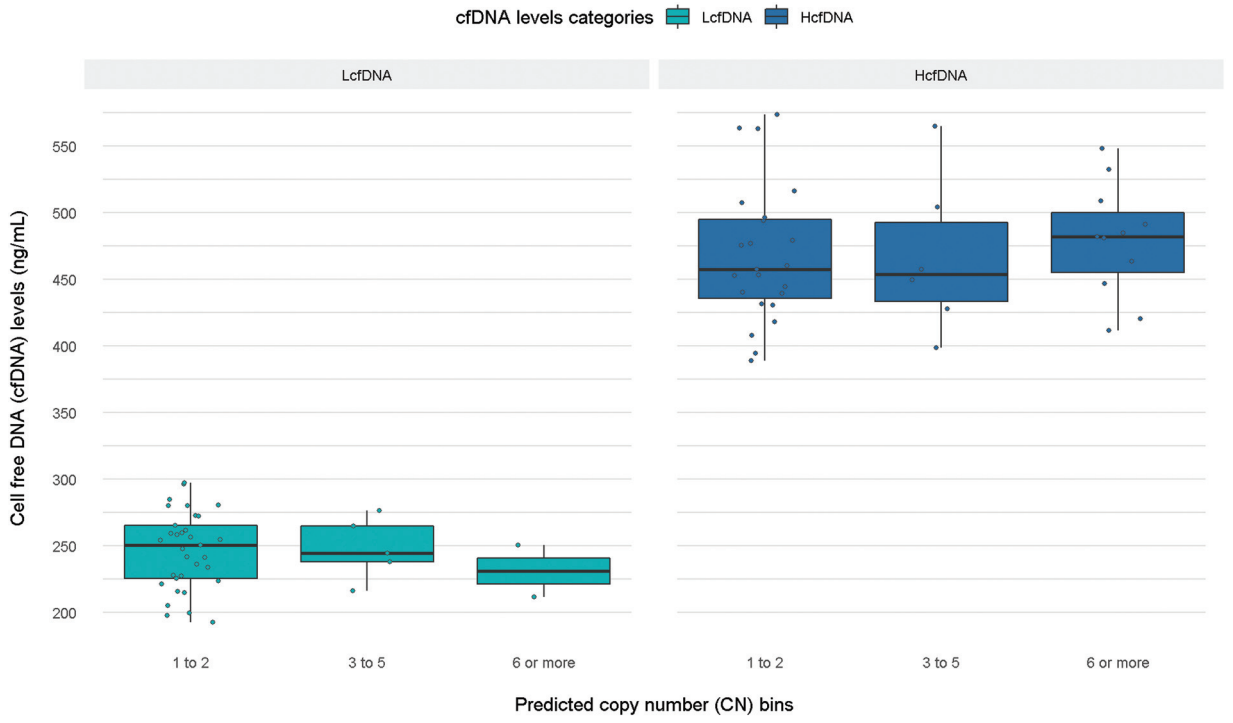
Supplementary Methods

Bioinformatic Analyses

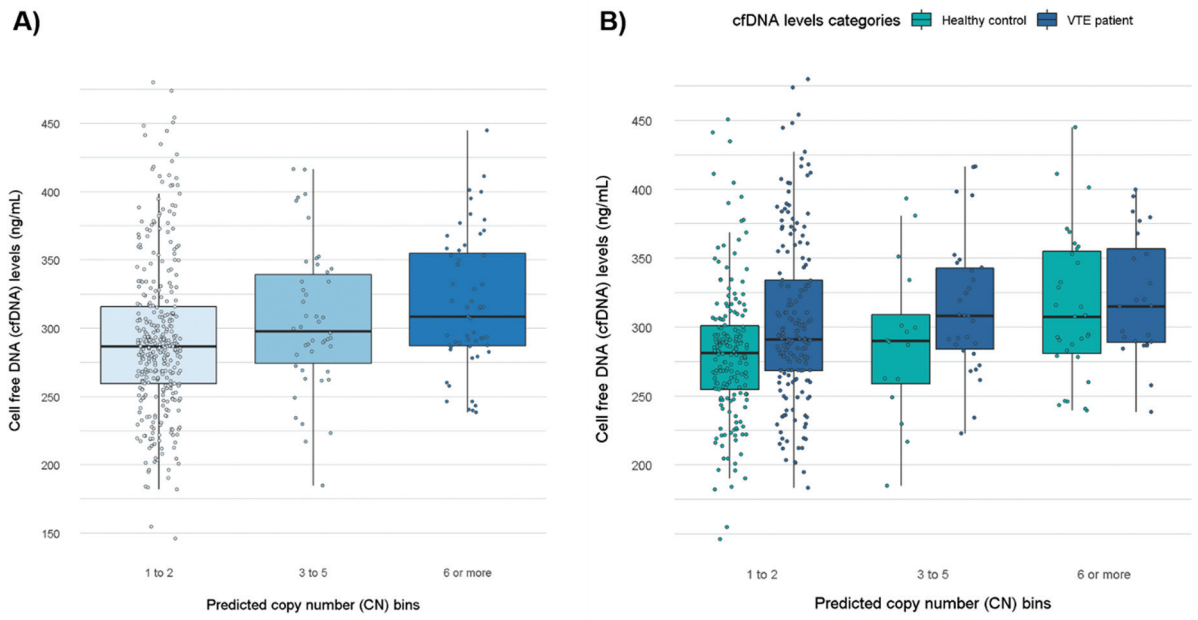
Indexed sequences were de-multiplexed and analyzed individually. Raw paired-end data (FASTQ format) were analyzed using CLC Genomic Workbench version 12 software (CLC Bio—Qiagen, Aarhus, Denmark). First, raw data were trimmed with ambiguous nucleotide (maximum, 2) and quality score (limit, 0.05) filters. An automatic read-through adapter trimming was also performed. Then, trimmed reads were aligned against the human genome reference (GRCh37/hg19). Read mapping was performed with specific parameter settings (match score, 1; mismatch cost, 2; indel cost, 3; length fraction, 0.7; similarity fraction, 0.9). Finally, the Basic Variant Detection tool was used for genetic variant calling (ignore broken pairs, yes; minimum coverage, 30 $\times$; minimum count, 2; minimum variant frequency, 25%; relative read direction filter, yes; significance, 1%).

Assessment of Putative Pathogenic Mutations

The parameters used to identify putative pathogenic mutations were as follows: (1) variant allele frequency $\leq$1% in all populations from the 1000 Genomes panel (May 2013 Phase 3 v5a),[1] (2) location referring to Variant Effect Predictor version 84 database[2] (excluded: intronic mutations located >10 bp from the flanking intronic regions of each exon, upstream, downstream and UTR variants, and synonymous variants not regulating splicing regions), and (3) minor allele frequency in our sample $\leq$5%.



Supplementary Fig. S1 Boxplot of the relationship between cfDNA levels and the mCNV in the *ORM1* locus in LcfDNA individuals ($n = 40$) and HcfDNA individuals ($n = 40$) of the discovery sample from the GAIT-2 Project. The *X*-axis corresponds to each CN bin of the mCNV, displayed in increasing order of dosage, and the *Y*-axis shows plasma cfDNA levels. Box midlines represent the median of the data, and the upper and lower limits of the box, the third and first quartiles (75th and 25th percentiles), respectively. Whiskers extend up to 1.5 times the interquartile range from the top/bottom of the box to the farthest datum within that distance. cfDNA, cell-free DNA; CN, copy number; HcfDNA, high cfDNA; LcfDNA, low cfDNA.



Supplementary Fig. S2 Boxplots of the relationship between cfDNA levels and the mCNV in the ORM1 locus in (A) the whole age- and sex-matched subset from the RETROVE project ($n = 460$) and (B) controls ($n = 230$) and VTE cases ($n = 230$) of the age- and sex-matched subset from the RETROVE Project. The *X*-axis corresponds to each CN bin of the mCNV, displayed in increasing order of CN, and the *Y*-axis shows plasma cfDNA levels. Box midlines represent the median of the data, and the upper and lower limits of the box represent the third and first quartiles (75th and 25th percentiles), respectively. Whiskers extend up to 1.5 times the interquartile range from the top/bottom of the box to the farthest datum within that distance; any data beyond that distance are represented individually as outlier points. cfDNA, cell-free DNA; CN, copy number; mCNV, multiallelic copy number variation; VTE, venous thromboembolism.

Supplementary Table S1 Primers used for Long-Range PCR amplification of the *ORM1* locus

LR-PCR	Forward primer	Forward primer position[a]	Reverse primer	Reverse primer position[a]	Size (bp)
ORM1-LR1	CTGCTCTCTGACAC TTATCATCAGAGG	chr9:117,082, 801–117,082,827	GGCCTAGGTGTGT CAGTGGCCACAC	chr9:117,086, 975–117,086,999	4,199
ORM1-LR2	GGAGCTGATGGAG CTGAACTGACAC	chr9:117,086, 757–117,086,781	GTGGCCCATTTGC GTTCCTACCTTAT	chr9:117,091, 348–117,091,373	4,617

[a]Genomic positions were based on the GRCh37/hg19 assembly.

Supplementary Table S2 PCR master mix for Long-Range PCR amplifications

LR-PCR components	Volumes for ORM1-LR1 (μL)	Volumes for ORM1-LR2 (μL)
SequalPrep 10X Reaction Buffer	2	2
DMSO	0.4	0.4
SequalPrep 10X Enhancer A	2	2
SequalPrep Long Polymerase 5 U/μL	0.36	0.36
Primer mix (6.5 μM)	2	1
Template DNA (50 ng/μL)	1	1
DNase-free water	12.24	13.24

Supplementary Table S3 Thermocycling conditions for Long-Range PCR amplifications

Thermal cycler steps	Temperature (°C)	Time	Cycles
Initial denaturation	94	2 minutes	1
Denaturation	94	10 seconds	×10
Annealing	60	30 seconds	
Elongation	68	18 minutes	
Denaturation	94	10 seconds	×20
Annealing	60	30 seconds	
Elongation	68	18 minutes (+ 20 seconds/cycle)	
Final elongation	72	5 minutes	1
Cooling	4	∞	

Supplementary Table S4 Thrombin generation parameters in the replication cohort

Trait	Mean in cases (Q1–Q3)	Mean in controls (Q1–Q3)
Lag time (minutes)	3.00 (2.67–3.66)	2.67 (2.33–3.33)
Time to peak (minutes)	6.50 (5.67–7.54)	6.33 (5.33–7.46)
Thrombin peak (nM)	242.41 (190.11–299.66)	224.26 (175.02–282.30)
ETP (nM*min)	1,461.51 (1,240.70–1,744.88)	1,346.61 (1,169.62–1,607.92)

Note: Q1–Q3: first and third quartiles.

Supplementary Table S5 Genetic variants detected by NGS in the *ORM1* locus

Genomic position[a]	Reference allele	Alternate allele	Gene location	dbSNP v155	Minor allele	MAF (%)[b]
chr9:117,083,180	G	A	Upstream	rs117036474	A	0.6
chr9:117,083,197	C	T	Upstream	rs543858329	T	0.6
chr9:117,083,493	A	G	Upstream	rs140189969	G	0.6
chr9:117,083,651	C	T	Upstream	rs7032449	T	20.6
chr9:117,083,725	A	G	Upstream	rs72746707	G	1.9
chr9:117,083,803	C	A	Upstream	rs150611042	A	14.4
chr9:117,083,806	A	G	Upstream	–	G	0.6
chr9:117,083,822	T	C	Upstream	–	C	0.6
chr9:117,083,830	A	G	Upstream	rs1030727886	G	0.6
chr9:117,083,848	CT	TGC	Upstream	–	TGC	0.6
chr9:117,083,862	A	G	Upstream	rs1444506598	G	0.6
chr9:117,083,885	T	C	Upstream	rs59980189	C	0.6
chr9:117,083,893	C	T	Upstream	rs182657892	T	0.6
chr9:117,083,906	T	A	Upstream	rs35240061	A	20.6
chr9:117,083,930	T	C	Upstream	–	C	1.2
chr9:117,083,947	A	G	Upstream	rs58791119	G	0.6
chr9:117,083,962	C	G	Upstream	–	G	1.9
chr9:117,083,968	T	C	Upstream	–	C	1.2
chr9:117,083,970	G	A	Upstream	rs187003700	A	0.6
chr9:117,083,976	A	G	Upstream	rs1391834027	G	1.2
chr9:117,083,992	A	G	Upstream	rs56742956	G	1.2

Supplementary Table S5 (*Continued*)

Genomic position[a]	Reference allele	Alternate allele	Gene location	dbSNP v155	Minor allele	MAF (%)[b]
chr9:117,084,020	G	A	Upstream	rs11791193	A	15.6
chr9:117,084,031	C	T	Upstream	rs10982149	T	20.0
chr9:117,084,032	G	A	Upstream	rs972794911	A	0.6
chr9:117,084,228	C	T	Upstream	rs2787337	T	20.6
chr9:117,084,275	G	A	Upstream	rs1017989	A	17.5
chr9:117,084,294	G	A	Upstream	rs1478367873	A	0.6
chr9:117,084,324	A	T	Upstream	–	T	0.6
chr9:117,084,390	A	T	Upstream	–	T	0.6
chr9:117,084,402	G	T	Upstream	–	T	0.6
chr9:117,084,447	C	G	Upstream	rs1738036077	G	0.6
chr9:117,084,462	C	T	Upstream	rs187920764	T	0.6
chr9:117,084,530	T	A	Upstream	rs1270130504	A	0.6
chr9:117,084,672	G	A	Upstream	rs116994374	A	14.4
chr9:117,084,722	G	A	Upstream	rs188924106	A	14.4
chr9:117,084,846	G	A	Upstream	rs890612897	A	0.6
chr9:117,084,855	A	T	Upstream	rs4510934	A	19.4
chr9:117,084,905	T	C	Upstream	rs1766076	C	22.5
chr9:117,084,913	T	C	Upstream	rs34272708	C	0.6
chr9:117,084,971	C	G	Upstream	rs140041983	G	14.4
chr9:117,084,988	TG	T	Upstream	rs201940520	T	4.4
chr9:117,085,019	C	T	Upstream	rs532735199	T	0.6
chr9:117,085,125	T	C	Upstream	rs143022373	C	11.2
chr9:117,085,186	C	G	Upstream	rs550478926	G	14.4
chr9:117,085,526	G	A	Exon 1	rs17650	G	41.2
chr9:117,085,549	G	A	Intron 1	rs374569633	A	1.2
chr9:117,085,635	T	C	Intron 1	rs1687381	T	41.2
chr9:117,086,035	C	T	Exon 2	rs700126	T	0.6
chr9:117,086,188	G	A	Intron 2	rs117221569	A	3.1
chr9:117,086,195	C	T	Intron 2	rs563773944	T	15.0
chr9:117,086,271	T	G	Intron 2	rs10982151	G	18.8
chr9:117,086,489	T	C	Intron 3	rs182101558	C	1.2
chr9:117,086,549	A	G	Intron 3	rs117428165	G	0.6
chr9:117,086,602	TTAGAAC	T	Intron 3	rs536264118	T	1.2
chr9:117,086,847	T	G	Intron 3	rs71503591	G	16.2
chr9:117,087,212	T	A	Intron 4	rs10982154	A	8.1
chr9:117,087,254	C	T	Intron 4	rs10982155	T	8.8
chr9:117,087,412	G	A	Exon 5	rs1126801	A	0.6
chr9:117,087,655	G	A	Intron 5	rs139920377	A	23.1
chr9:117,087,667	G	A	Intron 5	rs2070869	A	22.5
chr9:117,087,705	A	G	Intron 5	rs530319527	G	4.4
chr9:117,087,745	G	A	Intron 5	rs145835687	A	10.0
chr9:117,087,785	A	G	Intron 5	rs1766093	G	31.9
chr9:117,087,997	A	C	Intron 5	–	C	0.6

(*Continued*)

Supplementary Table S5 (*Continued*)

Genomic position[a]	Reference allele	Alternate allele	Gene location	dbSNP v155	Minor allele	MAF (%)[b]
chr9:117,088,064	T	A	Intron 5	rs10982156	A	8.8
chr9:117,088,076	T	C	Intron 5	rs2932682	C	24.4
chr9:117,088,120	T	C	Intron 5	rs1687384	T	42.5
chr9:117,088,134	C	G	Intron 5	rs115455945	G	0.6
chr9:117,088,155	A	G	Intron 5	rs118127002	G	3.8
chr9:117,088,166	A	T	Intron 5	rs10982158	T	17.5
chr9:117,088,177	C	T	Intron 5	rs1687385	T	3.8
chr9:117,088,197	C	T	Intron 5	rs1687386	T	5.0
chr9:117,088,311	T	C	Intron 5	rs184220363	C	3.8
chr9:117,088,313	G	T	Intron 5	rs190378866	T	3.1
chr9:117,088,320	C	A	Intron 5	rs796663400	A	3.8
chr9:117,088,324	A	T	Intron 5	rs796197661	T	3.8
chr9:117,088,329	A	T	Intron 5	rs541629030	T	3.8
chr9:117,088,345	A	G	Intron 5	rs796580277	G	3.8
chr9:117,088,354	A	G	Intron 5	rs2787340	G	3.8
chr9:117,088,504	T	C	Intron 5	rs1766097	C	14.4
chr9:117,088,668	TTGG	T	3′UTR	rs377324887	T	0.6
chr9:117,088,808	C	T	Downstream	rs1687387	T	10.6
chr9:117,088,961	A	G	Downstream	rs62559488	A	38.1
chr9:117,088,987	A	G	Downstream	rs879564860	G	15.6
chr9:117,089,183	C	A	Downstream	rs1439431972	A	0.6
chr9:117,089,348	C	CTG	Downstream	rs796174206	C	0.6
chr9:117,089,626	G	A	Downstream	rs749748577	A	0.6
chr9:117,089,833	T	C	Downstream	rs1687389	C	8.1
chr9:117,089,888	G	A	Downstream	rs1687390	A	8.8
chr9:117,089,921	G	T	Downstream	rs146077223	T	8.1
chr9:117,089,960	G	A	Downstream	rs1311594723	A	0.6
chr9:117,089,968	C	T	Downstream	rs569347615	T	0.6
chr9:117,089,971	C	T	Downstream	rs2636883	T	0.6
chr9:117,090,006	C	T	Downstream	rs1687391	T	8.8
chr9:117,090,042	T	C	Downstream	rs11788739	C	17.5
chr9:117,090,050	T	C	Downstream	rs536985980	C	0.6
chr9:117,090,376	C	G	Downstream	rs3762057	G	18.1
chr9:117,090,434	C	T	Downstream	rs3762056	T	8.8
chr9:117,090,575	T	C	Downstream	rs3762055	C	9.4
chr9:117,090,902	G	A	Downstream	rs547538130	A	1.9
chr9:117,091,002	C	T	Downstream	rs112164771	T	7.5
chr9:117,091,033	G	T	Downstream	rs10982163	T	8.1
chr9:117,091,067	G	A	Downstream	rs10982164	A	7.5
chr9:117,091,074	C	T	Downstream	rs7040440	T	10.6
chr9:117,091,169	G	A	Downstream	rs7040680	A	16.9

Abbreviation: MAF, minor allele frequency.
[a]Genomic positions were based on the human GRCh37/hg19 assembly.
[b]The minor allele frequency in our discovery sample of 80 individuals from the GAIT-2 project.

Supplementary Table S6 Predicted copy numbers of *ORM1* and *ORM2* loci in the 80 individuals from the GAIT-2 Project

Gene	Predicted copy numbers	*n* with LcfDNA (%)[a]	*n* with HcfDNA (%)[a]	*n* total (%)	Confidence value	Z-score
ORM1	2	33 (82.50%)	23 (57.50%)	56 (70.00%)	>0.99	0.02–1.53
ORM1	3	4 (10.00%)	1 (2.50%)	5 (6.25%)	>0.99	0.24–0.75
ORM1	4	1 (2.50%)	4 (10.00%)	5 (6.25%)	>0.99	0.03–0.49
ORM1	5	0 (0.00%)	1 (2.50%)	1 (1.25%)	0.71	1.52
ORM1	6	1 (2.50%)	0 (0.00%)	1 (1.25%)	<0.50	1.29
ORM1	7	1 (2.50%)	4 (10.00%)	5 (6.25%)	<0.50–0.90	0.04–1.31
ORM1	8	0 (0.00%)	5 (12.50%)	5 (6.25%)	<0.50–0.91	0.10–0.60
ORM1	9	0 (0.00%)	2 (5.00%)	2 (2.50%)	<0.50–0.83	0.18–0.74
ORM2	1	1 (2.50%)	0 (00.00%)	1 (1.25%)	>0.99	0.00
ORM2	2	38 (95.00%)	38 (98.00%)	76 (95.00%)	>0.99	0.02–2.59
ORM2	3	1 (2.50%)	2 (5.00%)	3 (3.75%)	>0.99	0.16–1.03

Abbreviations: HcfDNA, high cell-free DNA levels; LcfDNA, low cell-free DNA levels; *N,* number of individuals.
[a]The extreme sampling design included 40 individuals with LcfDNA and 40 individuals with HcfDNA.

Supplementary Table S7 Predicted copy numbers of *ORM1* in the replication cohort

Predicted copy number of *ORM1*	*N* cases (%)[a]	Mean age of cases (y)	*N* controls (%)[a]	Mean age of controls (y)	*N* total (%)	Mean age (y)	Confidence value	Z-score
1	0 (0.00)	–	1 (0.25)	37	1 (0.13)	37	>0.99	0.01
2	302 (75.50)	62	305 (76.25)	47	607 (75.88)	54	0.98 to >0.99	0.01–2.23
3	25 (6.25)	60	17 (4.25)	49	42 (5.25)	55	0.54 to >0.99	0.00–2.11
4	14 (3.50)	63	5 (1.25)	48	19 (2.38)	59	0.89 to >0.99	0.00–1.87
5	12 (3.00)	62	9 (2.25)	55	21 (2.63)	59	<0.50 to >0.99	0.01–1.70
6	9 (2.25)	59	7 (1.75)	43	16 (2.00)	52	<0.50 to >0.99	0.00–1.55
7	19 (4.75)	63	27 (6.75)	47	46 (5.75)	54	<0.50 to >0.99	0.00–1.39
8	13 (3.25)	65	17 (4.25)	40	30 (3.75)	51	<0.50–0.96	0.03–1.07
9	2 (0.50)	53	3 (0.75)	43	5 (0.63)	47	<0.50–0.70	0.07–0.74
10	2 (0.50)	49	1 (0.25)	43	3 (0.38)	47	<0.50–0.86	0.03–0.99
11	0 (0.00)	–	4 (1.00)	43	4 (0.5)	43	<0.50–0.69	0.04–0.61
12	1 (0.25)	74	2 (0.50)	59	3 (0.38)	64	<0.50–0.70	0.05–0.33
13	0 (0.00)	–	1 (0.25)	93	1 (0.13)	93	<0.50	0.13
14	1 (0.25)	85	0 (0.00)	–	1 (0.13)	85	<0.50	0.15
15	0 (0.00)	–	1 (0.25)	22	1 (0.13)	22	<0.50	0.13

Abbreviation: *N*, number of individuals; y, years.
[a]The replication cohort RETROVE included 400 VTE cases and 400 controls.

References

1. The 1000 Genomes Project Consortium; Auton A, Brooks LD, Durbin RM, et al. A global reference for human genetic variation. Nature 2015;526:68–74
2. McLaren W, Gil L, Hunt SE, et al. The Ensembl variant effect predictor. Genome Biol 2016;17(1):122

