

Генетическая номенклатура, введение

Екатерина Юрьевна Печёрина
Лаборатории ЦИР
Москва, январь 2024 г.

Ген

- HGNC (Hugo Gene Nomenclature Committee) сохраняет определение гена как «сегмента ДНК, который вносит вклад в фенотип/функцию». При отсутствии продемонстрированной функции ген может быть охарактеризован по последовательности, транскрипции или гомологии».

Ген

- В 1957 году международный комитет опубликовал рекомендации по генетическим символам и номенклатуре.
- Необходимость разработки официальных руководств по названиям и символам человеческих генов была признана в 1960-х годах, а полные руководства были выпущены в 1979 году (Эдинбургское собрание генома человека).

Альтернативные названия

- Для многих генов и соответствующих им белков в научной литературе и общедоступных биологических базах данных используется множество альтернативных названий, что создает проблему для эффективной организации и обмена биологической информацией.

Связь с номенклатурой белков

- Номенклатура генов и номенклатура белков являются аспектами одного и того же целого. Любое имя или символ, используемый для белка, потенциально может также использоваться для гена, который его кодирует, и наоборот.

Один к одному

- корреляция между генами и белками не всегда однозначна (в любом направлении); в некоторых случаях это соотношение «несколько к одному» или «один к нескольким», и тогда имена и символы могут быть в некоторой степени специфичными для гена или белка или перекрываться в использовании

Разные виды

- Многие механизмы жизни одинаковы или очень похожи у разных видов, родов, отрядов и типов (посредством гомологии и/или аналогии), так что данный белок может производиться во многих видах организмов; поэтому ученые часто используют один и тот же символ и название для данного белка у одного вида (например, мышей), как и у другого вида (например, человека).

F5 coagulation factor V [*Homo sapiens* (human)]

Gene ID: 2153, updated on 7-Jan-2024

Ген фактора 5 у Человека

Download Datasets

Summary

- Official Symbol** F5 provided by [HGNC](#)
 - Official Full Name** coagulation factor V provided by [HGNC](#)
 - Primary source** [HGNC:HGNC:3542](#)
 - See related** [Ensembl:ENSG00000198734](#) [MIM:612309](#); [AllianceGenome:HGNC:3542](#)
 - Gene type** protein coding
 - RefSeq status** REVIEWED
 - Organism** [Homo sapiens](#)
 - Lineage** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Placentalia; Eutheria; Chordata; Craniata; Vertebrata; Mammalia; Placentalia; Eutheria
 - Also known as** FVL; PCCF; THPH2; RPRGL1
 - Summary** This gene encodes an essential cofactor of the blood coagulation cascade of the activation peptide by thrombin during coagulation. This generates protein is a cofactor that participates with activated coagulation factor X recessive hemorrhagic diathesis or an autosomal dominant form of thrombophilia [provided by RefSeq, Apr 2015]
 - Expression** Biased expression in liver (RPKM 48.1), placenta (RPKM 23.5) and 2 other tissues [See more](#)
 - Orthologs** [mouse](#) [all](#)
- NEW** Try the new [Gene table](#)
Try the new [Transcript table](#)

F5 coagulation factor V [*Mus musculus* (house mouse)]

Gene ID: 14067, updated on 23-Nov-2023

Ген фактора 5 у МЫШИ

Download Datasets

Summary

- Official Symbol** F5 provided by [MGI](#)
 - Official Full Name** coagulation factor V provided by [MGI](#)
 - Primary source** [MGI:MGI:88382](#)
 - See related** [Ensembl:ENSMUSG00000026579](#) [AllianceGenome:MGI:88382](#)
 - Gene type** protein coding
 - RefSeq status** REVIEWED
 - Organism** [Mus musculus](#)
 - Lineage** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Myomorpha; Muroidea; Muridae; Murinae; Mus; Mus
 - Also known as** Cf5; Cf-5
 - Summary** This gene encodes a glycoprotein coagulation factor that plays a critical role in the process of blood coagulation and hemostasis. The encoded protein is activated by thrombin, to generate a heterodimer containing heavy and light chains held together by calcium ions. About half of the mice lacking the encoded protein die at an embryonic stage possible due to abnormal yolk-sac vasculature while the remaining animals succumbed to massive hemorrhage immediately after birth. A point mutation in this gene has been shown to cause disseminated intravascular thrombosis in the perinatal period, resulting in frequent deaths of newborn mice. [provided by RefSeq, Apr 2015]
 - Expression** Biased expression in liver adult (RPKM 34.5), liver E18 (RPKM 14.1) and 3 other tissues [See more](#)
 - Orthologs** [human](#) [all](#)
- NEW** Try the new [Gene table](#)
Try the new [Transcript table](#)

Наименования генов

- The HUGO Gene Nomenclature Committee
- Гены позвоночных имеют имена, которые представляют собой короткие идентификаторы (обычно от 3 до 8 символов). Например, ген белка 4, ассоциированного с цитотоксическими Т-лимфоцитами - CTLA4.
- Эти символы обычно создаются путем сокращения или аббревиатуры названия.

Корневой символ

- Корневая часть символов семейства генов (например, корень «SERPIN» в SERPIN1, SERPIN2, SERPIN3 и т. д.) называется корневым символом.

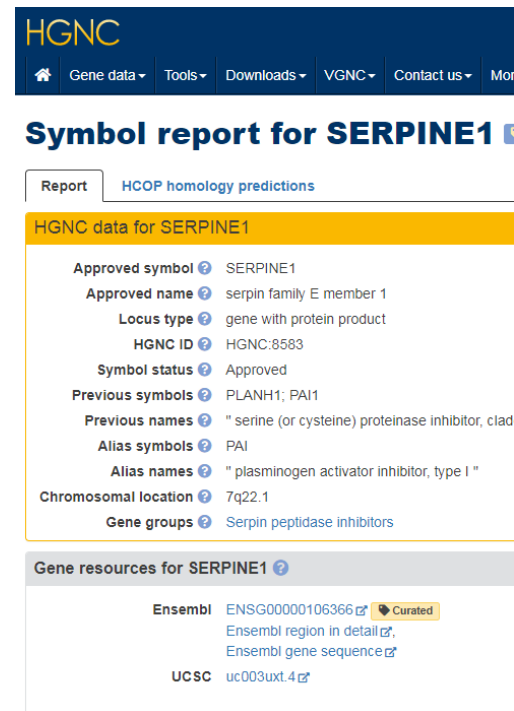
Наименования генов



- Каждому гену, который мы называем, присваивается уникальный символ, идентификатор HGNC (в формате HGNC:#) и описательное имя.
- Символы должны содержать только латинские буквы и арабские цифры
- Символы не должны содержать знаков препинания или буквы «G» для гена.
- Символы не содержат никаких ссылок на вид, в котором они закодированы, например, «H/h» для человека.
- «Генная номенклатура должна развиваться вместе с новыми технологиями, а не быть ограничительной, как это иногда происходит при применении исторических систем номенклатуры отдельных генов»

SERPIN


- Ген *SERPINE1* — белок PAI-1
- Ген *SERPINA1* — белок *alpha-1 antitrypsin*
- Ген *SERPINA3* — белок *SERPINA3, α -1-antichymotrypsin (AACT, ACT)*
- Ген *SERPINC1* — белок *антитромбин III*



The screenshot shows the HGNC Symbol report for SERPINE1. The page has a dark blue header with the HGNC logo and navigation links: Home, Gene data, Tools, Downloads, VGNC, Contact us, and More. Below the header, the title is "Symbol report for SERPINE1". There are two tabs: "Report" (selected) and "HCOP homology predictions". The main content area is titled "HGNC data for SERPINE1" and contains the following information:

Approved symbol	SERPINE1
Approved name	serpin family E member 1
Locus type	gene with protein product
HGNC ID	HGNC:8583
Symbol status	Approved
Previous symbols	PLANH1; PAI1
Previous names	" serine (or cysteine) proteinase inhibitor, clade 1 member 1
Alias symbols	PAI
Alias names	" plasminogen activator inhibitor, type 1 "
Chromosomal location	7q22.1
Gene groups	Serpin peptidase inhibitors

Below this, there is a section titled "Gene resources for SERPINE1" with the following links:

Ensembl	ENSG00000106366  Ensembl region in detail , Ensembl gene sequence
UCSC	uc003uxt.4

HLA

- HLA-DQ
 - α -chain кодируется HLA-DQA1
 - β -chain кодируется HLA-DQB1
- HLA-DR
 - α -chain кодируется HLA-DRA
 - 4 β -chains кодируется HLA-DRB1, DRB3, DRB4, DRB5

HGNC

Gene data ▾ Tools ▾ Downloads ▾ VGNC ▾ Contact us ▾ More ▾

Symbol report for HLA-DQA1 ?

Report **HCOP homology predictions**

HGNC data for HLA-DQA1

Approved symbol ?	HLA-DQA1
Approved name ?	major histocompatibility complex, class II, DQ alpha 1
Locus type ?	gene with protein product
HGNC ID ?	HGNC:4942
Symbol status ?	Approved
Previous symbols ?	HLA-DQA
Alias symbols ?	CELIAC1
Chromosomal location ?	6p21.32
Gene groups ?	C1-set domain containing Histocompatibility complex

Gene resources for HLA-DQA1 ?

Ensembl	ENSG00000196735 Ensembl region in detail , Ensembl gene sequence
UCSC	uc003obr.4 Alliar

CELIAC1

HGNC

CELIAC1



[Home](#) [Gene data](#) [Tools](#) [Downloads](#) [VGNC](#) [Contact us](#) [More](#)

[Request symbol](#)

Search results

Filter by type

Gene 2

Filter by gene entry status

Approved 2

Filter by gene locus type

Protein-coding gene 2

Download all results



20 items per page

Items: 1 to 2 of 2

HLA-DQA1: major histocompatibility complex, class II, DQ alpha 1

Gene HGNC ID HGNC:4942 **Locus type** Gene with protein product **Status** Approved
Matches Alias gene symbol: **CELIAC1**

HLA-DQB1: major histocompatibility complex, class II, DQ beta 1

Gene HGNC ID HGNC:4944 **Locus type** Gene with protein product **Status** Approved
Matches Alias gene symbol: **CELIAC1**

Items: 1 to 2 of 2

Псевдоген

- Псевдоген определяется как последовательность, которая не способна продуцировать функциональный белковый продукт, но имеет высокий уровень гомологии с функциональным геном. В общем, мы называем только псевдогены, которые сохраняют гомологию со значительной частью функционального предкового гена.
- Обработанным псевдогенам присваиваются имена на основе конкретного родительского гена, с добавлением буквы P и номера к символу родительского гена (например, NASAP10, «псевдоген НАСА 10»). Нумерация обычно зависит от вида.

CYP21A1 и CYP21A1P

- Особенности строения гена CYP21A2, а именно наличие в непосредственной близости псевдогена CYP21A1, приводят к тому, что около 90% мутантных аллелей являются следствием рекомбинации между этими генами и несут в себе одну или несколько частых мутаций.

Гены полиморфны

- Полиморфизмы генов могут возникать в любой области генома.
- Большинство полиморфизмов молчат, то есть не изменяют функцию или экспрессию гена
- Полиморфизмы — варианты аллелей с достаточно большой частотой
- Однонуклеотидные варианты (SNV) - это однонуклеотидные изменения, которые происходят в геноме в определенном месте, когда один нуклеотид заменяется на другой.

SNP

- Однонуклеотидный полиморфизм (SNP /snɪp/; множественное число SNP /snɪps/) представляет собой замену одного нуклеотида в определенном положении генома, которая присутствует в достаточно большой части рассматриваемой популяции.
- Например, нуклеотид G, присутствующий в определенном месте эталонного генома, может быть заменен нуклеотидом A у меньшинства людей. Две возможные вариации нуклеотидов этого SNP – G или A – называются аллелями.

SNP

- Однонуклеотидные полиморфизмы могут находиться в пределах кодирующих последовательностей генов, некодирующих участков генов или в межгенных участках (участках между генами).
- SNP в кодирующей последовательности не обязательно изменяют аминокислотную последовательность вырабатываемого белка из-за вырожденности генетического кода.
- SNP в кодирующей области бывают двух типов: синонимичные SNP и несинонимичные SNP. Синонимичные SNP не влияют на последовательность белка, тогда как несинонимичные SNP изменяют аминокислотную последовательность белка.
- В геноме человека обнаружено более 600 миллионов SNP.

Примеры

- SNP в гене F5 - «лейденская мутация», тромбофилия фактора V
- Интронный SNP в гене репарации несоответствия ДНК PMS2 (rs1059060, Ser775Asn) предполагалось, что связан с повышенным повреждением ДНК сперматозоидов и риском мужского бесплодия.

Биоинформационные базы данных

- dbSNP — это база данных SNP Национального центра биотехнологической информации (NCBI). По состоянию на 8 июня 2015 г. dbSNP перечислил 149 735 377 SNP у людей.[
- SNPeD — это база данных в стиле вики, поддерживающая аннотацию, интерпретацию и анализ личного генома.
- Kaviar представляет собой сборник SNP из нескольких источников данных, включая dbSNP.
- База данных OMIM описывает связь между полиморфизмами и заболеваниями (например, предоставляет заболевания в текстовой форме).
- dbSAP - база данных полиморфизма отдельных аминокислот для обнаружения вариаций белков
- База данных мутаций генов человека содержит мутации генов, вызывающие или связанные с наследственными заболеваниями человека, а также функциональные SNP.
- Международный проект ХарМар, в рамках которого исследователи идентифицируют SNP тегов, чтобы иметь возможность определить коллекцию гаплотипов, присутствующих у каждого субъекта.
- GWAS Central позволяет пользователям визуально опрашивать фактические данные об ассоциациях на сводном уровне в одном или нескольких исследованиях ассоциаций на уровне всего генома.

Номенклатура dbSNP

- Стандарт rs#### принят dbSNP и использует префикс «rs» для «эталонного SNP», за которым следует уникальный и произвольный номер

Общество по изменению генома человека (HGVS)

- использует стандарт, который передает больше информации об SNP:
 - c.76A>T: для кодирующей области, за которым следует номер положения нуклеотида, за которым следует однобуквенное сокращение нуклеотида (A, C, G, T или U), за которым следует знак «больше» («>») для обозначения замены, за которой следует сокращение нуклеотида, который заменяет первый
 - p.Ser123Arg: для белка, за которым следует трехбуквенное сокращение аминокислоты, за которым следует число, обозначающее положение аминокислоты, а затем сокращение аминокислоты, которое заменяет первое.

rs6025

- rs6025 представляет собой SNP в гене фактора V F5, кодирующий изменение белка с аргинина в положении 506 на глутамин.
- Полученный аллель rs6025(A) кодирует мутацию, известную как лейденская мутация, R506Q. Эту мутацию часто называют изменением G1691A или 1691G-A.
- p.Arg506Gln
- HGVS nomenclature NM_000130.4 c.1601G>A p.R534Q

rs6025Curr
Released Septe

Organism	<i>Homo sapiens</i>	Clinical Significance	Reported in ClinVar
Position	chr1:169549811 (GRCh38.p14) ⓘ	Gene : Consequence	F5 : Missense Variant
Alleles	C>A / C>T	Publications	184 citations
Variation Type	SNV Single Nucleotide Variation		LitVar² <small>1096</small>
Frequency	T=0.016744 (4432/264690, TOPMED) T=0.022585 (3828/169494, ALFA) T=0.017619 (2471/140250, GnomAD) (+ 5 more)	Genomic View	See rs on genome

Frequency

Variant
DetailsClinical
Significance

HGVS

Submissions

History

Publications

Fla

ALFA Allele Frequency

The ALFA project provide aggregate allele frequency from dbGaP. More information is available on the project [page](#) including descriptions, data access, use.

Release Version: 20230706150541Search:

Population	Group	Sample Size	Ref Allele	Alt Allele
Total	Global	185916	C=0.977592	T=0.022408
European	Sub	145872	C=0.974471	T=0.025529
African	Sub	15632	C=0.99507	T=0.00493

Frequency **Variant Details** Clinical Significance HGVS Submissions History Publicat

Genomic Placements

Sequence name	Change
F5 RefSeqGene (LRG_553)	NG_011806.1:g.41721G>T
F5 RefSeqGene (LRG_553)	NG_011806.1:g.41721G>A
GRCh37.p13 chr 1	NC_000001.10:g.169519049T>C
GRCh37.p13 chr 1	NC_000001.10:g.169519049T>A
GRCh38.p14 chr 1	NC_000001.11:g.169549811C>A
GRCh38.p14 chr 1	NC_000001.11:g.169549811C>T

Gene: F5, coagulation factor V (minus strand)

Molecule type	Change	Amino acid[Codon]
coagulation factor V preproprotein	NP_000121.2:p.Arg534Leu	R (Arg) > L (Leu)
coagulation factor V preproprotein	NP_000121.2:p.Arg534Gln	R (Arg) > Q (Gln)
F5 transcript	NM_000130.5:c.1601G>T	R [CGA] > L [CTA]
F5 transcript	NM_000130.5:c.1601G>A	R [CGA] > Q [CAA]

rs6025

Organism	<i>Homo sapiens</i>	Clinical Significance	Reported in ClinVar
Position	chr1:169549811 (GRCh38.p14) ?	Gene : Consequence	F5 : Missense Variant
Alleles	C>A / C>T	Publications	184 citations LitVar² <small>1096</small>
Variation Type	SNV Single Nucleotide Variation	Genomic View	See rs on genome
Frequency	T=0.016744 (4432/264690, TOPMED) T=0.022585 (3828/169494, ALFA) T=0.017619 (2471/140250, GnomAD) (+ 5 more)		

Curr
Released Septe

Frequency

Variant
DetailsClinical
Significance

HGVS

Submissions

History

Publications

Fla

ALFA Allele Frequency

The ALFA project provide aggregate allele frequency from dbGaP. More information is available on the project [page](#) including descriptions, data access, use.

Release Version: 20230706150541

Search:

Population	Group	Sample Size	Ref Allele	Alt Allele
Total	Global	185916	C=0.977592	T=0.022408
European	Sub	145872	C=0.974471	T=0.025529
African	Sub	15632	C=0.99507	T=0.00493

rs6025

SNPedia

Orientation minus

Stabilized minus

Geno [↕](#) Mag [↕](#) Summary [↕](#)

(A;A)	7	11.4x higher risk of thrombosis
(A;G)	4.1	3.5-4.4x risk of thrombosis
(G;G)	0	normal/common risk of thrombosis

Reference GRCh38 38.1/141

Chromosome1

Position 169549811

Gene F5

is a snp

is mentioned by

dbSNP rs6025

dbSNP rs6025

(classic) rs6025

ClinGen rs6025

ebi rs6025

HLI rs6025

Exac rs6025

F5

- В нормальном пути фактор V действует как кофактор, позволяя фактору Ха активировать протромбин, в результате чего образуется фермент тромбин. Тромбин, в свою очередь, расщепляет фибриноген с образованием фибрина, который полимеризуется, образуя плотную сетку, составляющую большую часть сгустка. Активированный протеин С является природным антикоагулянтом, который ограничивает степень свертывания крови путем расщепления и разрушения фактора V.

	Max Magnitude	Chromosome position
rs1018827	0	169,544,768
rs10919186	0	169,547,462
rs118203905	5.5	169,555,300
rs118203906	5.5	169,555,299
rs118203907	5.5	169,530,805
rs118203908	5.5	169,542,689
rs118203909	5.5	169,541,609
rs118203910	5.5	169,518,453
rs118203911	5.5	169,552,693
rs118203912	5.5	169,560,701
rs12040141	0	169,532,462
rs1800595	0	169,541,110
rs2269648	0	169,586,812
rs2420371	0	169,522,317
rs3766110	0	169,545,945
rs386834226	0	169,541,903
rs386834227	0	169,529,635
rs386834228	0	169,525,949
rs4524	0	169,542,517
rs4525	0	169,542,496
rs6019	0	169,572,275
rs6020	0	169,549,874
rs6022	0	169,560,588
rs6025	7	169,549,811
rs6028	0	169,582,444
rs6030	0	169,529,737
rs6427196	0	169,511,985
rs6703865	0	169,581,725
rs7542281	0	169,567,201
rs757953549	5.5	169,542,872
rs9332595	0	169,545,117
rs9332647	0	169,522,862

Анализ SNP

- SNP можно легко проанализировать, поскольку они содержат только два возможных аллеля и три возможных генотипа, включающих эти два аллеля: гомозиготный А, гомозиготный В и гетерозиготный АВ, что приводит к множеству возможных методов анализа. Некоторые из них включают: секвенирование ДНК; капиллярный электрофорез; масс-спектрометрии; одноцепочечный конформационный полиморфизм (SSCP); одиночное базовое расширение; электрохимический анализ; денатурирующая ВЭЖХ и гель-электрофорез; полиморфизма длин рестрикционных фрагментов; и гибридационный анализ.

Лаборатории ЦИР
Лаборатории репродукции

Анализы и цены | Адреса лабораторий | Библиотека | Правила сд

Пациентам | Врачам

Поиск по сайту | Например: Общий анализ крови

Главная > Анализы и цены > Молекулярно-генетические исследования > Риск развития тром

Полиморфизмы (два фактора):
Лейденская мутация (FV): 1691 G>A,
Ген протромбина (FII): 20210 G>A
Артикул: 00849

В этот блок анализов входят:

- [Полиморфизм FV G1691A \(rs6025\)](#)
- [Полиморфизм FII G20210A \(rs1799963\)](#)



Результаты исследований

Пациент

Пол

Возраст

Заказ

Д.п.м.

Показатель	Результат		Референсный интервал	
Молекулярно-генетические исследования				
[00849] Полиморфизмы (два фактора): Лейденская мутация (FV): 1691 G>A, Ген протромбина (FII): 20210 G>A - буккальный эпителий				
Гены тромбофилии				
Ген	Кодируемый белок	Полиморфизм	Результат	Шифр
F2	II коагуляционный фактор (протромбин)	G20210A (rs1799963)	G/G	1
Ген F2 кодирует фактор II свертывания крови (протромбин). Наличие генотипа AA повышает риск венозного тромбоза в 6.7 раз, AG - в 2,8 раза. Частота аллелей: G=0.9964 Ориентация цепи: плюс Синонимы: c.20210, i3002432 Заключение: По гену фактора II свертывания крови (FII, протромбин) выявлен генотип G/G, встречающийся до 99% пациентов. Данный вариант не увеличивает риск тромботических осложнений, патологии сердечно-сосудистой системы и гестации.				
F5	V коагуляционный фактор свертываемости крови (фактор Лейдена)	G1691A (rs6025)	G/G	1
Наличие лейденской мутации в гетерозиготном варианте (AG) повышает риск тромбоза в 3,5-4,4 раза. В гомозиготном варианте (AA) - в 11,4 раза. Частота аллелей: G=0.98052 Синонимы: лейденская мутация, R506Q, c.1691 Заключение: По гену фактора V свертывания крови (FV) выявлен генотип G/G, встречающийся в популяции до 98% пациентов. Данный вариант не увеличивает риск тромботических осложнений, патологии сердечно-сосудистой системы и осложнений беременности. Шифр: 1 - гомозигота частый аллель, 2 - гетерозигота, 3 - гомозигота редкий аллель				

Результаты исследований не являются диагнозом и интерпретируются лечащим врачом с учетом всех данных о пациенте (лабораторных, инструментальных и



Handwritten signature or initials.

Легко сдать удалённо в ЦИР



Буккальный эпителий



Буккальный эпителий



«Капля крови» на промокашке

Спасибо за внимание!