-0.008	38522887	22.385229	

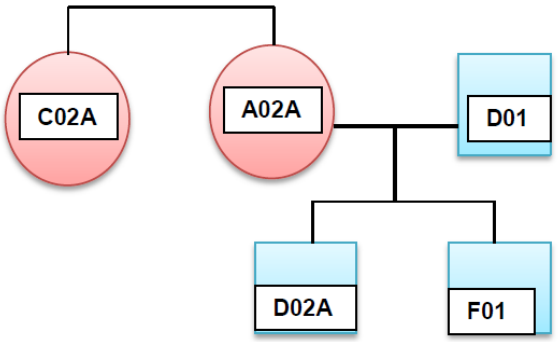
proportion of linkage families since it is only one the proportion is between 0 and 1

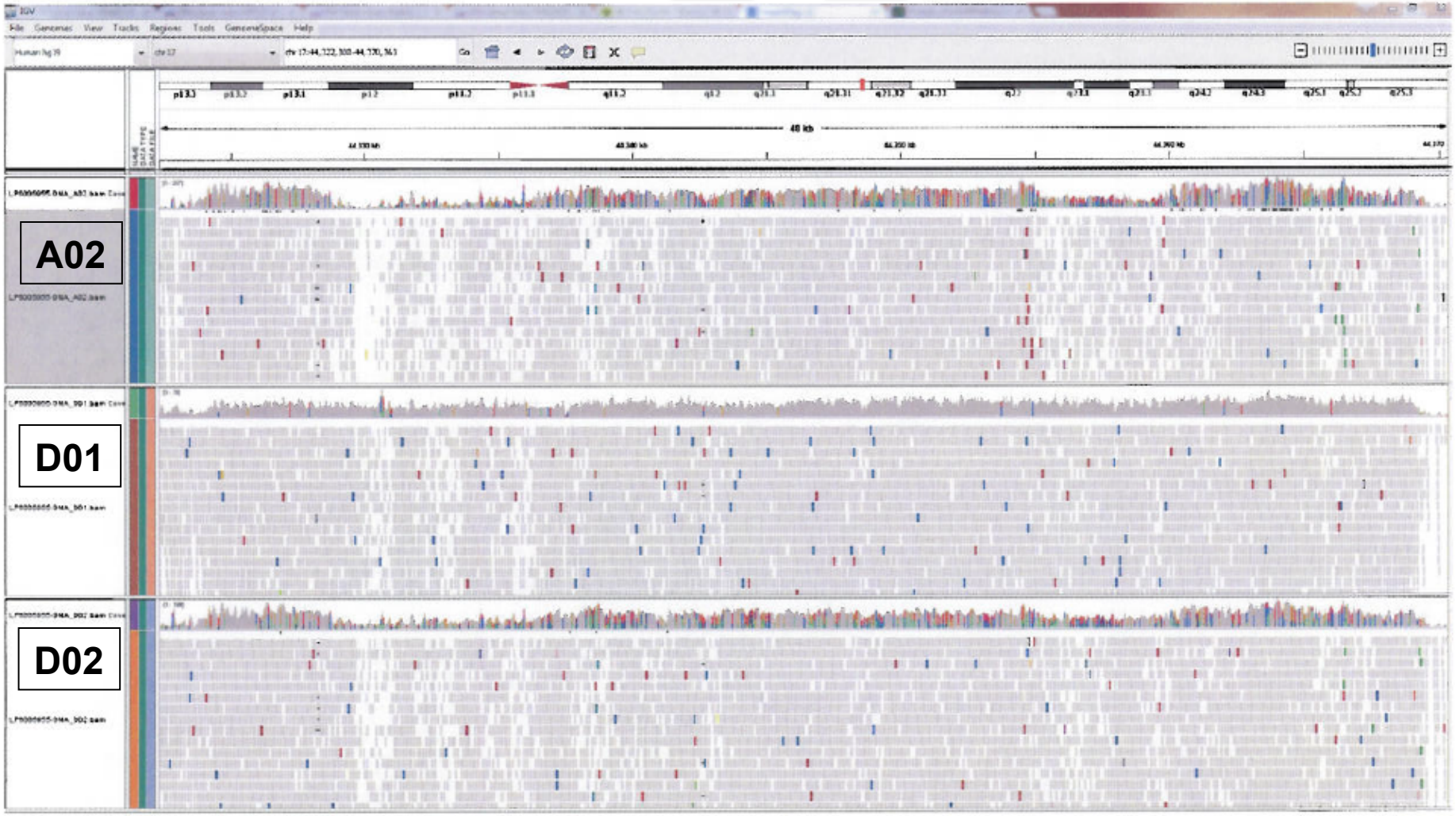
model assumes prevalence of disease of 0.0001, and a dominant model with 0.90 penetrance

MAYO CLINIC

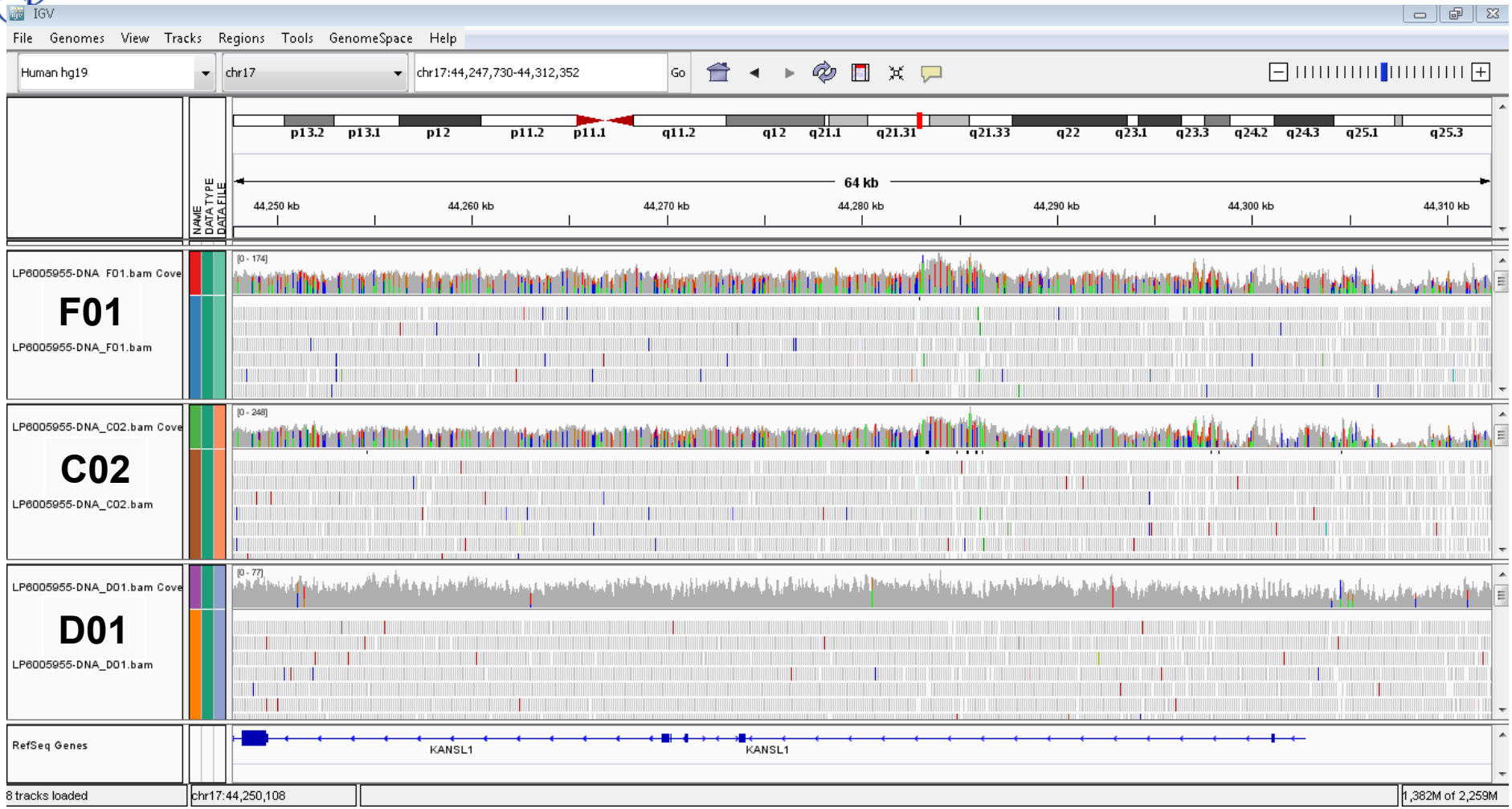

Dominant Model, Chromosome=8
Dominant Model, Chromosome=17
Dominant Model, Chromosome=15
LOD
Position
Position
Position
1
1
1
2
0
−2
−4
−6
1
0
−1
−2
−3
−4
12726013
39430955
66135897
92840839
119545781
146250723
25640
16230457
32435274
48640092
64844909
81049726
20191741
36626359
53060978
69495596
85930215
102364833

MAYO CLINIC





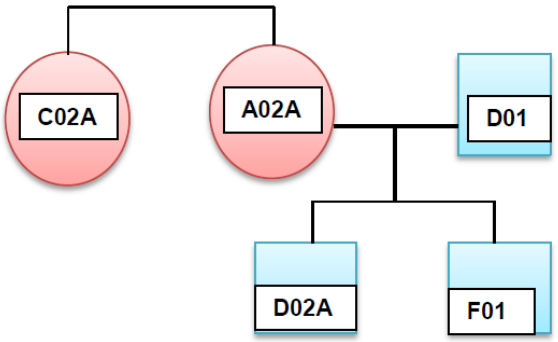
MAYO CLINIC

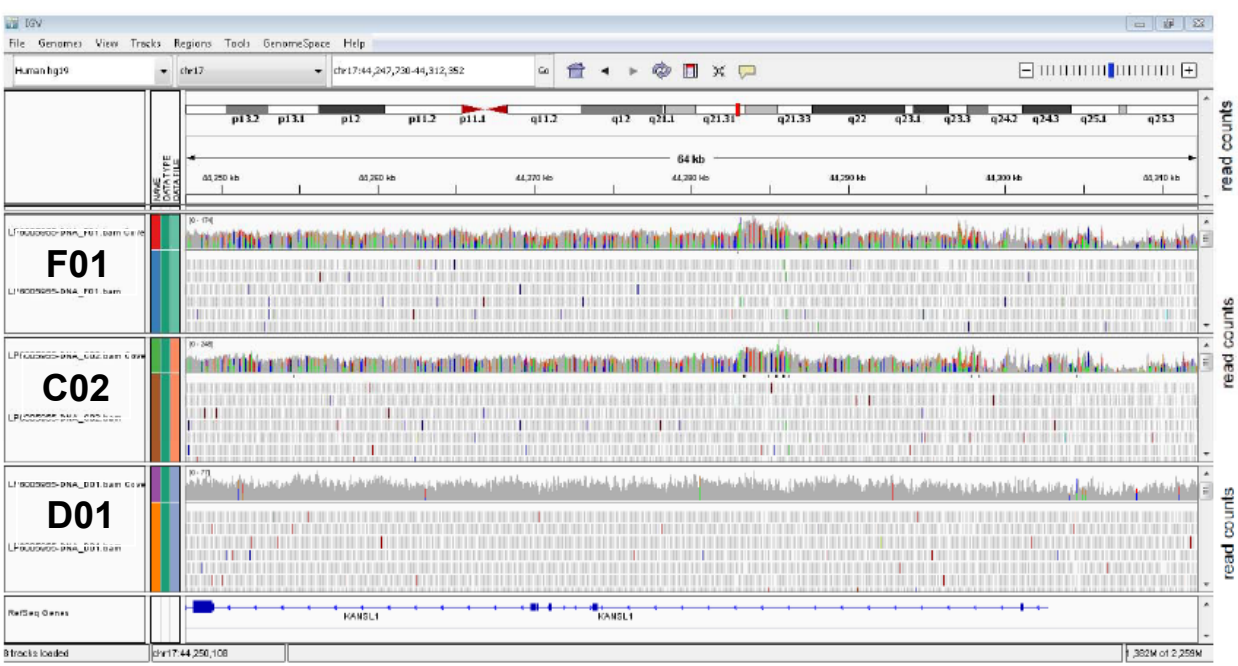


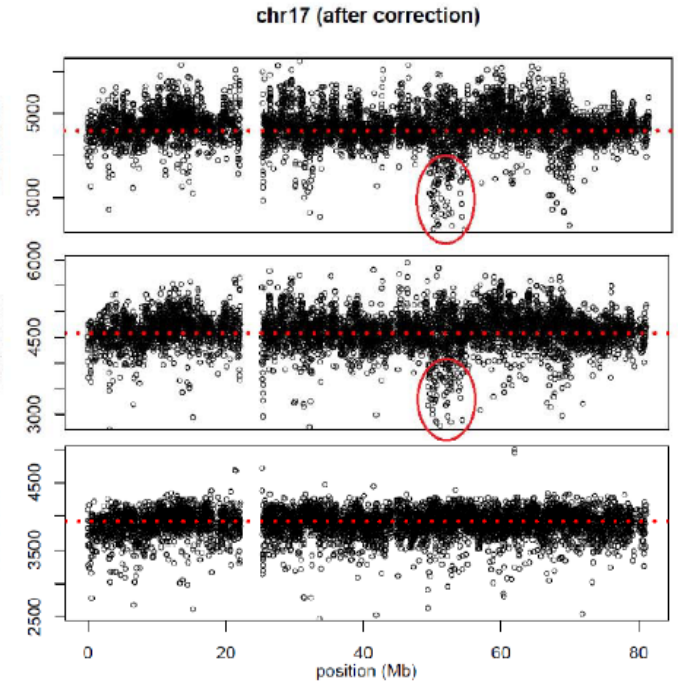
The 17q21.31 haplotype plays a role in the affected members in the family except for the D01 member that is the control for this family. This is a megabase-long inversion polymorphism has many uncharacterized CNVs and markers that are associate with female fertility, meiotic recombination and neurological disease. This inverted H2 form of 17q21.31 appears to be positively selected in Europeans. Nine segregating structural form of 17q21.31 were identified. The independently derived, partial duplications of the *KANSL1* gene; and their allele frequencies have risen to high allele frequency (26% and 19%) in Europeans.

Boettger LM et al. Nat Genet.;44(8):881-885.

MAYO CLINIC



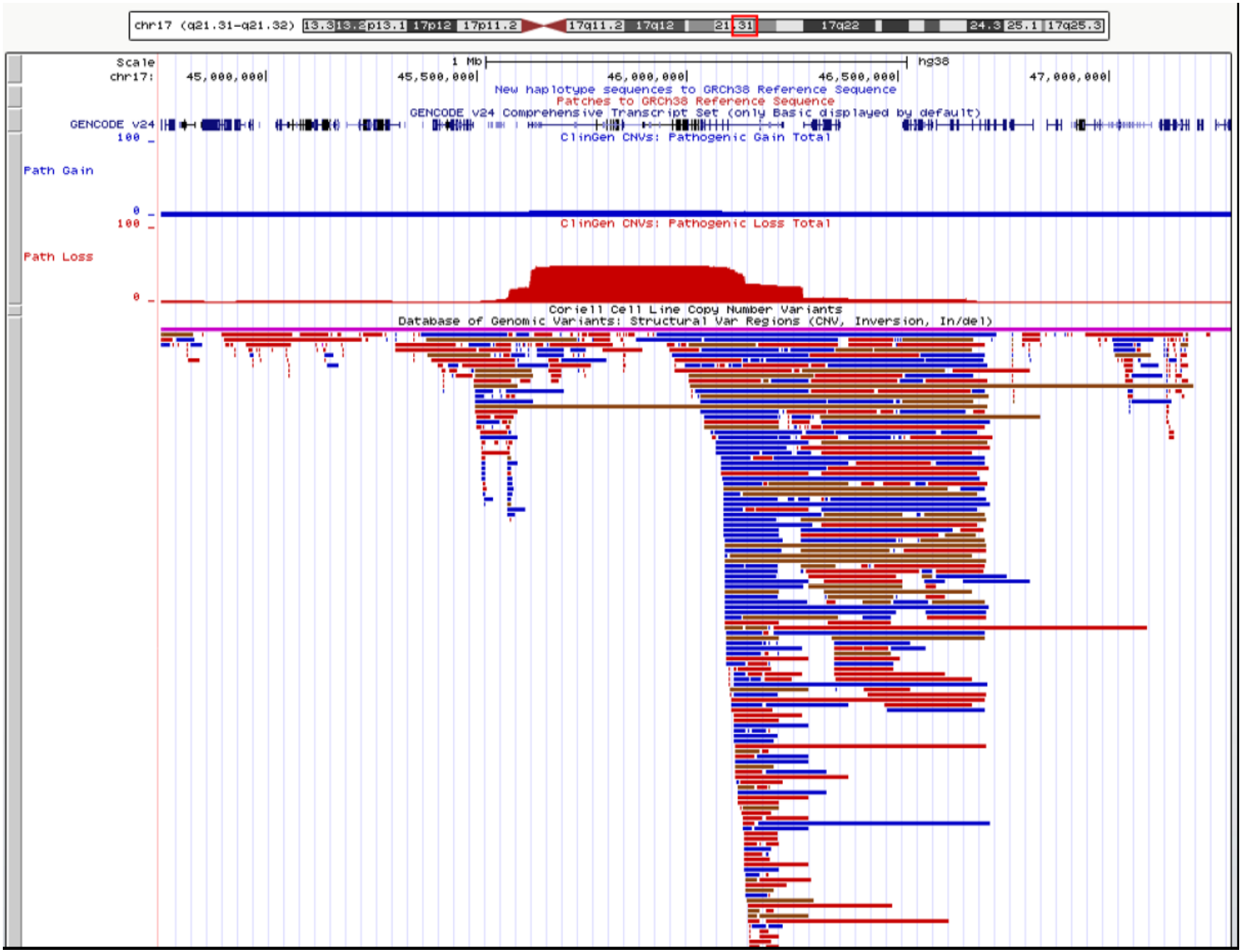


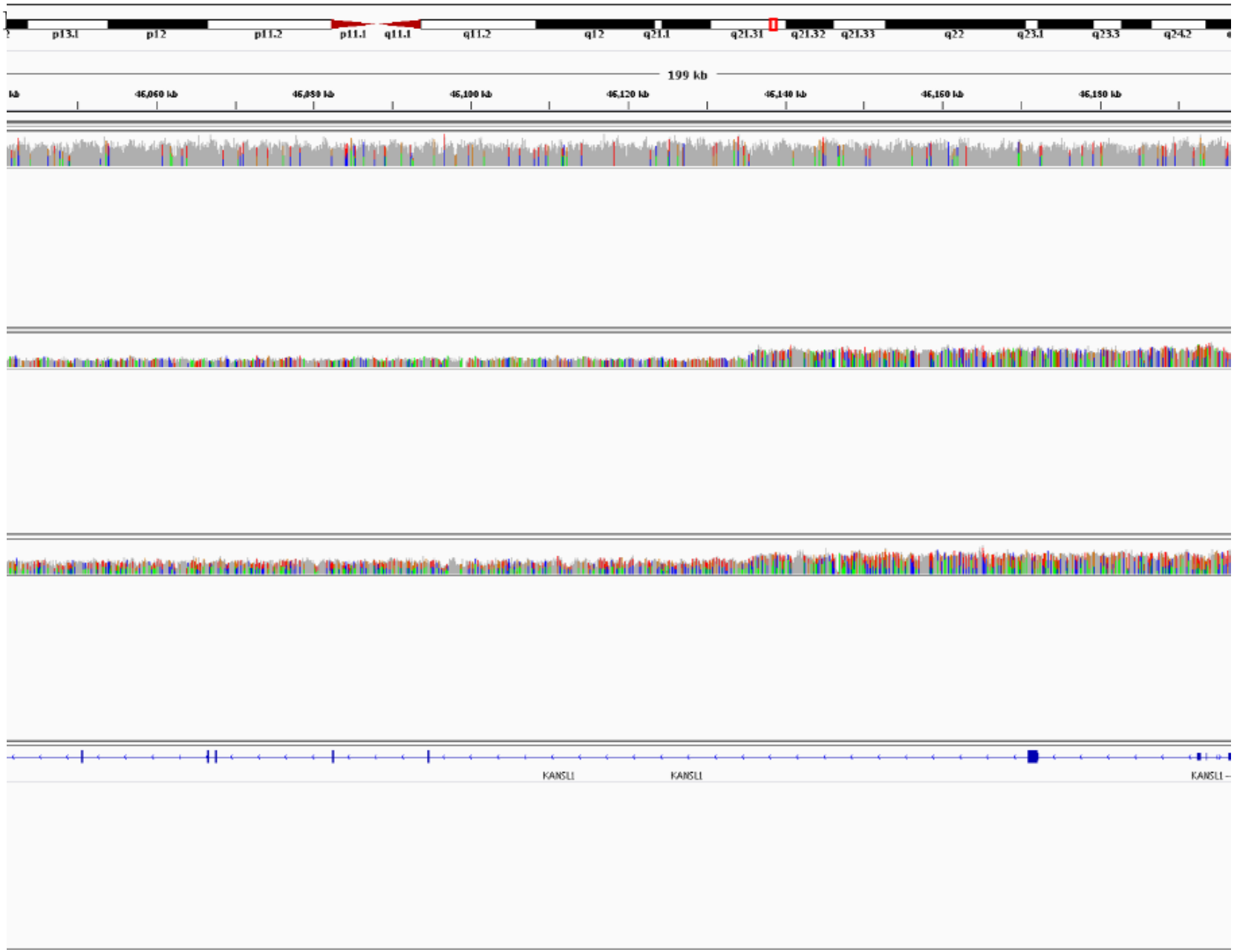


FO1 is the son of D01, both do not have VTE, but F01 is carrier from the mother, A02. He has CNV event without loss of heterozygosity (he has the deletion from the mother and normal from the father, D01). The sisters, C02 and A02, have the deletion in 17q21.31. Both of the, have CNV with loss of heterozygosity.

Across the TOPMed VTE samples 4.7% have CNV similar to C02 and 34.9% similar to F01.

MAYO CLINIC


chr17 (q21.31-q21.32)
13.3 13.2 p13.1 17p12 17p11.2 17q11.2 17q12 21.31 17q22 24.3 25.1 17q25.3
Scale
chr17:
1 Mb
hg38
45,000,000
45,500,000
46,000,000
46,500,000
47,000,000
New haplotype sequences to GRCh38 Reference Sequence
Patches to GRCh38 Reference Sequence
GENCODE v24 Comprehensive Transcript Set (only Basic displayed by default)
GENCODE v24
ClinGen CNVs: Pathogenic Gain Total
100
Path Gain
0
100
ClinGen CNVs: Pathogenic Loss Total
Path Loss
0
Coriell Cell Line Copy Number Variants
Database of Genomic Variants: Structural Var Regions (CNV, Inversion, In/del)



In the MayoTOPMed samples, 39.6% has similar feature as the middle and bottom panels (n=1355).

MAYO CLINIC

Chromosome 17 Abnormality:
There is a significant loosely defined hyper-variant region on chromosome 17 spanning at least 10 million base pairs (chr17:35,000,000-50,000,000). This region is "loosely" defined because no clear breakpoints exist to show the exact duplication event. Suggesting that this might have been one event in the past that integrated into the human genome and over evolutionary time has been sporadically repaired. A more likely explanation however, is that these affected patients have an additional large piece of chromosome 17 or several additional smaller sporadic pieces of chromosome 17 free floating in their nucleus. The genetic implications of such scenarios should be examined by investigating these patients' RNA and protein expression profiles.

This image highlights the hyper-variant observation mentioned above. There are 3 samples shown in this figure, the top sample and bottom samples are affected family members. Notice the rainbow colors within the coverage plots and the y-axis for each coverage plot. The unaffected family member has half the coverage and far less variants than the affected family members. Also the affected family members have many variant allele frequencies that do not appear 50-50 but rather 33-66, further suggesting they have more than 2 alleles for this region. This image is only ~48kb of the 10 million base pairs impacted but this evidence is throughout the larger impacted region.

MAYO CLINIC

SNVs:

A total of 3,380 SNVs are only enriched within the affected samples and only within the regions of interest previously defined on chr15, chr17, and chr18. From those SNVs 36 are predicted to have a high or moderate functional impact and 4 of those were also identified previously by VAAST as likely relevant. Here are those 36 SNVs:

#CHROM	POS	REF	ALT
chr15	42028820	A	G
chr15	42134097	C	T
chr15	42744094	A	G
chr15	49285001	G	C
*chr17	37840860	G	C
chr17	38457151	G	A
chr17	38640744	C	T
chr17	38645125	A	G
*chr17	38910115	T	A
chr17	39023396	G	A
chr17	39084504	T	C
chr17	39135207	A	G
chr17	39137297	C	T
chr17	39140272	A	G
chr17	39140272	A	G
chr17	39155969	A	G
chr17	39156084	T	C
chr17	39186168	A	C
chr17	39884065	C	A
chr17	40179048	G	C
chr17	40255743	T	C
chr17	40714804	C	A
chr17	40818699	A	G
chr17	41131645	C	A
chr17	41245471	C	T
chr17	41738823	G	A
chr17	42828484	T	A
chr17	42849660	G	A
chr17	44714881	C	A
chr17	45360730	T	C
chr17	48155425	G	A
*chr17	48559722	G	C
chr17	48614426	G	A
*chr17	48626804	A	T
chr17	49371340	T	C
chr18	5145609	G	T

The variants highlighted in yellow show less than 0.1% prevalence within 1kgenomes and are not present within the biobank WES samples. The variants with asterisks were the 4 also identified independently by VAAST. The genes of interest from the SNVs are mainly these; PGAP3, KRT25, RSAD1, SPATA20, and BRAC1. Specifically the gene, PGAP3, seems most interesting because it is known to play a role in hyperphosphatasia.

MAYO CLINIC

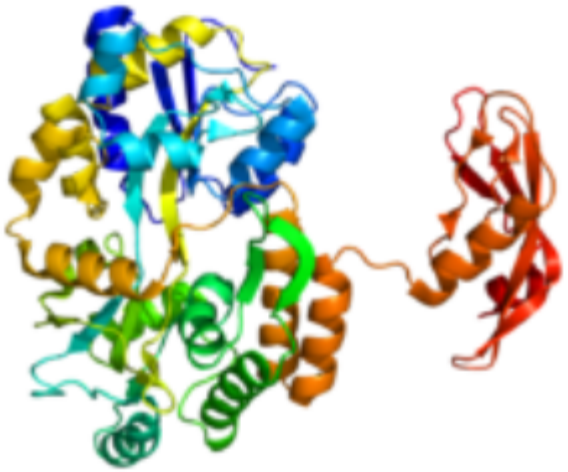
InDels and SVs:

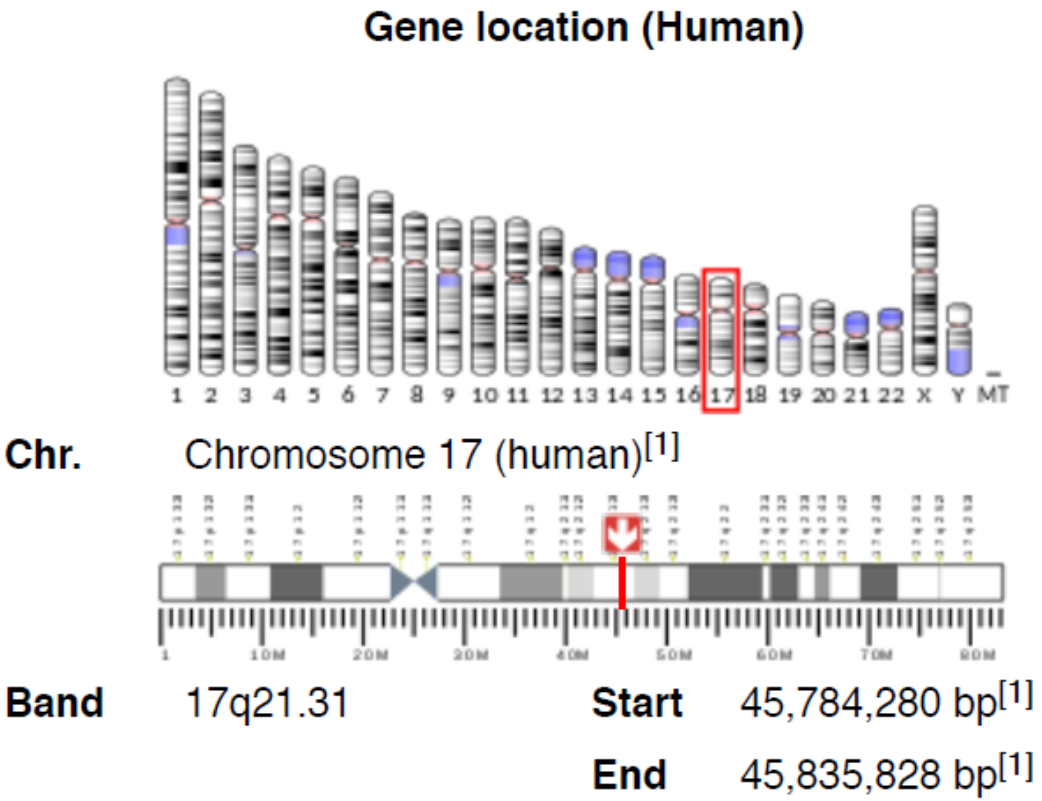
A total of 264 breakpoints from SoftSearch were examined from the previously defined regions of interest. From those regions 1 moderately sized deletion and 5 small InDels were found overlapping regulatory elements. These regulatory elements were defined by HaploReg and Encode. These events appear to be all heterozygous and are all only enriched within the affected samples. Here are those 6 events:

chr15:49,544,846-49,547,879	4.5kb Deletion	GALK2
chr15:48,622,608	InDel	DUT
chr17:46,531,619	InDel	lncRNA -> SKAP1 / HOXB1
chr17:43,908,574	InDel	CRHR1
chr17:40,719,893	InDel	MLX
chr17:39,973,196	InDel	FKBP10

MAYO CLINIC

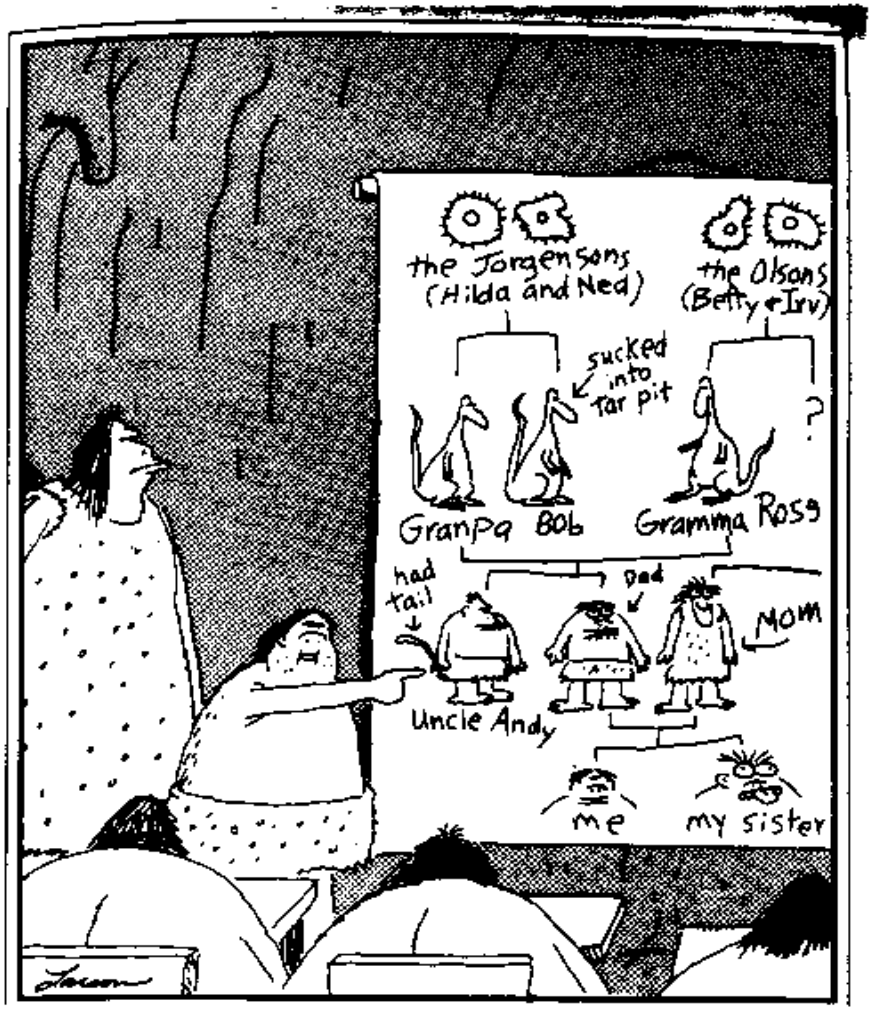
Corticotropin-releasing hormone receptor 1 (CRHR1)





The human CRHR1 contains 14 exons over 20kb of DNA, and its full gene product is a peptide composed of 44 amino acids. This gene is alternatively spliced into a series of variants, which are generated through deletion of one of the 14 exons that may causes a frame-shift in the open reading frame enconde corresponding isoforms of CRF1.

MAYO CLINIC



Dirk brings his family tree to class.