

Tutorial 1. Introduction

The Chinese University of Hong Kong

BMEG3102 Bioinformatics

TA: Yizhen Chen



TA Info

- Yizhen CHEN (Channie)
- Office: SHB 1023
- Email: yzchen@cse.cuhk.edu.hk
- Consultation hour: Tue 14:00-16:00

Agenda

- Course Assessment
- Introduction to Bioinformatics
- Demonstration on how to do exercise related to Web tool
- Case Study: Mutation and disease (Sickle Cell Anemia)

Course assessment

- Assignments: 40%
- Class participation: 5%
- Reading presentation: 5%
- Mid-term examination: 20%
- Final examination: 30%

What is “Bioinformatics”?

- Biology + Informatics
- For Biology, you will learn / see:
 - Some biology terms
 - Basic but important biology concepts
 - Some medical problems (e.g. find genetic variants that are related to a disease)
- For Informatics, you will learn / see:
 - Some engineering terms
 - Large-scale data sets
 - Some systematic ways to solve problems and analyze results

Basic terms in Biology

Term	Description
Bioinformatics	Application of computer science and information technology to the field of biology and medicine
Cell	Basic functional unit of life
Chromosome	Organized structure of DNA and protein found in cells
DNA	Carrying genetic information / double helix / A, T, C, G
RNA	Single strand / A, U, C, G
Protein	A strand of amino acids, folded into a particular structure
Amino acid	20 common types
DNA Replication	DNA → DNA (mitosis and meiosis)
Transcription	DNA → RNA
Translation	RNA → Protein
Codon	Specify which amino acid will be translated for each DNA triplet

Basic terms in Biology

Term	Description
Coding Strand	Appears the same (except $T \rightarrow U$) as the RNA transcribed
Template Strand	Used as the template for RNA synthesis
Mutation	Change of DNA sequence, e.g. substitution, <u>insertion</u> or <u>deletion</u> (indel) in some genomic regions
Transition	Nucleotide changes within purine (nucleotides with two carbon rings, i.e., $A \leftrightarrow G$) or within pyrimidine (nucleotides with one carbon ring, i.e., $C \leftrightarrow T$)
Transversion	Nucleotide changes other than $A \leftrightarrow G$ and $C \leftrightarrow T$ (more change in size, leading to more drastic consequences)
Synonymous Mutation	Different DNA sequences produce the same protein
Non-synonymous Mutation	Different DNA sequences produce different proteins
Missense Mutation	Change one amino acid to another (Non-synonymous)
Nonsense Mutation	Change one amino acid to stop codon (Non-synonymous)
Single Nucleotide Polymorphism (SNP)	a substitution of a single nucleotide at a specific position in the genome in a sufficiently large fraction of the population

Central Dogma

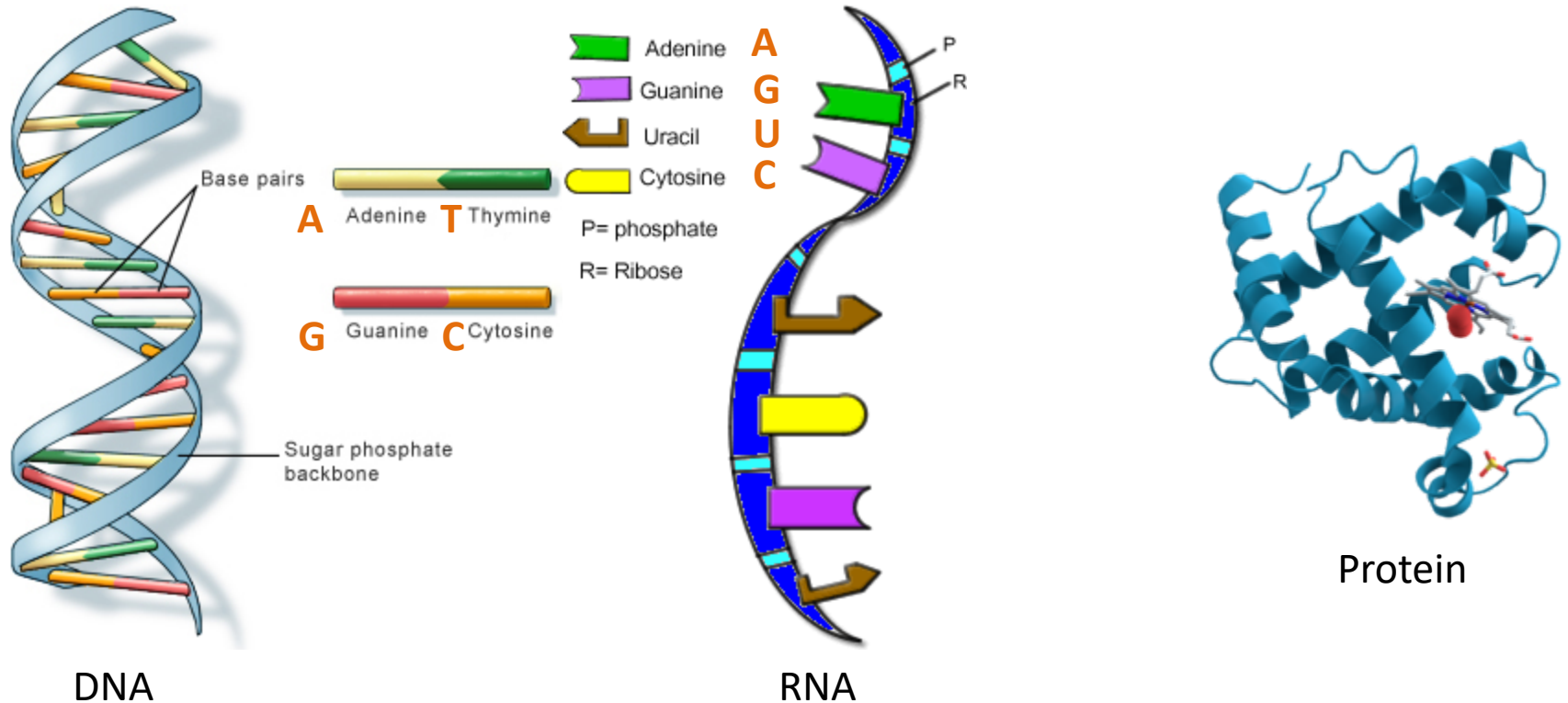


Image credit: <http://ghr.nlm.nih.gov/handbook/basics/dna>, <http://www.biologycorner.com/bio1/DNA.html>, wiki image

Central Dogma

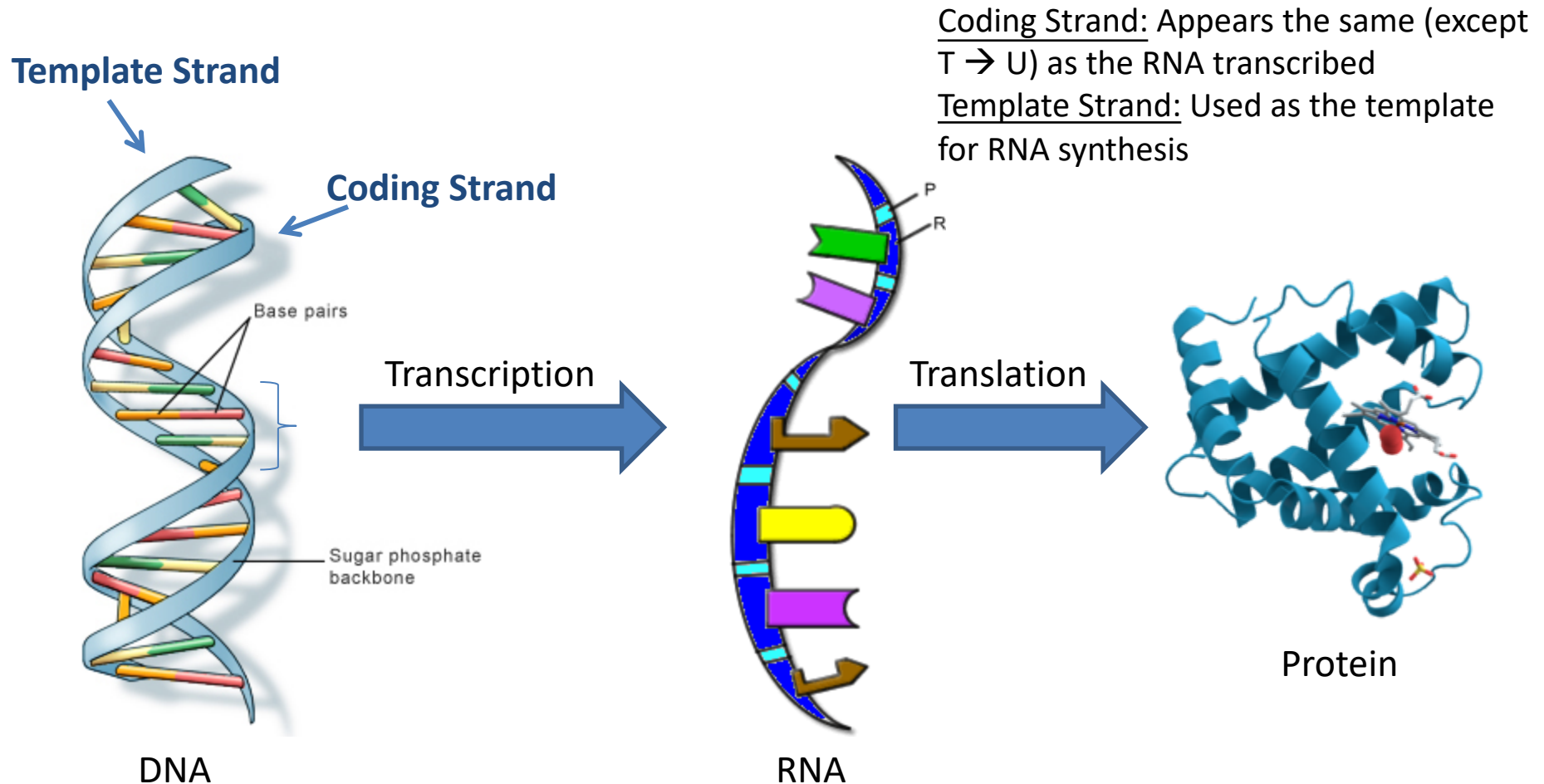


Image credit: <http://ghr.nlm.nih.gov/handbook/basics/dna>, <http://www.biologycorner.com/bio1/DNA.html>, wiki image

Some useful videos

DNA Replication: <https://www.youtube.com/watch?v=TNKWgcFPHqw>

Transcription: <http://www.youtube.com/watch?v=WsofH466lqk>

Translation: <http://www.youtube.com/watch?v=5bLEDd-PSTQ>

Mitosis: <https://www.youtube.com/watch?v=f-lDPgEfAHI&t=32s>

Meiosis: <https://www.youtube.com/watch?v=VzDMG7ke69g>

Youtube channel

Amoeba Sisters

<https://www.youtube.com/user/AmoebaSisters>

Codon table

		Second base				
		U	C	A	G	
First base	U	UUU } Phenyl- alanine F UUC UUA } Leucine L UUG	UCU } UCC } Serine S UCA UCG	UAU } Tyrosine Y UAC UAA Stop codon UAG Stop codon	UGU } Cysteine C UGC UGA Stop codon UGG Tryptophan W	U C A G
	C	CUU } CUC } Leucine L CUA CUG	CCU } CCC } Proline P CCA CCG	CAU } Histidine H CAC CAA } Glutamine Q CAG	CGU } CGC } Arginine R CGA CGG	U C A G
	A	AUU } Isoleucine I AUC AUA AUG Methionine start codon M	ACU } ACC } Threonine T ACA ACG	AAU } Asparagine N AAC AAA } Lysine K AAG	AGU } Serine S AGC AGA } Arginine R AGG	U C A G
	G	GUU } GUC } Valine V GUA GUG	GCU } GCC } Alanine A GCA GCG	GAU } Aspartic acid D GAC GAA } Glutamic acid E GAG	GGU } GGC } Glycine G GGA GGG	U C A G

Exercise: Transcription and translation

- Given the following **coding strand** of a DNA:
123456789012345678901234567890
5' - TGTTGCATGAGATATGTACAGAGGTAGTGA - 3'
- Work out the **pre-mRNA** and **mature mRNA**, if position from 16 to 21 are **intron**, and the remaining are **exons**
- Deduce a **protein sequence** from the **mature mRNA**, and explain briefly how the sequence is deduced

Answer: Transcription and translation

- Pre-mRNA:
 - 5' - UGUUGCAUGAGAUUAUGUACAGAGGUAGUGA - 3'
- Mature mRNA:
 - 5' - UGUUGC **AUG** **AGA** **UAU** **AGG** **UAG** UGA - 3'
- Protein:
 - NH₃ - **M** **R** **Y** **R** * - COOH
- A coding sequence will always start with **start codon** (AUG) and end with **stop codon** (UAA, UAG or UGA).

Exercise: Six reading frames

- You are provided with a DNA sequence **within a coding region** below:
 - Write down the sequence of reverse strand, and
 - Deduce **ALL** possible protein sequences (you may **NOT** worry about the start codon and stop codon)
- DNA sequence:
5'-AACTTGTTTCGTACA-3'

Six reading frames

- Given a sequence of DNA in coding region, we can have **6** ways to translate the DNA to protein.

- Example:

- +3
L V R T
- +2
T C S Y
- +1
N L F V

5' -AACTTGTTCGTACA-3'

3' -TTGAACAAGCATGT-5'

- 1
K N T C
- 2
S T R V
- 3
V Q E Y

		Second base				
		U	C	A	G	
First base	U	UUU } Phenylalanine F UUC } UUA } Leucine L UUG }	UCU } Serine S UCC } UCA } UCG }	UAU } Tyrosine Y UAC } UAA } Stop codon UAG } Stop codon	UGU } Cysteine C UGC } UGA } Stop codon UGG } Tryptophan W	U C A G
	C	CUU } Leucine L CUC } CUA } CUG }	CCU } Proline P CCC } CCA } CCG }	CAU } Histidine H CAC } CAA } Glutamine Q CAG }	CGU } Arginine R CGC } CGA } CGG }	U C A G
	A	AUU } Isoleucine I AUC } AUA } AUG } Methionine start codon M	ACU } Threonine T ACC } ACA } ACG }	AAU } Asparagine N AAC } AAA } Lysine K AAG }	AGU } Serine S AGC } AGA } Arginine R AGG }	U C A G
	G	GUU } Valine V GUC } GUA } GUG }	GCU } Alanine A GCC } GCA } GCG }	GAU } Aspartic acid D GAC } GAA } Glutamic acid E GAG }	GGU } Glycine G GGC } GGA } GGG }	U C A G

Six reading frames

- Given a sequence of DNA in coding region, we can have 6 ways to translate the DNA to protein.

- Example:

- +3

L V R T

- +2

T C S Y

- +1

N L F V

5' - AAC FTG TTC GTACA - 3'

3' - TTGAACAAGCATGT - 5'

- 1

K N T C

- 2

S T R V

- 3

V Q E Y

		Second base				
		U	C	A	G	
First base	U	UUU Phenylalanine F UUC UUA Leucine L UUG	UCU Serine S UCC UCA UCG	UAU Tyrosine Y UAC UAA Stop codon UAG Stop codon	UGU Cysteine C UGC UGA Stop codon UGG Tryptophan W	Third base U C A G U C A G U C A G U C A G
	C	CUU Leucine L CUC CUA CUG	CCU Proline P CCC CCA CCG	CAU Histidine H CAC CAA Glutamine Q CAG	CGU Arginine R CGC CGA CGG	
	A	AUU Isoleucine I AUC AUA AUG Methionine start codon M	ACU Threonine T ACC ACA ACG	AAU Asparagine N AAC AAA Lysine K AAG	AGU Serine S AGC AGA Arginine R AGG	
	G	GUU Valine V GUC GUA GUG	GCU Alanine A GCC GCA GCG	GAU Aspartic acid D GAC GAA Glutamic acid E GAG	GGU Glycine G GGC GGA GGG	

Six reading frames

- Given a sequence of DNA in coding region, we can have 6 ways to translate the DNA to protein.

- Example:

+3

L V R T

+2

T C S Y

+1

N L F V

5' - A A C T T G T T C G T A C A - 3'

3' - T T G A A C A A G C A T G T - 5'

-1

K N T C

-2

S T R V

-3

V Q E Y

		Second base				
		U	C	A	G	
First base	U	UUU] Phenyl- alanine F UUC] UUA] Leucine L UUG]	UCU] UCC] Serine S UCA] UCG]	UAU] Tyrosine Y UAC] UAA] Stop codon UAG] Stop codon	UGU] Cysteine C UGC] UGA] Stop codon UGG] Tryptophan W	U C A G
	C	CUU] Leucine L CUC] CUA] CUG]	CCU] CCC] Proline P CCA] CCG]	CAU] Histidine H CAC] CAA] Glutamine Q CAG]	CGU] Arginine R CGC] CGA] CGG]	U C A G
	A	AUU] Isoleucine I AUC] AUA] AUG] Methionine start codon M	ACU] Threonine T ACC] ACA] ACG]	AAU] Asparagine N AAC] AAA] Lysine K AAG]	AGU] Serine S AGC] AGA] Arginine R AGG]	U C A G
	G	GUU] Valine V GUC] GUA] GUG]	GCU] Alanine A GCC] GCA] GCG]	GAU] Aspartic acid D GAC] Glutamic acid E GAA] GAG]	GGU] Glycine G GGC] GGA] GGG]	U C A G

Six reading frames

- Given a sequence of DNA in coding region, we can have 6 ways to translate the DNA to protein.

- Example:

• +3

L V R T

• +2

T C S Y

• +1

N L F V

5' -AACTTGTTCGTACA-3'

3' -TTGAACAAGCATGT-5'

• -1

K N T C

• -2

S T R V

• -3

V Q E Y

		Second base				
		U	C	A	G	
First base	U	UUU] Phenyl-alanine F UUC] UUA] Leucine L UUG]	UCU] UCC] Serine S UCA] UCG]	UAU] Tyrosine Y UAC] UAA] Stop codon UAG] Stop codon	UGU] Cysteine C UGC] UGA] Stop codon UGG] Tryptophan W	Third base U C A G U C A G U C A G U C A G
	C	CUU] Leucine L CUC] CUA] CUG]	CCU] CCC] Proline P CCA] CCG]	CAU] Histidine H CAC] CAA] Glutamine Q CAG]	CGU] Arginine R CGC] CGA] CGG]	
	A	AUU] Isoleucine I AUC] AUA] AUG] Methionine start codon M	ACU] Threonine T ACC] ACA] ACG]	AAU] Asparagine N AAC] AAA] Lysine K AAG]	AGU] Serine S AGC] AGA] Arginine R AGG]	
	G	GUU] Valine V GUC] GUA] GUG]	GCU] Alanine A GCC] GCA] GCG]	GAU] Aspartic acid D GAC] GAA] Glutamic acid E GAG]	GGU] Glycine G GGC] GGA] GGG]	

Six Reading Frames

- Given a sequence of DNA in coding region, we can have 6 ways to translated the DNA to protein

- Example:

- +3

L V R T

- +2

T C S Y

- +1

N L F V

5' -AACTTGTTCGTACA-3'

3' -TTGAACAAAGCA⁴TGT-5'

- 1

K N T C

- 2

S T R V

- 3

V Q E Y

		Second base				
		U	C	A	G	
First base	U	UUU] Phenyl-alanine F UUC] UUA] Leucine L UUG]	UCU] UCC] Serine S UCA] UCG]	UAU] Tyrosine Y UAC] UAA] Stop codon UAG] Stop codon	UGU] Cysteine C UGC] UGA] Stop codon UGG] Tryptophan W	U C A G
	C	CUU] Leucine L CUC] CUA] CUG]	CCU] CCC] Proline P CCA] CCG]	CAU] Histidine H CAC] CAA] Glutamine Q CAG]	CGU] Arginine R CGC] CGA] CGG]	U C A G
	A	AUU] Isoleucine I AUC] AUA] AUG] Methionine start codon M	ACU] Threonine T ACC] ACA] ACG]	AAU] Asparagine N AAC] AAA] Lysine K AAG]	AGU] Serine S AGC] AGA] Arginine R AGG]	U C A G
	G	GUU] Valine V GUC] GUA] GUG]	GCU] Alanine A GCC] GCA] GCG]	GAU] Aspartic acid D GAC] Glutamic acid E GAA] GAG]	GGU] Glycine G GGC] GGA] GGG]	U C A G

Six Reading Frames

- Given a sequence of DNA in coding region, we can have 6 ways to translated the DNA to protein

- Example:

- +3

L V R T

- +2

T C S Y

- +1

N L F V

5' -AACTTGTTCGTACA-3'

3' -TTGAACAAAGCATGT-5'

- 1

K N T C

- 2

S T R V

- 3

V Q E Y

		Second base				
		U	C	A	G	
First base	U	UUU] Phenyl-alanine F UUC] UUA] Leucine L UUG]	UCU] UCC] Serine S UCA] UCG]	UAU] Tyrosine Y UAC] UAA] Stop codon UAG] Stop codon	UGU] Cysteine C UGC] UGA] Stop codon UGG] Tryptophan W	U C A G
	C	CUU] Leucine L CUC] CUA] CUG]	CCU] CCC] Proline P CCA] CCG]	CAU] Histidine H CAC] CAA] Glutamine Q CAG]	CGU] Arginine R CGC] CGA] CGG]	U C A G
	A	AUU] Isoleucine I AUC] AUA] AUG] Methionine start codon M	ACU] Threonine T ACC] ACA] ACG]	AAU] Asparagine N AAC] AAA] Lysine K AAG]	AGU] Serine S AGC] AGA] Arginine R AGG]	U C A G
	G	GUU] Valine V GUC] GUA] GUG]	GCU] Alanine A GCC] GCA] GCG]	GAU] Aspartic acid D GAC] GAA] Glutamic acid E GAG]	GGU] Glycine G GGC] GGA] GGG]	U C A G

Six reading frames

- Given a sequence of DNA in coding region, we can have 6 ways to translate the DNA to protein.

- Example:

- +3
L V R T
- +2
T C S Y
- +1
N L F V

5' -AACTTGTTCGTACA-3'

3' -TTGAACAAAGCATGT-5'

- 1
K N T C
- 2
S T R V
- 3
V Q E Y

		Second base				
		U	C	A	G	
First base	U	UUU } Phenylalanine F UUC } UUA } Leucine L UUG }	UCU } Serine S UCC } UCA } UCG }	UAU } Tyrosine Y UAC } UAA } Stop codon UAG } Stop codon	UGU } Cysteine C UGC } UGA } Stop codon UGG } Tryptophan W	Third base U C A G U C A G U C A G U C A G
	C	CUU } Leucine L CUC } CUA } CUG }	CCU } Proline P CCC } CCA } CCG }	CAU } Histidine H CAC } CAA } Glutamine Q CAG }	CGU } Arginine R CGC } CGA } CGG }	
	A	AUU } Isoleucine I AUC } AUA } AUG } Methionine start codon M	ACU } Threonine T ACC } ACA } ACG }	AAU } Asparagine N AAC } AAA } Lysine K AAG }	AGU } Serine S AGC } AGA } Arginine R AGG }	
	G	GUU } Valine V GUC } GUA } GUG }	GCU } Alanine A GCC } GCA } GCG }	GAU } Aspartic acid D GAC } GAA } Glutamic acid E GAG }	GGU } Glycine G GGC } GGA } GGG }	

Exercise: Six reading frames

- Write down the six reading frames of the following DNA strand:

5'-CACTATCGCTTGTCC-3'

		Second base					
		U	C	A	G		
First base	U	UUU Phenyl-alanine F UUC UUA Leucine L UUG	UCU UCC Serine S UCA UCG	UAU Tyrosine Y UAC UAA Stop codon UAG Stop codon	UGU Cysteine C UGC UGA Stop codon UGG Tryptophan W	Third base	U
	C	CUU CUC Leucine L CUA CUG	CCU CCC Proline P CCA CCG	CAU Histidine H CAC CAA Glutamine Q CAG	CGU CGC Arginine R CGA CGG		C
	A	AUU Isoleucine I AUC AUA AUG Methionine start codon M	ACU ACC Threonine T ACA ACG	AAU Asparagine N AAC AAA Lysine K AAG	AGU Serine S AGC AGA Arginine R AGG		A
	G	GUU GUC Valine V GUA GUG	GCU GCC Alanine A GCA GCG	GAU Aspartic acid D GAC GAA Glutamic acid E GAG	GGU GGC Glycine G GGA GGG		G

Answer: Six reading frames

- +3
L S L V
- +2
T I A C
- +1
H Y R L S

5' -CACTATCGCTTGTCC-3'

3' -GTGATAGCGAACAGG-5'

- -1
V I A Q G
- -2
* R K D
- -3
S D S T

Problems of previous exercise

- No splicing
- The 5'- and 3'-untranslated regions (UTR) are missing
- The protein does not start at “start codon”, and does not finish at “stop codon”
- The protein is too short (usually a protein should contain several hundred amino acids)

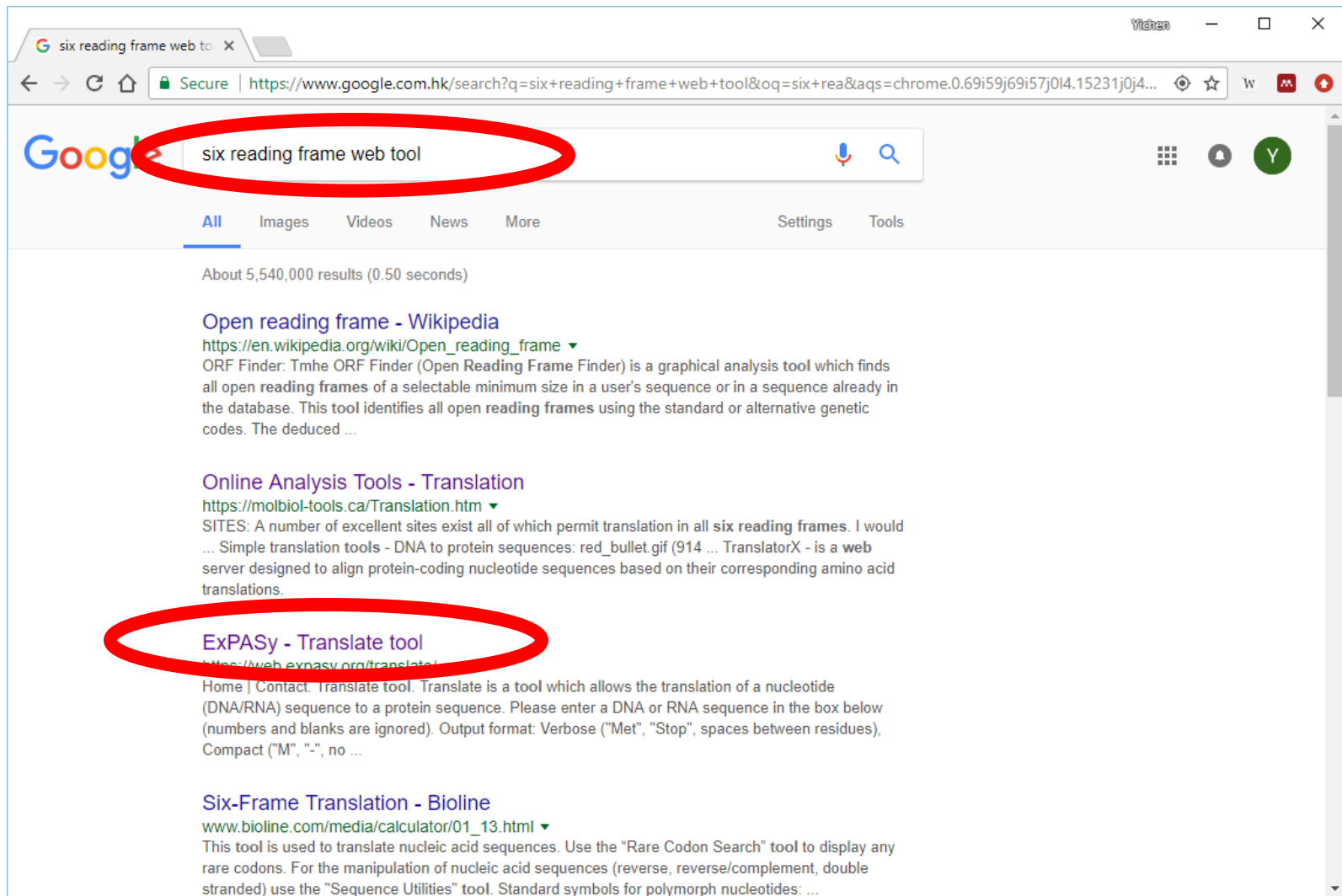
Demonstration: Web exercise

- Find a Web tool to perform six-frame translation on the sequence below:

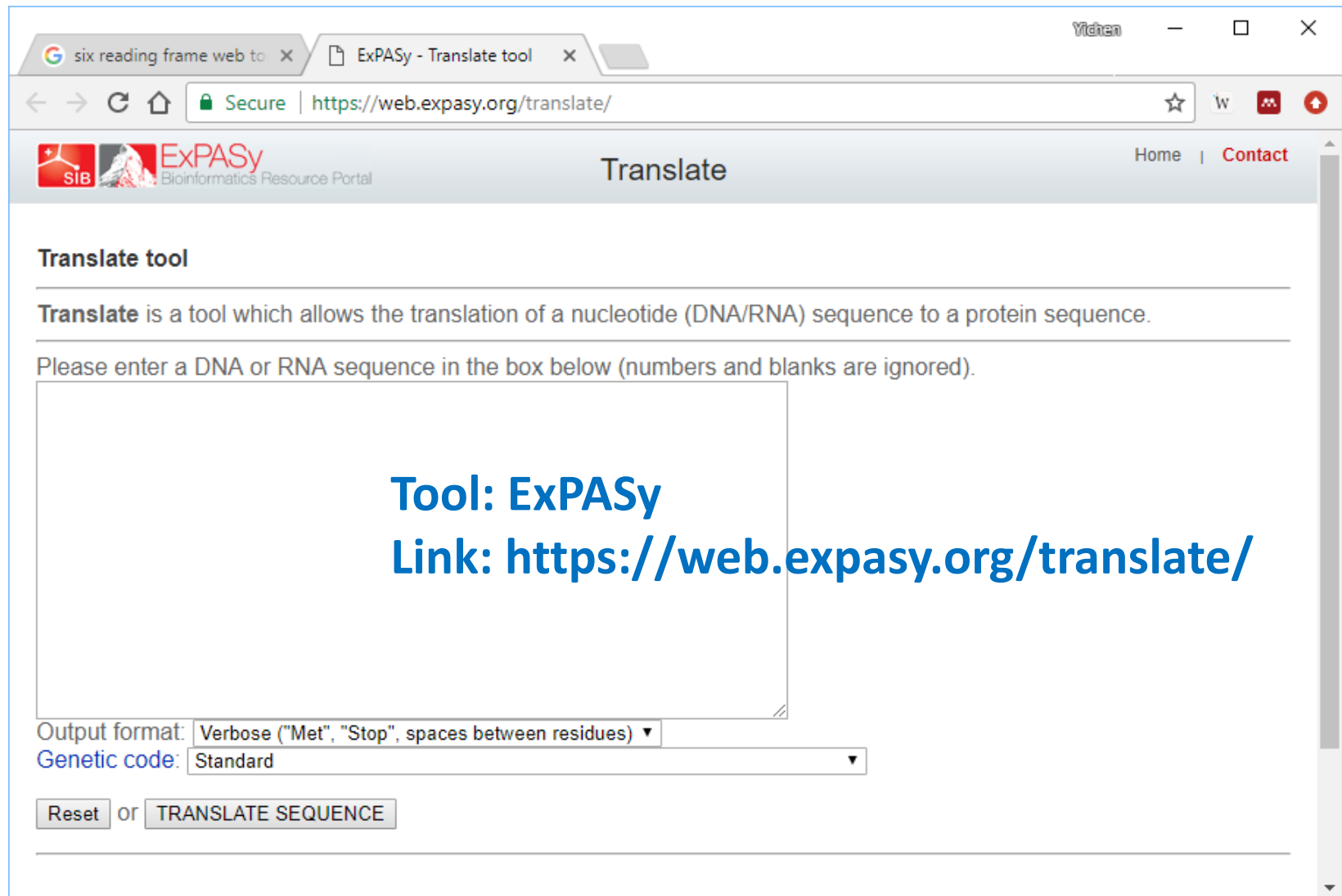
5' -CACTATCGCTTGTCC-3'

- Remember to record details of tools you used including the **input data**, **parameters** used, and full screen shots of the **outputs** of **every step**.

Step 1: Find tools from Google



Step 2: Describe the tool, 'ExPASy'



The screenshot shows a web browser window with the ExPASy Translate tool. The browser's address bar displays the URL <https://web.expasy.org/translate/>. The page header includes the ExPASy logo and the text "Bioinformatics Resource Portal". The main heading is "Translate". Below this, a description states: "Translate is a tool which allows the translation of a nucleotide (DNA/RNA) sequence to a protein sequence." A prompt asks the user to "Please enter a DNA or RNA sequence in the box below (numbers and blanks are ignored)." A large text box is provided for sequence input. Below the text box, there are two dropdown menus: "Output format:" set to "Verbose ('Met', 'Stop', spaces between residues)" and "Genetic code:" set to "Standard". At the bottom, there are two buttons: "Reset" and "TRANSLATE SEQUENCE".

Tool: ExPASy
Link: <https://web.expasy.org/translate/>

Step 3: State the input and parameters

The screenshot shows the ExPASy Translate tool interface. The browser address bar shows the URL <https://web.expasy.org/translate/>. The page title is "Translate". The ExPASy logo and "Bioinformatics Resource Portal" are visible. The "Translate tool" section describes the tool's function. The input field contains the sequence "CACTATCGCTTGTCC", which is circled in red. The output format is set to "Compact ('M', '-', no spaces)" and the genetic code is set to "Standard", both of which are circled in red. The "Reset" and "TRANSLATE SEQUENCE" buttons are at the bottom. To the right of the interface, the input sequence and parameters are listed: "Input: CACTATCGCTTGTCC", "Parameter:", "Output Format: Compact ('M', '-', no spaces)", and "Genetic code: Standard". A note at the bottom right explains that "M" means start codon and "-" means stop codon.

Translate tool

Translate is a tool which allows the translation of a nucleotide (DNA/RNA) sequence to a protein sequence.

Please enter a DNA or RNA sequence in the box below (numbers and blanks are ignored).

CACTATCGCTTGTCC

Input: CACTATCGCTTGTCC

Parameter:

Output Format: Compact ("M", "-", no spaces)

Genetic code: Standard

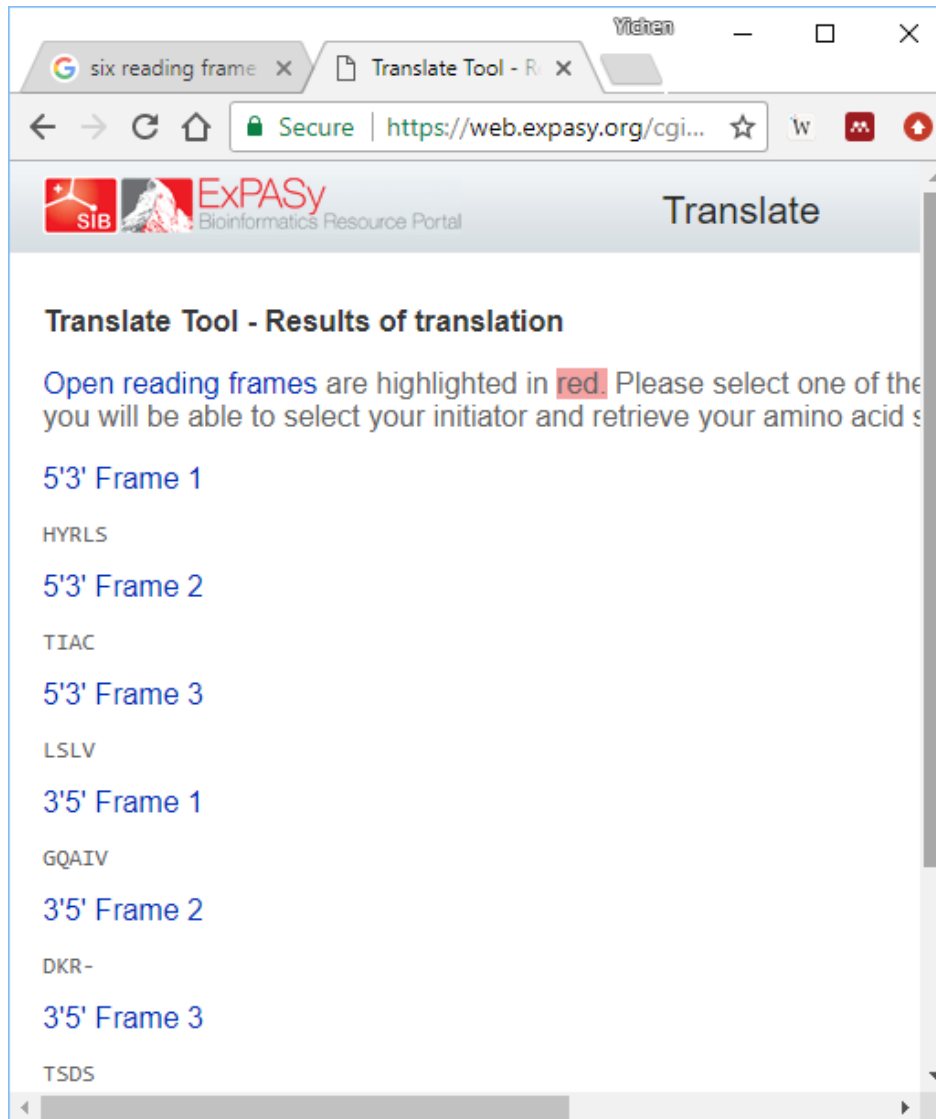
Output format: Compact ("M", "-", no spaces)

Genetic code: Standard

Reset or TRANSLATE SEQUENCE

NOTE: "M" means start codon, while "-" means stop codon

Step 4: Show the results



- Six reading frames:

– +1 HYRLS

– +2 TIAC

– +3 LSLV

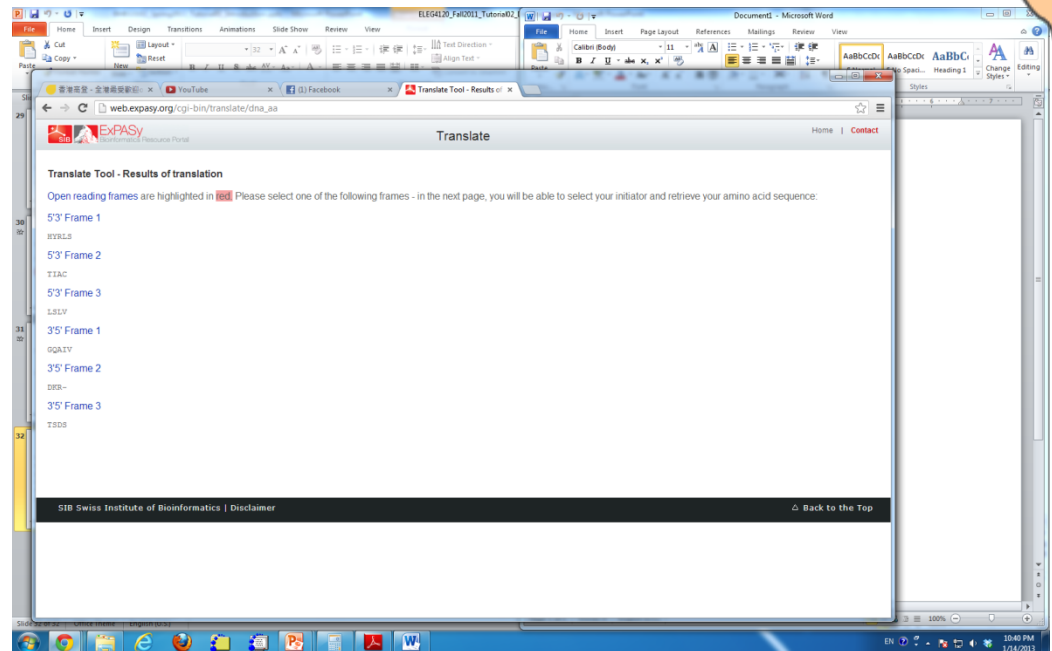
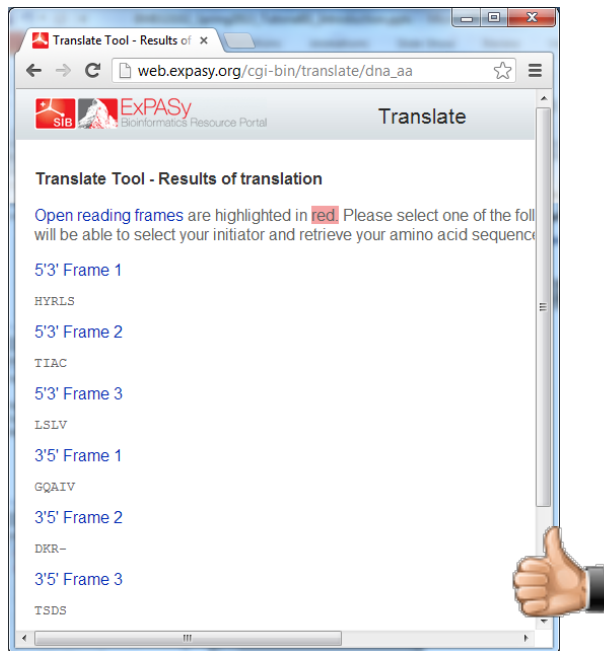
– -1 GQAIV

– -2 DKR-

– -3 TSDS

Tips for making screenshot

- No extra background
- Good quality of words / figures
- Reasonable size of screenshots

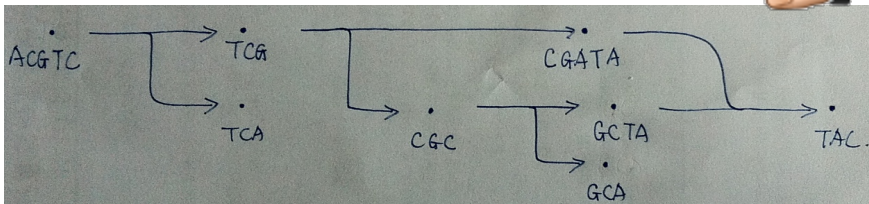


Tools for making screenshot

- For Windows users:
 - Press Alt-PrtScn, copy it to Word/PowerPoint, then crop (Format > Crop)
 - Use Snipping Tool to select the desired area
- For Mac users:
 - Press Shift+Command+4 to select the desired area
 - <https://support.apple.com/en-hk/HT201361>

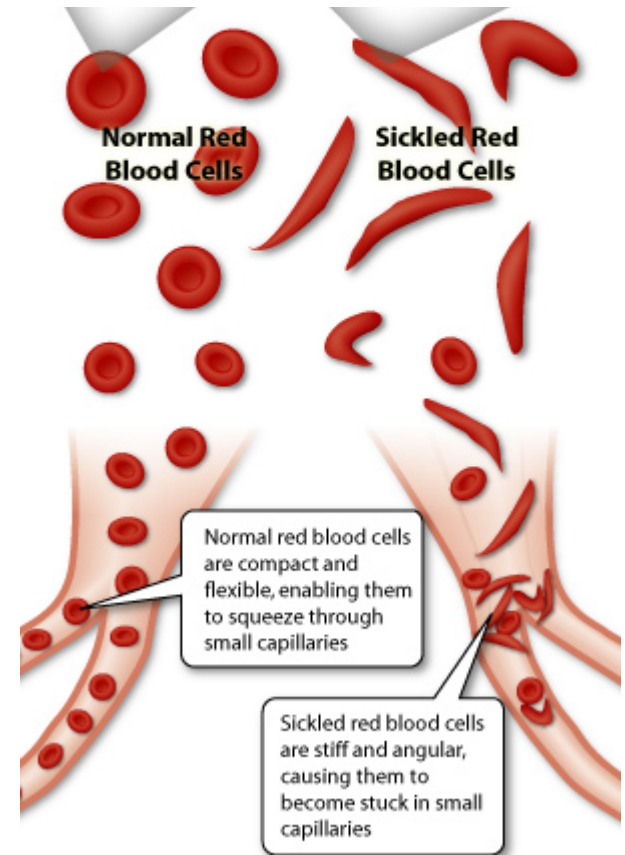
Tips for submitting assignments

- Only accept **softcopy** for the assignment
- Submit your work using the Blackboard System
- When you are working with some *written problems*, you may:
 - Type and draw the figures in the computers, or
 - Write on paper and scan your work, or
 - Write on paper, take a photo and attach the photo in Word
- However,



Case study: Mutation and disease

- Classic example:
 Sickle-cell Anemia
 - Normal red blood cells are round and flexible, able to carry oxygen and travel through blood vessels
 - People with Sickle-cell Anemia have red blood cells that look like a sickle and get stuck in blood vessels



Case study: Mutation and disease

- What went wrong?
 - Hemoglobin molecules, which are responsible for carrying oxygen, form a long chain in affected people
 - Why? There is a mutation in one of the hemoglobin genes on chromosome 11

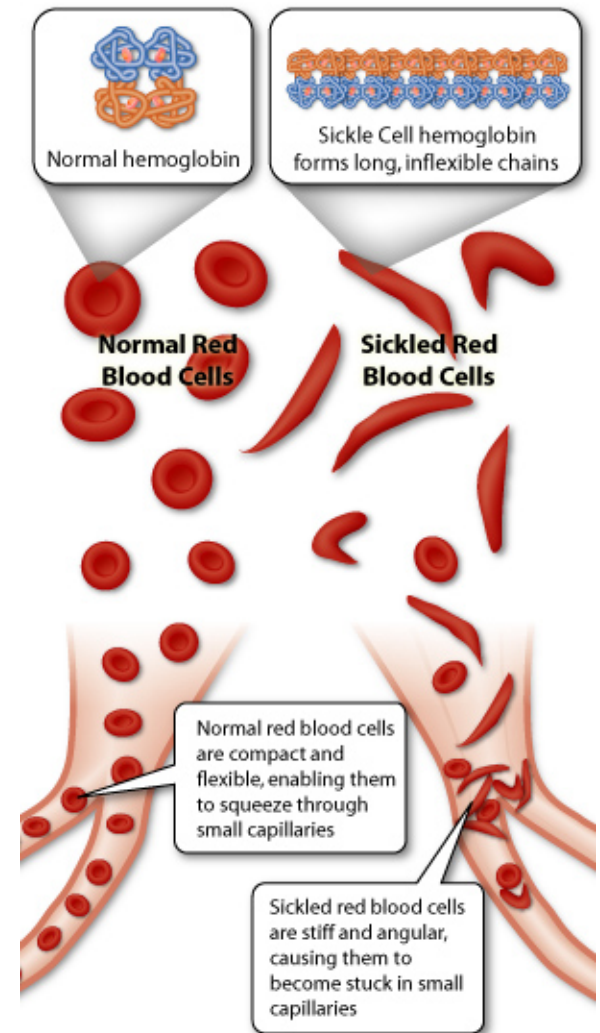


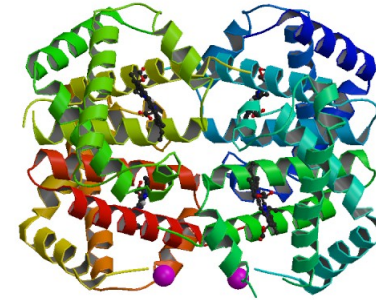
Image credit: <http://learn.genetics.utah.edu/content/disorders/whataregd/sicklecell/images/sicklecell.jpg>

Case study: Mutation and disease

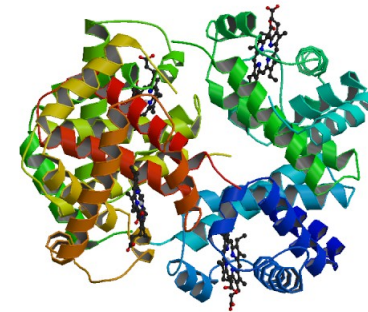
- Mutation in DNA causes change in protein structure

HBB Sequence in Normal Adult Hemoglobin (Hb A):							
Nucleotide	CTG	ACT	CCT	GAG	GAG	AAG	TCT
Amino Acid	Leu	Thr	Pro	Glu	Glu	Lys	Ser
	3			6			9

HBB Sequence in Mutant Adult Hemoglobin (Hb S):							
Nucleotide	CTG	ACT	CCT	GTG	GAG	AAG	TCT
Amino Acid	Leu	Thr	Pro	Val	Glu	Lys	Ser
	3			6			9



Normal human hemoglobin



A common mutated form of human hemoglobin that causes Sickle-cell Anemia

Image credit: Protein Data Bank, entries 4HHB and 2HBS,
http://www.ornl.gov/sci/techresources/Human_Genome/posters/chromosome/Gifs/HBBmutseq_2.gif

Find the gene that causes the disease

sickle cell anemia - Google Search x Sick cell disease - Genetics Home Reference x

https://www.google.com/search?newwindow=1&rlz=1C1GCEU_enHK822HK823&ei=hNg-XNDZFM...

Google

sickle cell anemia

All Images News Videos Books More Settings Tools

About 16,500,000 results (0.54 seconds)

Sickle cell disease is a group of disorders that affects hemoglobin, the molecule in red blood **cells** that delivers oxygen to **cells** throughout the body. People with this disorder have atypical hemoglobin molecules called hemoglobin S, which can distort red blood **cells** into a **sickle**, or crescent, shape.



Sickle cell disease - Genetics Home Reference - NIH
<https://ghr.nlm.nih.gov/condition/sickle-cell-disease>

About this result Feedback

People also ask

- What happens to a person with sickle cell anemia? ▾
- Can a white person have sickle cell anemia? ▾
- How does sickle cell anemia make you feel? ▾
- Is Thalassemia the same as sickle cell anemia? ▾

Feedback

Sickle cell anemia - Symptoms and causes - Mayo Clinic
<https://www.mayoclinic.org/diseases-conditions/sickle-cell-anemia/.../syc-20355876> ▾

Mar 8, 2018 - In sickle cell anemia, the abnormal hemoglobin causes red blood cells to become rigid,

Find the gene that causes the disease

The screenshot shows a web browser window with two tabs: 'sickle cell anemia - Google Search' and 'Sickle cell disease - Genetics Home Reference'. The address bar shows the URL 'https://ghr.nlm.nih.gov/condition/sickle-cell-disease#genes'. The page header includes the NIH logo and 'U.S. National Library of Medicine', a 'Share This Page' button, the 'Genetics Home Reference' logo, and the tagline 'Your Guide to Understanding Genetic Conditions'. A 'MENU' button and a search icon are also present. The main content area is titled 'Sickle cell disease' and includes buttons for 'Printable PDF', 'Open All', and 'Close All'. Below the title are three expandable sections: 'Description', 'Frequency', and 'Causes'. The 'Causes' section is expanded and circled in red, containing the text: 'Mutations in the [HBB](#) gene cause sickle cell disease.' Below this, a paragraph explains that hemoglobin consists of four protein subunits (two alpha-globin and two beta-globin) and that the *HBB* gene provides instructions for making beta-globin. It states that various mutations in the *HBB* gene lead to different versions of beta-globin, including hemoglobin S (HbS), hemoglobin C (HbC), and hemoglobin E (HbE). A 'Related Information' sidebar on the right contains links to 'What is a gene?', 'What is a gene mutation and how do mutations occur?', and 'How can gene mutations affect...'. The page footer indicates 'BMEG3012 Bioinformatics Tutorial Notes | Spring 2021'.

sickle cell anemia - Google Search x Sickle cell disease - Genetics Home Reference x

https://ghr.nlm.nih.gov/condition/sickle-cell-disease#genes

NIH U.S. National Library of Medicine Share This Page

Genetics Home Reference Your Guide to Understanding Genetic Conditions

MENU

Sickle cell disease

Printable PDF Open All Close All

► Description

► Frequency

▼ Causes

Mutations in the [HBB](#) gene cause sickle cell disease.

Hemoglobin consists of four protein subunits, typically, two subunits called alpha-globin and two subunits called beta-globin. The *HBB* gene provides instructions for making beta-globin. Various versions of beta-globin result from different mutations in the *HBB* gene. One particular *HBB* gene mutation produces an abnormal version of beta-globin known as hemoglobin S (HbS). Other mutations in the *HBB* gene lead to additional abnormal versions of beta-globin such as hemoglobin C (HbC) and hemoglobin E (HbE). *HBB* gene

Related Information

[What is a gene?](#)

[What is a gene mutation and how do mutations occur?](#)

[How can gene mutations affect...](#)

UCSC Genome Browser

The screenshot shows a Google search interface with the query 'ucsc genome browser'. The top result, 'UCSC Genome Browser Home', is circled in red. Below the main search results, there are several related links including 'Genome Browser', 'Genome Browser Mirror Sites', 'Training', 'Citing Us', 'Genome Browser User's Guide', 'Contact Us', and 'UCSC Genome Browser - Wikipedia'. On the right side, there is a knowledge panel for 'UCSC Genome Browser' with a description, primary citation, and research center information.

ucsc genome browser - Google S x +

https://www.google.com/search?q=ucsc+genome+browser&rlz=1C1GCEU_enHK822HK823&oq=ucsc+genome+&aqs=c...

Google

ucsc genome browser

All Images Videos News Maps More Settings Tools

About 1,690,000 results (0.49 seconds)

UCSC Genome Browser Home

https://genome.ucsc.edu/

UCSC Genome Browser. ... combine data sources from the Genome Browser database. Gene Sorter ... run the Genome Browser on your laptop or server.

You've visited this page many times. Last visit: 1/9/19

Genome Browser
UCSC Genome Browser assembly
ID: hg38. Sequencing/Assembly ...

Genome Browser Mirror Sites
UCSC Genome Browser Mirror Sites.
Sponsored mirrors. In ...

Training
Online training and tutorials. Our
video tutorials address some ...

[More results from ucsc.edu »](#)

Citing Us
Citing the UCSC Browser in a
Publication or Web Page ...

Genome Browser User's Guide
Genome Browser User Guide.
Contents. What does the ...

Contact Us
Contact Us. UCSC Genome
Informatics Group UCSC ...

UCSC Genome Browser - Wikipedia
https://en.wikipedia.org/wiki/UCSC_Genome_Browser ▼
The UCSC Genome Browser is an on-line, and downloadable, genome browser hosted by the University of California, Santa Cruz (UCSC). It is an interactive ...

UCSC Genome Browser

The UCSC Genome Browser is an on-line, and downloadable, genome browser hosted by the University of California, Santa Cruz. [Wikipedia](#)

Description: The UCSC Genome Browser

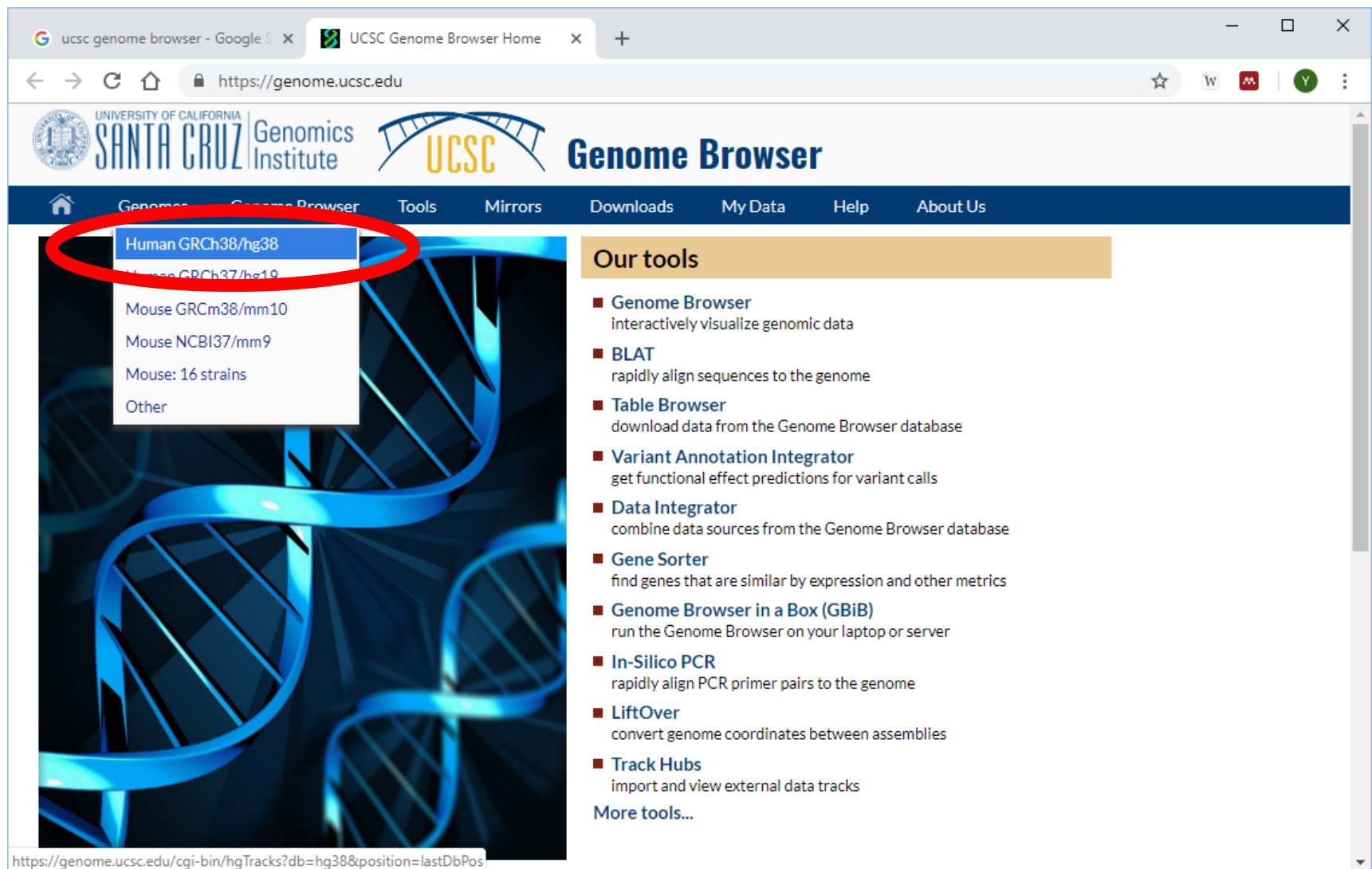
Primary citation: Tyner & al. (2016)

Research center: [University of California Santa Cruz](#)

People also search for View 5+

BLAST Mapping via blat

UCSC Genome Browser



The screenshot shows the UCSC Genome Browser homepage in a web browser. The browser's address bar displays `https://genome.ucsc.edu`. The page header includes the University of California Santa Cruz Genomics Institute logo and the UCSC Genome Browser title. A navigation bar contains links for Home, Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Help, and About Us. A dropdown menu is open under the 'Genomes' link, listing options: Human GRCh38/hg38 (highlighted with a red circle), Human GRCh37/hg19, Mouse GRCm38/mm10, Mouse NCBI37/mm9, Mouse: 16 strains, and Other. The main content area features a large blue DNA double helix graphic on the left and a 'Our tools' section on the right. The 'Our tools' section lists various tools with brief descriptions: Genome Browser, BLAT, Table Browser, Variant Annotation Integrator, Data Integrator, Gene Sorter, Genome Browser in a Box (GBiB), In-Silico PCR, LiftOver, and Track Hubs. A 'More tools...' link is at the bottom of this list. The browser's status bar at the bottom shows the URL `https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&position=lastDbPos`.

ucsc genome browser - Google S x UCSC Genome Browser Home x +

← → ↻ 🏠 🔒 `https://genome.ucsc.edu` ☆ W Y

UNIVERSITY OF CALIFORNIA SANTA CRUZ Genomics Institute UCSC Genome Browser

🏠 Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

Human GRCh38/hg38
Human GRCh37/hg19
Mouse GRCm38/mm10
Mouse NCBI37/mm9
Mouse: 16 strains
Other

Our tools

- **Genome Browser**
interactively visualize genomic data
- **BLAT**
rapidly align sequences to the genome
- **Table Browser**
download data from the Genome Browser database
- **Variant Annotation Integrator**
get functional effect predictions for variant calls
- **Data Integrator**
combine data sources from the Genome Browser database
- **Gene Sorter**
find genes that are similar by expression and other metrics
- **Genome Browser in a Box (GBiB)**
run the Genome Browser on your laptop or server
- **In-Silico PCR**
rapidly align PCR primer pairs to the genome
- **LiftOver**
convert genome coordinates between assemblies
- **Track Hubs**
import and view external data tracks

More tools...

`https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&position=lastDbPos`

Search on UCSC Genome Browser

The screenshot displays the UCSC Genome Browser interface. The browser's address bar shows the URL: <https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=&virtModeTyp...>. The main header reads "UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly". Below this, navigation controls include "move" and "zoom in/out" buttons. The search bar contains the coordinates "chr1:11,102,837-11,267,747" and the gene name "HBB", which is highlighted with a red box. A red arrow points to the search bar with the text "Type 'HBB'". The main track displays the HBB gene structure, including the "New haplotype sequences to GRCh38 Reference Sequence" and "Patches to GRCh38 Reference Sequence". Below the main track, there are various tracks such as "OMIM Alleles", "Cons 100 Verts", and "Simple Nucleotide Polymorphisms (dbSNP 151)". The bottom section, titled "Mapping and Sequencing", contains a grid of tracks with dropdown menus for each, including "Base Position", "GRC Patch Release", "Alt Map...", "Assembly", "Centromeres", "Chromosome Band", "Clone Ends", "FISH Clones", "Gap", "GC Percent", "GRC Contigs", "GRC Incident", "Hg19 Diff", "INSDC", "LRG Regions", "Mappability...", "RefSeq Acc", "Restr Enzymes", "Scaffolds", "Short Match", "STS Markers", and "hide" buttons for each track.

View target gene on Genome Browser

ucsc genome browser - Google S x Human hg38 HBB UCSC Genome x +

https://genome.ucsc.edu/cgi-bin/hgTracks?hgtgroup_map_close=0&hgtgroup_genes_close=0&hgtgroup_phenDis_close...

Known Genes

HBB (ENST00000647020.1) at chr11:5225464-5227197 - Involved in oxygen transport from the lung to the various peripheral tissues. (from UniProt P68871)

HBB (ENST00000633227.1) at chr11:5225467-5227072 - hemoglobin subunit beta (from HGNC HBB)

HBB (ENST00000633227.1) at chr11:5225467-5227072 - hemoglobin subunit beta (from HGNC HBB)

HBB (ENST00000475226.1) at chr11:5225655-5226823 - hemoglobin subunit beta (from HGNC HBB)

HBB (ENST00000380315.2) at chr11:5226620-5229395 - Belongs to the globin family. (from UniProt F8W6P5)

HBB (ENST00000335295.4) at chr11:5225464-5227071 - Homo sapiens hemoglobin subunit beta (HBB), mRNA. (from RefSeq NM_000518)

HBD (ENST00000380299.3) at chr11:5232838-5234648 - Homo sapiens hemoglobin subunit delta (HBD), mRNA. (from RefSeq NM_000519)

RBM17 (ENST00000446108.5) at chr10:6089346-6117457 - Homo sapiens RNA binding motif protein 17 (RBM17), transcript variant 2, mRNA. (from RefSeq NM_00114)

RBM17 (ENST00000379888.8) at chr10:6088987-6116880 - Homo sapiens RNA binding motif protein 17 (RBM17), transcript variant 1, mRNA. (from RefSeq NM_03290)

HBA1 (ENST00000320868.9) at chr16:176680-177522 - Homo sapiens hemoglobin subunit alpha 1 (HBA1), mRNA. (from RefSeq NM_000558)

HBA2 (ENST00000251595.11) at chr16:172876-173710 - Homo sapiens hemoglobin subunit alpha 2 (HBA2), mRNA. (from RefSeq NM_000517)

HBBP1 (ENST00000454892.2) at chr11:5241105-5243533 - hemoglobin subunit beta pseudogene 1 (from HGNC HBBP1)

HBBP1 (ENST00000433329.1) at chr11:5242120-5243537 - Homo sapiens hemoglobin subunit beta pseudogene 1 (HBBP1), non-coding RNA. (from RefSeq NR_001589)

TMEM158 (ENST00000503771.1) at chr3:45224466-45226278 - Homo sapiens transmembrane protein 158 (gene/pseudogene) (TMEM158), mRNA. (from RefSeq NM_015444)

NCBI RefSeq genes, curated subset (NM_*, NR_*, and YP_*)

NM_000518.4 at chr11:5225466-5227071

NR_001589.1 at chr11:5241954-5243592

RefSeq Genes

HBB at chr11:5225464-5227071 - (NM_000518) hemoglobin subunit beta

HBBP1 at chr11:5241955-5243592 - (NR_001589)

NCBI RefSeq other annotations (not NM_*, NR_*, XM_*, XR_*, or YP_*)

KRT89P at chr12:52340997-52354463

Non-Human RefSeq Genes

View target gene on Genome Browser

ucsc genome browser - Google S x Human hg38 chr11:5225464-522 x +

https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=&virtModeTyp...

Genomes Genome Browser Tools Mirrors Downloads My Data View Help About Us

UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x 100x

chr11:5,225,464-5,227,197 1,734 bp. chr11:5,225,464-5,227,197 go

chr11 (p15.4) p15.4 15.1 p13 11p12 11.2 13.4 11q14.1 q21 22.1 q22.3 q23.3 q25

Scale chr11: 500 bases | hg38 5,225,000 | 5,226,000 | 5,227,000

New haplotype sequences to GRCh38 Reference Sequence
Patches to GRCh38 Reference Sequence
GENCODE v29 Comprehensive Transcript Set (only Basic displayed by default)

HBB HBB OMIM Allelic Variants OMIM Allelic Variants

OMIM Alleles 4.68 -
Cons 100 Verts -4.5 -
Common SNPs(151)

Simple Nucleotide Polymorphisms (dbSNP 151) Found in >= 1% of Samples

move start < 2.0 > move end < 2.0 >

track search default tracks default order hide all add custom tracks track hubs configure multi-region reverse resize refresh

collapse all

Use drop-down controls below and press refresh to alter tracks displayed.
Tracks with lots of items will automatically be displayed in more compact modes.

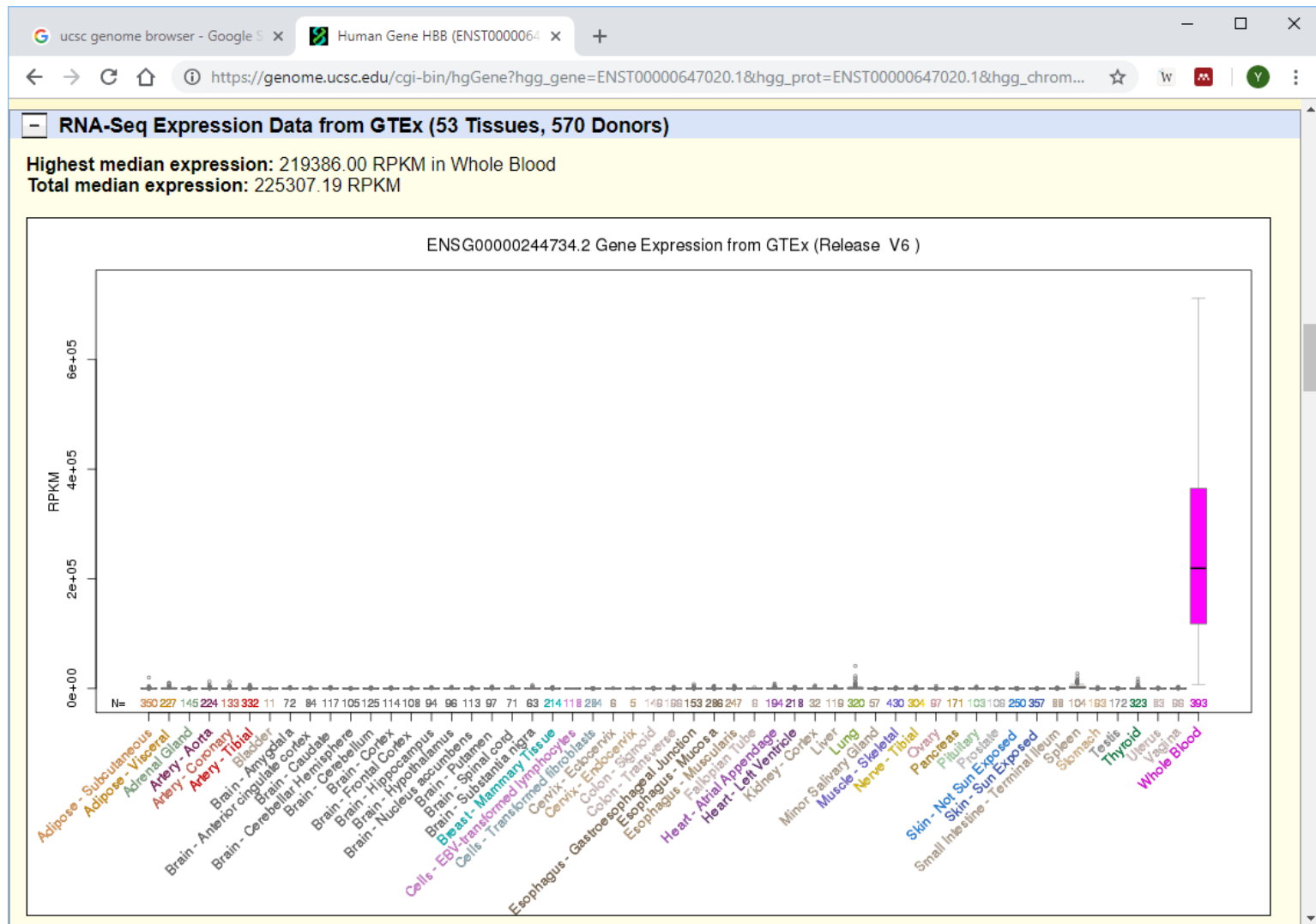
expand all

Mapping and Sequencing

refresh

Base Position dense ▾	GRC Patch Release full ▾	Alt Map... hide ▾	Assembly hide ▾	Centromeres hide ▾	Chromosome Band hide ▾
Clone Ends hide ▾	FISH Clones hide ▾	Gap hide ▾	GC Percent hide ▾	GRC Contigs hide ▾	GRC Incident hide ▾
Hg19 Diff hide ▾	INSDC hide ▾	LRG Regions hide ▾	Mappability... hide ▾	RefSeq Acc hide ▾	Restr Enzymes hide ▾

View RNA expression level



Change settings of Genome Browser

The screenshot shows the UCSC Genome Browser interface with the following tracks and settings:

- Comparative Genomics** (refresh button):
 - Conservation: full
 - Cons 7 Verts: hide
 - Cons 20 Mammals: hide
 - Cons 30 Primates: hide
 - Primate Chain/Net: hide
 - Placental Chain/Net: hide
 - Vertebrate Chain/Net: hide
 - Chicken Chain/Net: hide
- Variation** (refresh button):
 - New Common SNPs(151): dense
 - Common SNPs(150): hide
 - Common SNPs(147): hide
 - Common SNPs(146): hide
 - Common SNPs(144): hide
 - Common SNPs(142): hide
 - New All SNPs(151)**: hide (circled in red with a blue arrow and text "click here")
 - All SNPs(150): hide
 - All SNPs(147): hide
 - All SNPs(146): hide
 - All SNPs(144): hide
 - All SNPs(142): hide
 - All SNPs(141): hide
 - New Flagged SNPs(151): hide
 - Flagged SNPs(150): hide
 - Flagged SNPs(147): hide
 - Flagged SNPs(146): hide
 - Flagged SNPs(144): hide
 - Flagged SNPs(142): hide
 - Flagged SNPs(141): hide
 - New Mult. SNPs(151): hide
 - Mult. SNPs(150): hide
 - Mult. SNPs(147): hide
 - Mult. SNPs(146): hide
 - Mult. SNPs(144): hide
 - Mult. SNPs(142): hide
 - Mult. SNPs(141): hide
 - DGV Struct Var: hide
- Repeats** (refresh button):
 - RepeatMasker: hide
 - Interrupted Rpts: hide
 - RepeatMasker Viz.: hide
 - Microsatellite: hide
 - Segmental Dups: hide
 - Self Chain: hide
 - Simple Repeats: hide
 - WM + SDust: hide

A "refresh" button is located at the bottom center of the interface.

Change settings of Genome Browser

ucsc genome browser - Google S x All SNPs(151) Track Settings rs334 RefSNP Report - dbSNP - x +

https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=707068335_uwhm1yLANg1AN77tDfiSE6rQGdad&c=chr11&g=snp151

Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

All SNPs(151) Track Settings

Simple Nucleotide Polymorphisms (dbSNP 151) (▲All Variation tracks)

Display mode: pack Submit

Include Chimp state and observed human alleles in name: ☐
(If enabled, chimp allele is displayed first, then '>', then human alleles).

Show alleles on strand of reference genome reported by dbSNP: ☐

+ **Use Gene Tracks for Functional Annotation**

+ **Filtering Options**

- **Coloring Options**

SNP Feature for Color Specification: Function Set defaults

The selected "Feature for Color Specification" above has the selection of colors below for each attribute. Only the color options for the feature selected above will be used to color items; color options for other features will not be shown. If a SNP has more than one of these attributes, the stronger color will override the weaker color. The order of colors, from strongest to weakest, is red, green, black, gray, and black.

Unknown	black	Locus	black	Coding - Synonymous	green	Coding - Non-Synonymous	red
Untranslated	black	Intron	black	Splice Site	black		

[View table schema](#)

Data last updated: 2018-04-19

Show the non-synonymous mutations

ucsc genome browser - Google x Human hg38 chr11:5225464-522 rs334 RefSNP Report - dbSNP - x +

https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=&virtModeTyp...

Genomes Genome Browser Tools Mirrors Downloads My Data View Help About Us

UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly

move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x 100x

chr11:5,225,464-5,227,197 1,734 bp. enter position, gene symbol, HGVS or search terms go

chr11 (p15.4) p15.4 15.1 p13 11p12 11.2 13.4 11q14.1 q21 22.1 q22.3 q23.3 q25

Scale chr11: 500 bases | hg38 5,226,000 | 5,226,500 | 5,227,000

New haplotype sequences to GRCh38 Reference Sequence
Patches to GRCh38 Reference Sequence
GENCODE v29 Comprehensive Transcript Set (only Basic displayed by default)

HBB HBB

OMIM Alleles OMIM Allelic Variants

Vertebrate Multiz Alignment & Conservation (100 Species)

Simple Nucleotide Polymorphisms (dbSNP 151) Found in >= 1% of Samples

Common SNPs (151)

rs1420779550	rs53750694	rs193922560	rs558554234	rs576852971	rs999594627	rs946957516	rs35002698	rs11549407	rs33956879	rs63750400
rs1224765274	rs34240441	rs1421375702	rs908372322	rs35920381	rs187507944	rs996887983	rs1803195	rs1135101	rs33915112	rs63750953
rs528009939	rs34407387	rs181743523	rs941178010	rs922176223	rs952327061	rs7946748	rs35890360	rs1141370	rs35474880	rs33931746
rs1473799128	rs35669628	rs911937119	rs1461234906	rs34214308	rs532489997	rs7480526	rs33918483	rs1135071	rs33936254	rs371164553
rs753171158	rs35492035	rs944935116	rs773343845	rs553945255	rs985479299	rs373152208	rs33985847	rs35004220	rs34716811	rs261864518
rs989928447	rs35825479	rs63750433	rs966818537	rs1318746626	rs765573500	rs974387932	rs33969853	rs770452060	rs33931746	rs33931746
rs33985472	rs34095019	rs193922561	rs1392082436	rs892640661	rs753003330	rs368176939	rs35902963	rs11703179	rs334	rs339418
rs63750954	rs33946775	rs1257175417	rs1172356241	rs1334784243	rs113152027	rs35630218	rs34718174	rs9851334	rs713040	rs637518
rs281864532	rs34398074	rs1232651664	rs978207135	rs908075557	rs928354330	rs751644989	rs33947415	rs74718	rs341065501	rs33931746
rs35949130	rs33953406	rs1468940510	rs1241261843	rs1250456244	rs879578966	rs759795721	rs36187977	rs962266675	rs33931746	rs33931746
rs63751128	rs63750320	rs1489533338	rs554914198	rs190369729	rs892029919	rs368604295	rs34151706	rs377447760	rs762692752	rs33931746
rs63750205	rs34188626	rs346909599	rs1373924860	rs1364368636	rs947313558	rs1328092438	rs34974709	rs372778054	rs34784528	rs33931746
rs33976907	rs63749615	rs930802593	rs1417673135	rs162729393	rs375049787	rs1353813923	rs33990253	rs369341845	rs63750628	rs33931746
rs281864905	rs33950778	rs1083790835	rs569636767	rs1313875484	rs577036437	rs10768583	rs33991472	rs773415134	rs35352549	rs33931746
rs1228808616	rs33910209	rs1315014741	rs1427204895	rs964137770	rs565840445	rs756567448	rs35395625	rs375672426	rs34308195	rs33931746
rs34171453	rs33971634	rs533807852	rs911884715	rs908980587	rs1200964989	rs770630710	rs35276874	rs372716225	rs386134236	rs33931746
rs1264983767	rs33957286	rs1184042209	rs1337114930	rs1001129219	rs529931134	rs778382385	rs33988732	rs372467642	rs1463214708	rs33931746
rs770911771	rs34139813	rs63751175	rs63750105	rs1470539687	rs973955533	rs35614355	rs35094013	rs776904725	rs1418232313	rs33931746
rs34029390	rs34502690	rs1291160651	rs1194851586	rs1363355137	rs996301151	rs1799901	rs35171933	rs781015056	rs921599411	rs33931746
rs981033731	rs63750022	rs956857115	rs1468076084	rs112408237	rs915743646	rs35099062	rs34960334	rs369168168	rs954327894	rs33931746
rs193922549	rs33925391	rs1444028845	rs1405170526	rs551285871	rs948460713	rs63750842	rs34037627	rs565131843	rs281864525	rs33931746
rs1356409255	rs34643844	rs886156115	rs1450341299	rs1055395965	rs981149711	rs63750283	rs35165357	rs63751076	rs34598529	rs33931746
rs1223731404	rs33983276	rs1011427406	rs1251286087	rs531567967	rs1371816847	rs33921589	rs33969727	rs41486348	rs33980857	rs33931746
rs745310118	rs35238478	rs1461634734	rs1322482679	rs1484625119	rs1392181924	rs35209591	rs34676051	rs41539955	rs397509430	rs33931746
rs369101035	rs33987957	rs1005042281	rs1200755137	rs537557714	rs191037579	rs33937393	rs63750336	rs369482557	rs33981098	rs33931746
rs1265505530	rs34303736	rs35328027	rs1340330454	rs1428270127	rs934286084	rs33965000	rs35519054	rs372001957	rs33931746	rs33931746

Click on rs334

More about the SNP rs334

ucsc genome browser - Google S x Simple Nucleotide Polymorphism x rs334 RefSNP Report - dbSNP - N x +

https://genome.ucsc.edu/cgi-bin/hgc?hgsid=707068335_uwhm1yLAnG1AN77tDfISE6rQGdad&c=chr11&l=5225463&r=5...

Genomes Genome Browser Tools Mirrors Downloads My Data Help About Us

Simple Nucleotide Polymorphisms (dbSNP 151)

dbSNP build 151 rs334

dbSNP: **rs334** [Click here](#)
rs334 build 151 chr11:5227002-5227002
Band: 11p15.4
Genomic Size: 1
[View DNA for this feature](#) (hg38/Human)

strand: -
Observed: A/C/G/T
Reference allele: A

[Class](#) single
[Validation](#) by-cluster,by-frequency,by-1000genomes
[Function](#) [missense_variant](#)
[Molecule Type](#) genomic
[Average Heterozygosity](#) 0.009 +/- 0.065
[Weight](#) 1
[Submitter Handles](#) [1000GENOMES](#), [AFFY](#), [APPLERA_GI](#), [CGAP-GAI](#), [CLINSEQ_SNP](#), [COMPLETE_GENOMICS](#), [EVA_EXAC](#), [EXOME_CHIP](#), [GNOMAD](#), [HBVAR](#), [HUMAN_LONGEVITY](#), [ILLUMINA](#), [ITMI](#), [JMKIDD_LAB](#), [LEE](#), [MILLER_NIDDK](#), [NCBI](#), [NHLBI-ESP](#), [OMIM-CURATED-RECORDS](#), [PERLEGEN](#), [RSG_UW](#), [SEATTLESEQ](#), [SEQUENOM](#), [TOPMED](#), [WEILL_CORNELL_DGM](#)
[Allele Frequencies](#) A: 99.471% (125679 / 126348); T: 0.530% (669 / 126348)

Miscellaneous properties annotated by dbSNP:

HBB Sequence in Normal Adult Hemoglobin (Hb A):

Nucleotide	CTG	ACT	CCG	GAG	GAG	AAG	TCT
Amino Acid	Leu	Thr	Pro	Glu	Glu	Lys	Ser
	3			6			9

HBB Sequence in Mutant Adult Hemoglobin (Hb S):

Nucleotide	CTG	ACT	CCG	GTG	GAG	AAG	TCT
Amino Acid	Leu	Thr	Pro	Val	Glu	Lys	Ser
	3			6			9

Go to dbSNP to view the SNP

The screenshot shows a web browser window with the URL <https://www.ncbi.nlm.nih.gov/snp/rs334>. The page header includes the NIH logo and the text "U.S. National Library of Medicine National Center for Biotechnology Information". A "Log in" button is in the top right. Below the header, the "dbSNP Short Genetic Variations" logo is on the left, and a search bar with the text "Search for rs" and a "Search" button is on the right. Below the search bar, it says "Example: rs268". The main content area is titled "Reference SNP (rs) Report". On the left, there is a "Switch to classic site" link and a "rs334" label. On the right, there is a "Download" link and social media icons. Below these, it says "Current Build 152 Released October 2, 2018". The main content is a table with two columns. The left column has labels: "Organism", "Position", "Alleles", "Variation Type", and "Frequency". The right column has labels: "Clinical Significance", "Gene : Consequence", "Publications", and "Genomic View". The "Genomic View" label is circled in red, and the link "See rs on genome" is also circled in red.

Organism	<i>Homo sapiens</i>	Clinical Significance	Reported in ClinVar
Position	chr11:5227002 (GRCh38.p12) ?	Gene : Consequence	HBB : Missense Variant
Alleles	T>A / T>C / T>G	Publications	100 citations
Variation Type	SNV Single Nucleotide Variation	Genomic View	See rs on genome
Frequency	A=0.00342 (842/245928, GnomAD) A=0.00438 (532/121340, ExAC) A=0.0116 (360/30968, GnomAD) (+ 1 more)		

Go to dbSNP to view the SNP

[illegible]

Check List

- What is “bioinformatics”?
- Describe a problem in bioinformatics, with relevant data, databases, and algorithms.
- What is Central Dogma?
- What is RNA splicing?
- What is the unit of sequence data?
- How can we perform six-frame translation with a given DNA sequence?