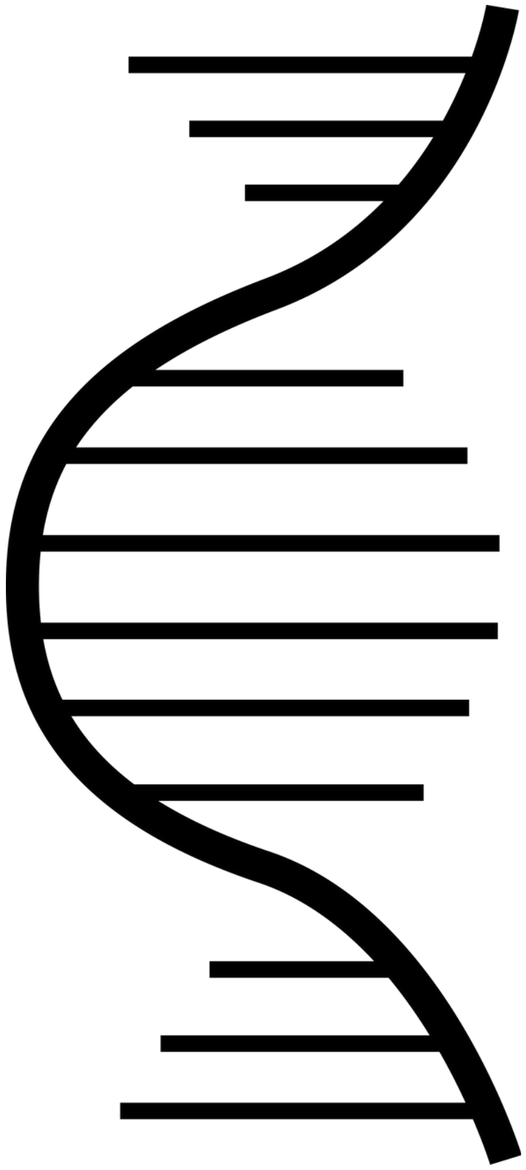
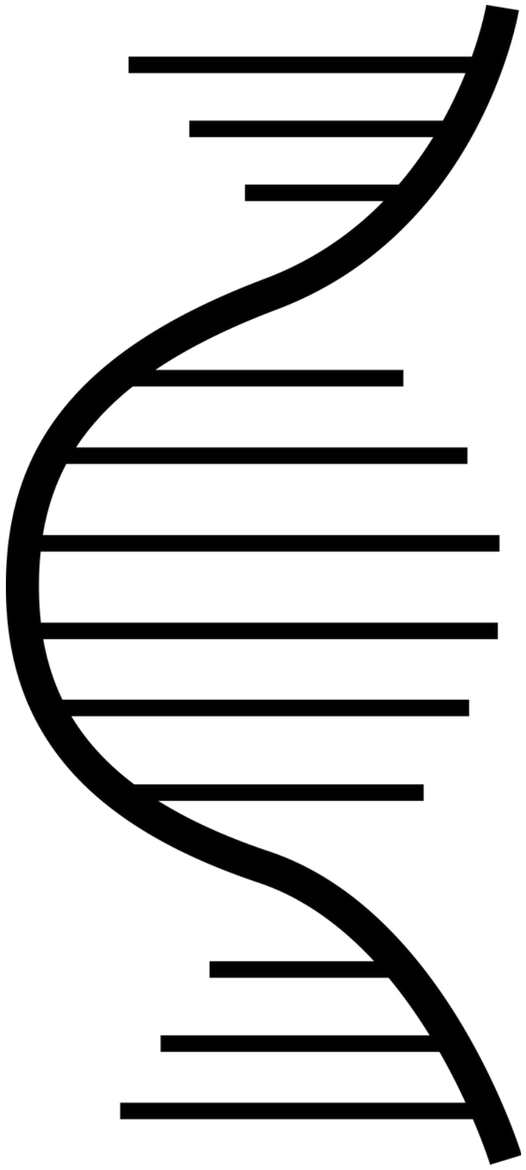
VARIANT
CURATION

LIST OF CONTENTS

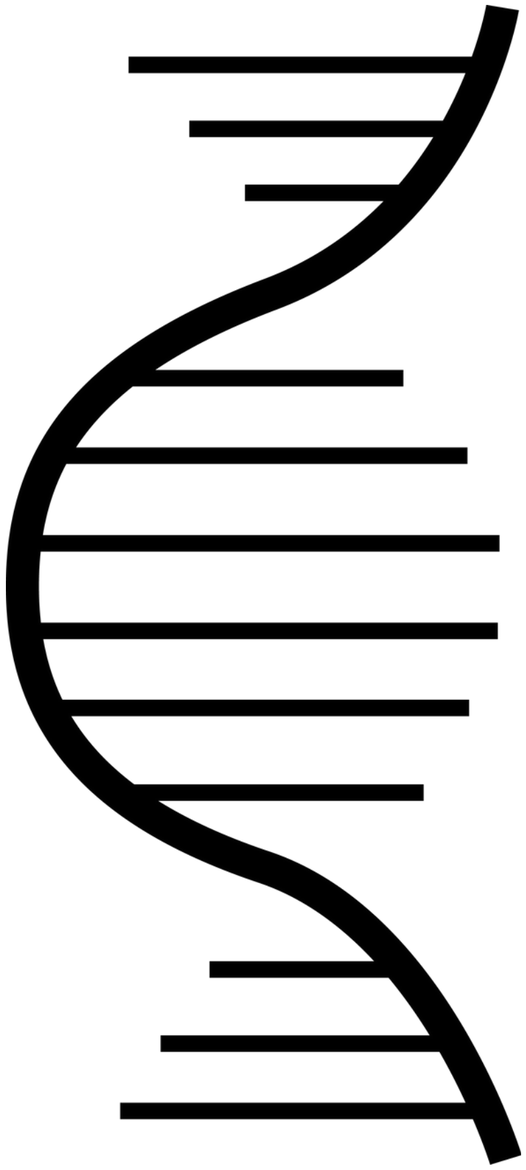


HOW MANY NUCLEOTIDES DOES A HUMAN GENOME HAVE?

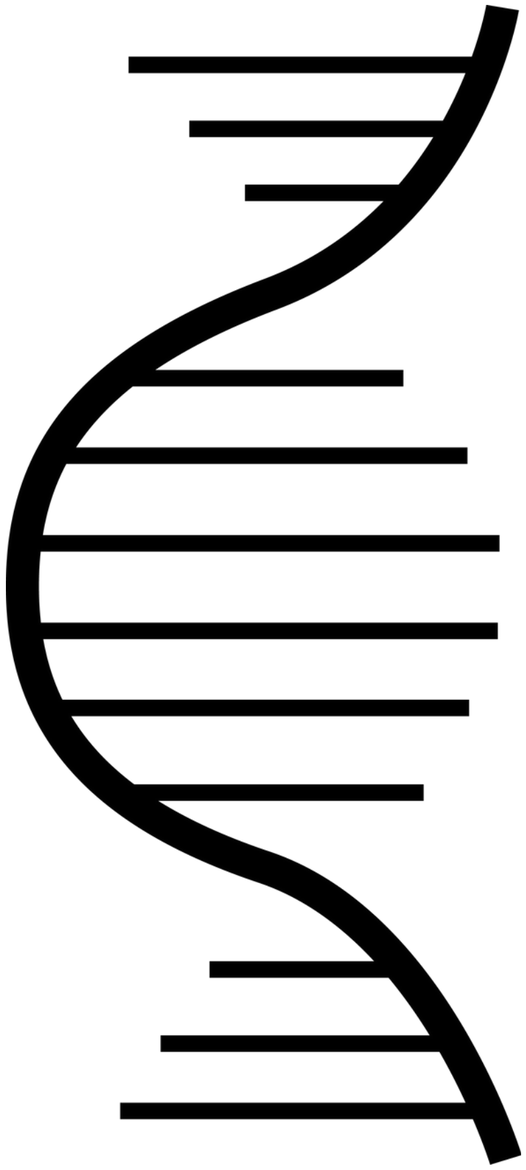


HOW MANY NUCLEOTIDES DOES A HUMAN GENOME HAVE?

3.2 billion



HOW MANY GENES DOES A HUMAN GENOME HAVE?



HOW MANY GENES DOES A HUMAN GENOME HAVE?

20,000

~1% of the genome

99.6 % identical genome

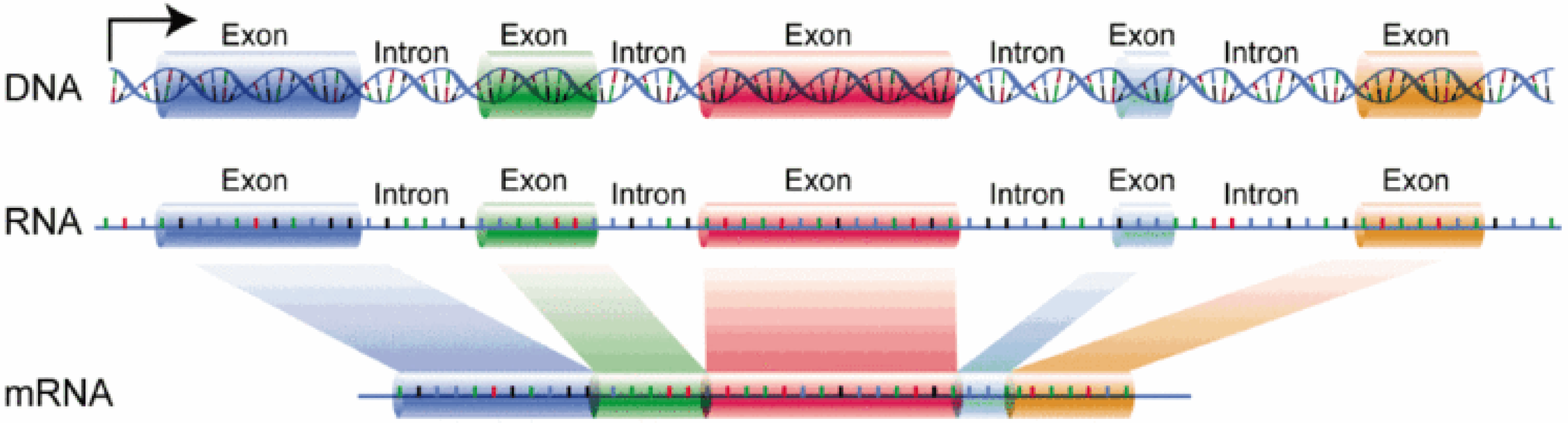
0.4% variation

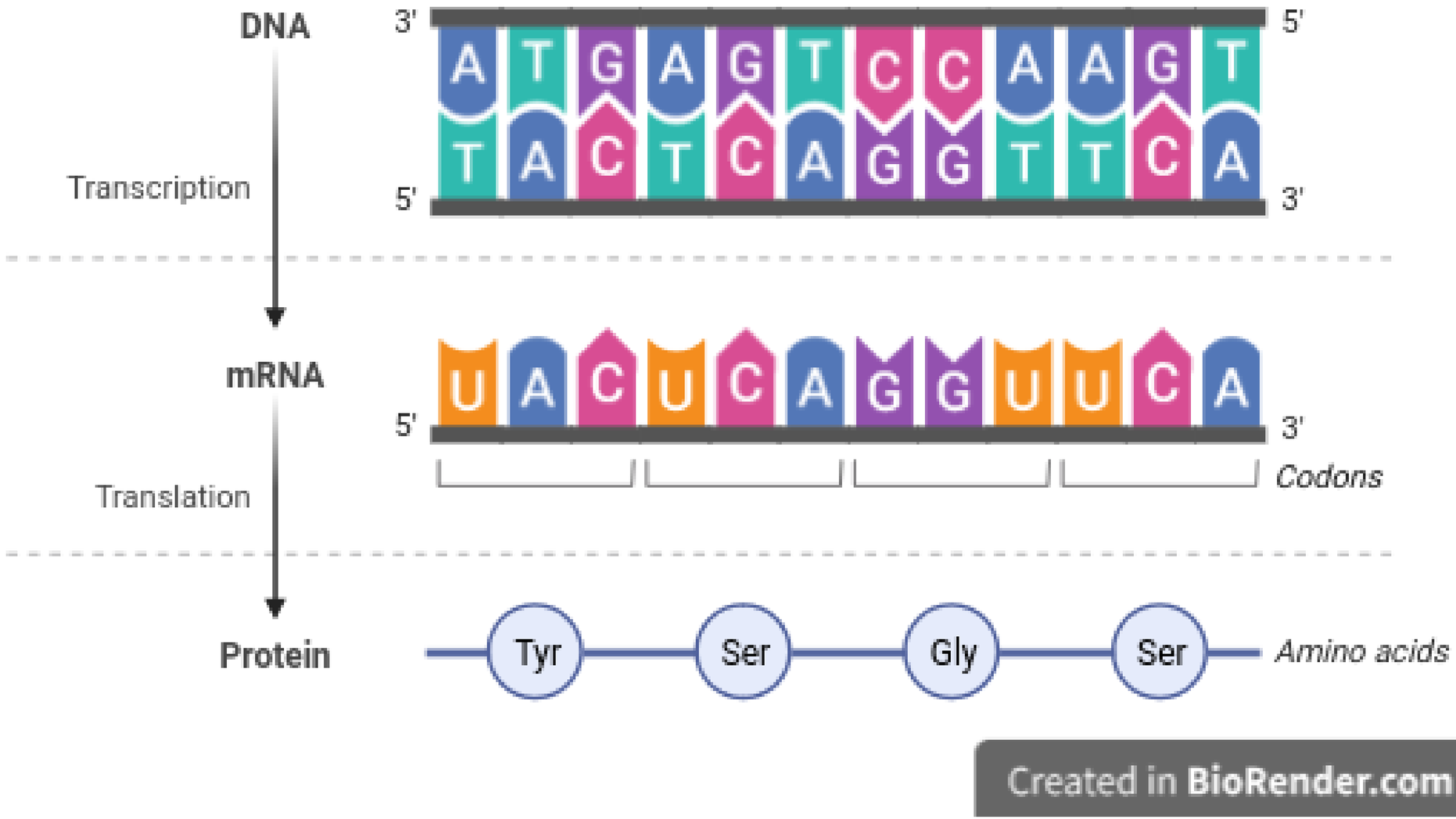
12 million base pairs

~4 to 5 million SNPs in a person's genome

01

Central Dogma: From DNA to Proteins




DNA
3'
5'
A T G A G T C C A A G T
T A C T C A G G T T C A
5'
3'
Transcription
mRNA
5'
U A C U C A G G U U C A
3'
Codons
Translation
Protein
Tyr
Ser
Gly
Ser
Amino acids
Created in BioRender.com
bio

Codon Chart

First Base	Second Base: U	Second Base: C	Second Base: A	Second Base: G	Third Base
U	UUU, UUC – Phenylalanine (Phe/F) UUA, UUG – Leucine (Leu/L)	CUU, CCU, CAU, CGU – Serine (Ser/S)	AUU, ACU – Tyrosine (Tyr/Y) **AAU** – STOP **AGU** – STOP	GUU, GCU – Cysteine (Cys/C) **GAU** – STOP GGU – Tryptophan (Trp/W)	U C A G
C	CUU, CUC, CUA, CUG – Leucine (Leu/L)	CUC, CCC, CAC, CGC – Proline (Pro/P)	AUC, ACC – Histidine (His/H) AAC, AGC – Glutamine (Gln/Q)	GUC, GCC, GAC, GGC – Arginine (Arg/R)	U C A G
A	AUU, AUC, AUA – Isoleucine (Ile/I) **AUG** – Methionine (Met/M)	CUA, CCA, CAA, CGA – Threonine (Thr/T)	AUA, ACA – Asparagine (Asn/N) AAA, AGA – Lysine (Lys/K)	GUA, GCA – Serine (Ser/S) GAA, GGA – Arginine (Arg/R)	U C A G
G	GUU, GUC, GUA, GUG – Valine (Val/V)	CUG, CCG, CAG, CGG – Alanine (Ala/A)	AUG, ACG – Aspartic acid (Asp/D) AAG, AGG – Glutamic acid (Glu/E)	GUG, GCG, GAG, GGG – Glycine (Gly/G)	U C A G

02

Main Assumptions

- Seen as rare or not seen in the healthy population
- At the exonic or splicing region
- Affect the function of the protein
  - Found at;
    - a mutational hotspot
    - an important functional domain
  - The affected gene to be intolerant to the type of mutation of interest
  - Not shown as benign previously

MUTATION TYPES

Synonymous

Missense

In-frame indel

Loss of Function

- Nonsense
- Frameshift indels
- Splicing

03

ACMG Guidelines

- The American College of Medical Genetics and Genomics
- Internationally accepted guidelines for the interpretation of variants

BENIGN CRITERIA

Stand-alone

- Frequency of 0.005 for dominant genes
- Frequency of 0.01 for recessive genes

Strong

- Healthy example for fully penetrant diseases
- Experimental evidences
- Lack of segregation

Supporting

- Missense when LoF is known mechanism
- Inframe indels in repetitive regions
- In silico tool predictions
- Synonym mutations
- Mutations in; UTR, intron, intergenic

PATHOGENICITY CRITERIA

Very Strong

- Loss of function
- If LoF is known mechanism!

Strong

- Known pathogenic missense
- De novo
- Experimentally proven damage at variant level

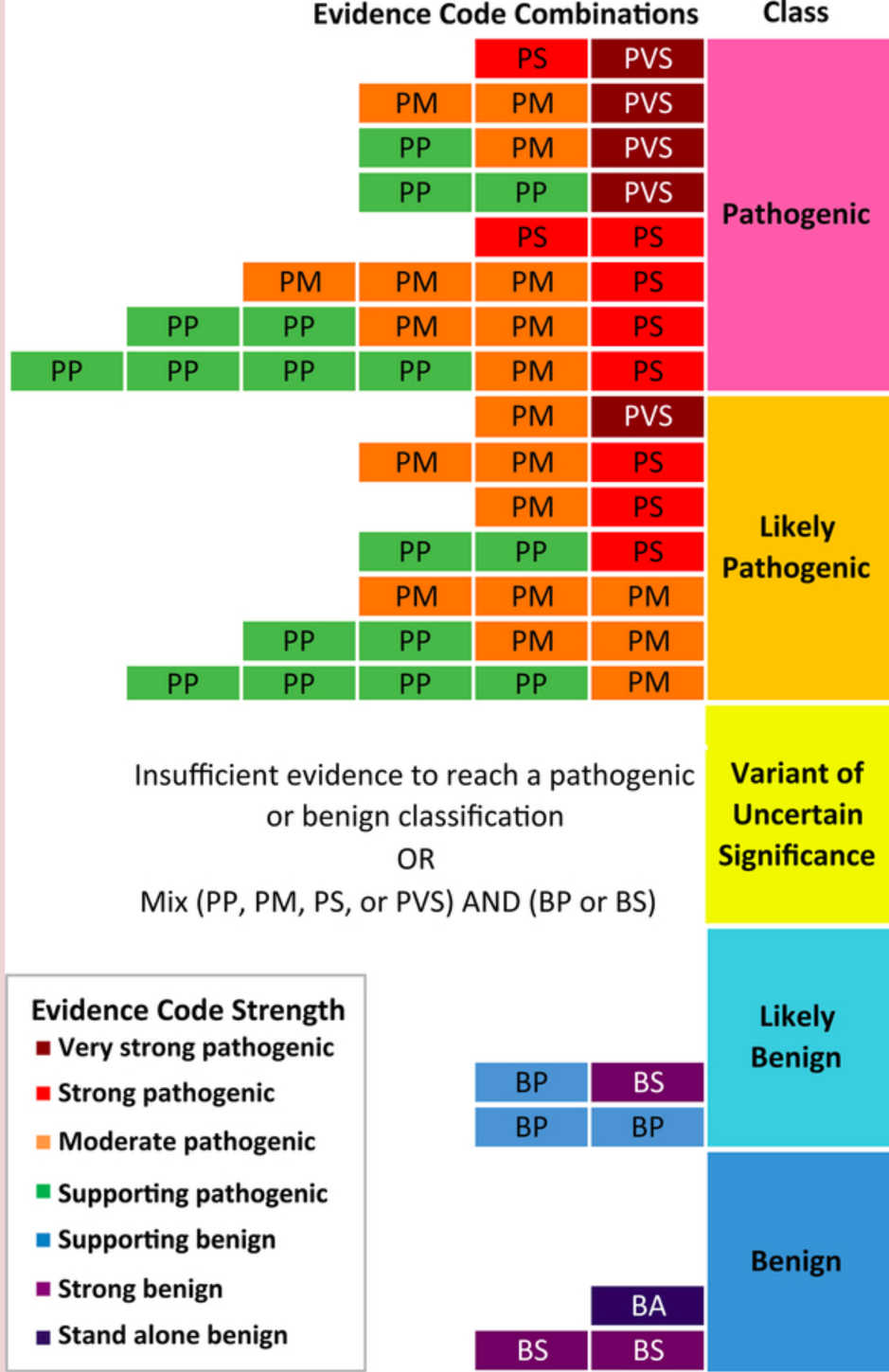
Moderate

- Mutational hotspots or well-known functional domain
- Inframe indels

Supporting

- Missense in a gene that has a low missense rate
- In silico tool predictions
- Experimentally proven damage at gene level

PATHOGENICITY INTERPRETATION



We want to hear about how you use gnomAD and your wish list! Please take 5 minutes to fill out our user survey.

genome aggregation database

gnomAD v2.1.1 Search by gene, region, or variant

Please note that gnomAD v2.1.1 and v3.1.1 have substantially different but overlapping sample compositions and are on different genome builds. For more information, see "Should I switch to the latest version of gnomAD?"

Examples

- Gene: PCSK9
- Transcript: ENST00000302118
- gnomAD v2.1.1 variant: 1-55516888-G-GA
- gnomAD v3.1.1 variant: 1-55051215-G-GA

DLG2 discs large MAGUK scaffold protein 2

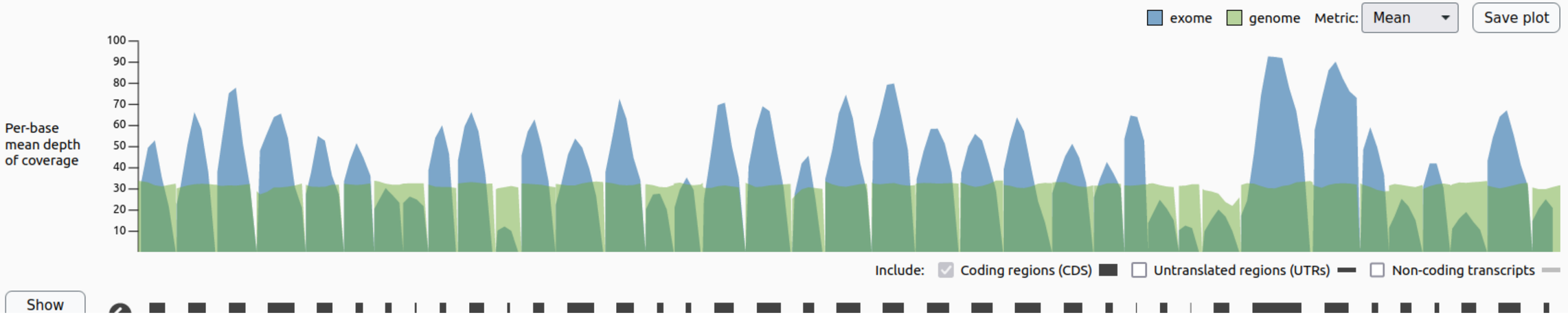
Dataset: gnomAD v2.1.1 | gnomAD SVs v2.1

Genome build GRCh37 / hg19
Ensembl gene ID ENSG00000150672.12
Ensembl canonical transcript ENST00000376104.2
Other transcripts ENST00000376106.3, ENST00000532653.1, and 30 more
Region 11:83166055-85338966
External resources Ensembl, UCSC Browser, and more

Constraint

Category	Expected SNVs	Observed SNVs	Constraint metrics
Synonymous	195.5	182	Z = 0.76 o/e = 0.93 (0.82 - 1.05)
Missense	535.2	400	Z = 2.08 o/e = 0.75 (0.69 - 0.81)
pLoF	61.2	13	pLI = 0.71 o/e = 0.21 (0.14 - 0.34)

Constraint metrics based on Ensembl canonical transcript (ENST00000376104.2).



DECIPHER GRCh38

About | Browse | DDD (UK)

Search DECIPHER

Help | Join | Log in

Mapping the clinical genome

Explore DECIPHER

It's free and you don't need to log in

DECIPHER is used by the clinical community to share and compare phenotypic and genotypic data. The DECIPHER database contains data from 46,207 patients who have given consent for broad data-sharing; DECIPHER also supports more limited sharing via consortia. Have a look at the numbers.

Anyone can browse publicly-available patient data on DECIPHER and request to be put in contact with the responsible clinician. Why? Because sharing benefits everyone.

Explore DECIPHER's genome browser

Delve into the Human Phenotype Ontology

Search all open-access DECIPHER data

Join DECIPHER

Be part of the sharing community

Projects affiliated to DECIPHER can deposit and share patients, variants, and phenotypes to invite collaboration and facilitate diagnosis. Once deposited, you can use DECIPHER to identify and prioritise potential matches, and you can request notifications as soon as new matches arrive.

As well as influencing individual patient outcomes, use of DECIPHER has contributed to over 2600 published articles since 2004. It's still free, and you are in control of what data to make public.

Join now

Find out more

Already a member?

Log in to access your patient data

Email address

Email

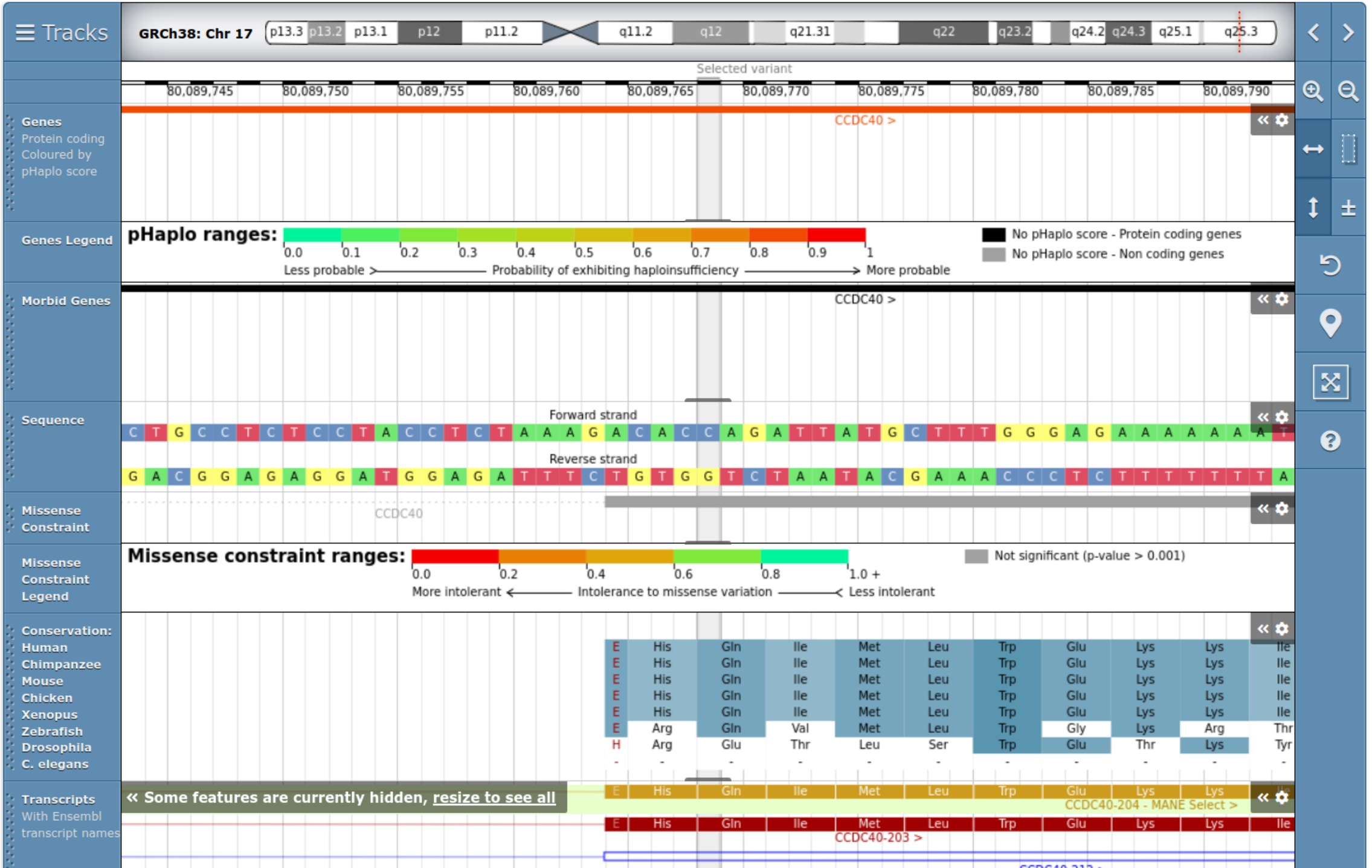
Password

Password

Log in

Reset your password

Feedback


Tracks
GRCh38: Chr 17
p13.3 p13.2 p13.1 p12 p11.2 q11.2 q12 q21.31 q22 q23.2 q24.2 q24.3 q25.1 q25.3
Selected variant
80,089,745 80,089,750 80,089,755 80,089,760 80,089,765 80,089,770 80,089,775 80,089,780 80,089,785 80,089,790
Genes
Protein coding
Coloured by
pHaplo score
CCDC40 >
Genes Legend
pHaplo ranges:
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Less probable
Probability of exhibiting haploinsufficiency
More probable
No pHaplo score - Protein coding genes
No pHaplo score - Non coding genes
Morbid Genes
CCDC40 >
Sequence
Forward strand
C T G C C T C T C C T A C C T C T A A A G A C A C C A G A T T A T G C T T T G G G A G A A A A A A A A T
Reverse strand
G A C G G A G A G G A T G G A G A T T T C T G T G G T C T A A T A C G A A A C C C T C T T T T T T T A
CCDC40
Missense Constraint
Missense constraint ranges:
0.0 0.2 0.4 0.6 0.8 1.0 +
More intolerant
Intolerance to missense variation
Less intolerant
Not significant (p-value > 0.001)
Missense Constraint Legend
Conservation:
Human
Chimpanzee
Mouse
Chicken
Xenopus
Zebrafish
Drosophila
C. elegans
E His Gln Ile Met Leu Trp Glu Lys Lys Ile
E His Gln Ile Met Leu Trp Glu Lys Lys Ile
E His Gln Ile Met Leu Trp Glu Lys Lys Ile
E His Gln Ile Met Leu Trp Glu Lys Lys Ile
E His Gln Ile Met Leu Trp Glu Lys Lys Ile
E Arg Gln Val Met Leu Trp Gly Lys Arg Thr
H Arg Glu Thr Leu Ser Trp Glu Thr Lys Tyr
Transcripts
With Ensembl
transcript names
« Some features are currently hidden, resize to see all
CCDC40-204 - MANE Select >
CCDC40-203 >

50 YEARS OMIM
Human Genetics Knowledge for the World

OMIM®

An Online Catalog of Human Genes and Genetic Disorders

Updated May 8, 2023

Search OMIM for clinical features, phenotypes, genes, and more...

Advanced Search : OMIM, Clinical Synopses, Gene Map

Need help? : Example Searches, OMIM Search Help, OMIM Video Tutorials

Mirror site : https://mirror.omim.org

OMIM is supported by a grant from NHGRI, licensing fees, and generous contributions from people like you.

Make a donation!

McKUSICK-NATHANS Department of Genetic Medicine

JOHNS HOPKINS MEDICINE

04

Case study

- 6-year-old boy
- Global developmental delay
- Generalized hypotonia
- Speech delay
- Strabismus: bilateral esotropia
- Decreased pain response
- Hyperactive DTR
- Mild facial dysmorphisms: frontal bossing, low-set ears

FILTERING CRITERIA

- Max allele frequency **< 0.001**
- Exclude variants with ClinVar clinical significance of **benign** or **likely benign**
- Primary candidates -> knock-down (homozygous LoF)
- Check ClinVar Disease Name