



HUMAN GENOME

Nov 16 2024

Giảng viên: TS. Lưu Phúc Lợi
Luu.p.loi@googlemail.com

Nội Dung Bài Báo Cáo

1. Human Genome Build
2. Introduction to UCSC Genome Browser
3. Introduction to GENCODE

Human Genome Build

Human	hs1	Jan. 2022	T2T Consortium CHM13v2.0	Available
	hg38	Dec. 2013	Genome Reference Consortium GRCh38	Available
	hg19	Feb. 2009	Genome Reference Consortium GRCh37	Available
	hg18	Mar. 2006	NCBI Build 36.1	Available
	hg17	May 2004	NCBI Build 35	Available
	hg16	Jul. 2003	NCBI Build 34	Available
	hg15	Apr. 2003	NCBI Build 33	Archived
	hg13	Nov. 2002	NCBI Build 31	Archived
	hg12	Jun. 2002	NCBI Build 30	Archived
	hg11	Apr. 2002	NCBI Build 29	Archived (data only)
	hg10	Dec. 2001	NCBI Build 28	Archived (data only)
	hg8	Aug. 2001	UCSC-assembled	Archived (data only)
	hg7	Apr. 2001	UCSC-assembled	Archived (data only)
	hg6	Dec. 2000	UCSC-assembled	Archived (data only)
	hg5	Oct. 2000	UCSC-assembled	Archived (data only)
	hg4	Sep. 2000	UCSC-assembled	Archived (data only)
	hg3	Jul. 2000	UCSC-assembled	Archived (data only)
	hg2	Jun. 2000	UCSC-assembled	Archived (data only)
	hg1	May 2000	UCSC-assembled	Archived (data only)

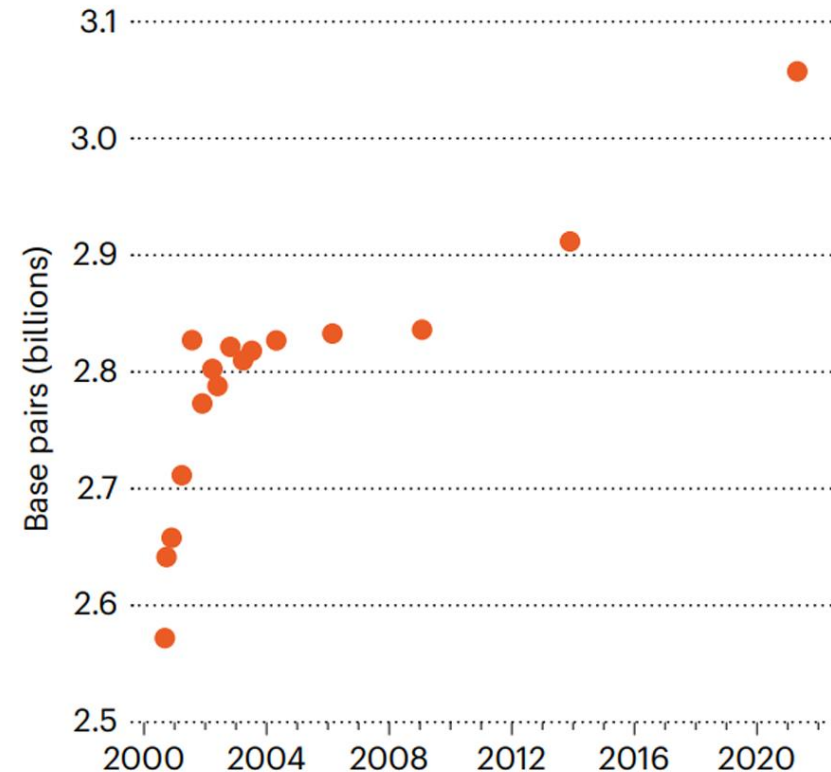
<https://genome.ucsc.edu/FAQ/FAQreleases.html>

A COMPLETE HUMAN GENOME IS CLOSE: HOW THE GAPS WERE FILLED

Researchers added 200 million DNA base pairs and 115 genes – but they've yet to finish the Y chromosome.

COMPLETING THE HUMAN GENOME

Researchers have been filling in incompletely sequenced parts of the human reference genome for 20 years, and have now almost finished it, with 3.05 billion DNA base pairs.



0.3% of sequence might still have errors. Includes X but not Y chromosome. Count excludes mitochondrial DNA.

SOURCE: ADAM PHILLIPPY

Introduction to UCSC Genome Browser



The screenshot shows the UCSC Genome Browser homepage. At the top, the University of California Santa Cruz Genomics Institute logo is on the left, and the UCSC Genome Browser logo is on the right. Below the logos is a navigation bar with links: Genomes, Genome Browser, Tools, Mirrors, Downloads, My Data, Projects, Help, and About Us. A large banner image displays a genomic track with various data visualizations, including a purple curve, a red and yellow heatmap, and a bar chart. Below the banner, there is a search bar with the placeholder text "Search genes, data, help docs and more..." and a "Search" button. To the left of the search bar is a red-bordered box containing the text "Try our new clinical tutorial!" and a magnifying glass icon. Below the search bar, the page is divided into two main sections: "Tools" and "News". The "Tools" section lists several tools with icons: hg38 (a red circle with a white 'h'), hg19 (a red circle with a white 'h'), and mm39 (a red circle with a white 'm'). The tools listed are: Genome Browser (Interactively visualize genomic data), BLAT (Rapidly align sequences to the genome), In-Silico PCR (Rapidly align PCR primer pairs to the genome), Table Browser (Download and filter data from the Genome Browser), LiftOver (Convert genome coordinates between assemblies), REST API (Returns data requested in JSON format), Variant Annotation Integrator (Annotate genomic variants), and More tools... The "News" section lists several news items with dates: Nov. 8, 2024 - New GENCODE gene tracks: Human V47 (hg19/hg38) - Mouse M36; Nov. 4, 2024 - GIAB Problematic Regions tracks for human (hg38 and hs1); Oct. 23, 2024 - New GENCODE gene tracks: V47 (hg38) - VM36 (mm39); Oct. 9, 2024 - CADD v1.7 and ClinGen CSPEC for hg19 and hg38; Oct. 1, 2024 - New clinical tutorial; Sep. 30, 2024 - New gnomAD v4.1 tracks for hg38. Below the news items are buttons for "More news..." and "Subscribe". At the bottom of the page, there is a blue banner with the text "Meetings and Workshops: Come see us in person!". Below this banner, there are two news items: CSHL Genome Informatics 2024 -- Hinxton, UK. Nov 13-15, 2024 and American Society of Human Genetics (ASHG) -- Denver, CO. November. A blue circle with a white question mark is located in the bottom right corner of the page.

UNIVERSITY OF CALIFORNIA
SANTA CRUZ Genomics Institute
UCSC **Genome Browser**

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

Try our new clinical tutorial!

Search genes, data, help docs and more... Search

Tools

-  **Genome Browser** - Interactively visualize genomic data
-  **BLAT** - Rapidly align sequences to the genome
-  **In-Silico PCR** - Rapidly align PCR primer pairs to the genome
-  **Table Browser** - Download and filter data from the Genome Browser
-  **LiftOver** - Convert genome coordinates between assemblies
-  **REST API** - Returns data requested in JSON format
-  **Variant Annotation Integrator** - Annotate genomic variants
- More tools...**

News

- Nov. 8, 2024 - **New GENCODE gene tracks: Human V47 (hg19/hg38) - Mouse M36**
- Nov. 4, 2024 - **GIAB Problematic Regions tracks for human (hg38 and hs1)**
- Oct. 23, 2024 - **New GENCODE gene tracks: V47 (hg38) - VM36 (mm39)**
- Oct. 9, 2024 - **CADD v1.7 and ClinGen CSPEC for hg19 and hg38**
- Oct. 1, 2024 - **New clinical tutorial**
- Sep. 30, 2024 - **New gnomAD v4.1 tracks for hg38**

More news... Subscribe

Meetings and Workshops: Come see us in person!

-  **CSHL Genome Informatics 2024** -- Hinxton, UK. Nov 13-15, 2024
-  **American Society of Human Genetics (ASHG)** -- Denver, CO. November

?

<https://genome.ucsc.edu/index.html>

Getting Help with the UCSC Genome Browser

Online training and tutorials

Our video tutorials address some common questions we've gathered from our [mailing list](#). Along the way we try to show you interesting features of the Browser you may not have found on your own.

Visit our [YouTube channel](#) or use the links below.

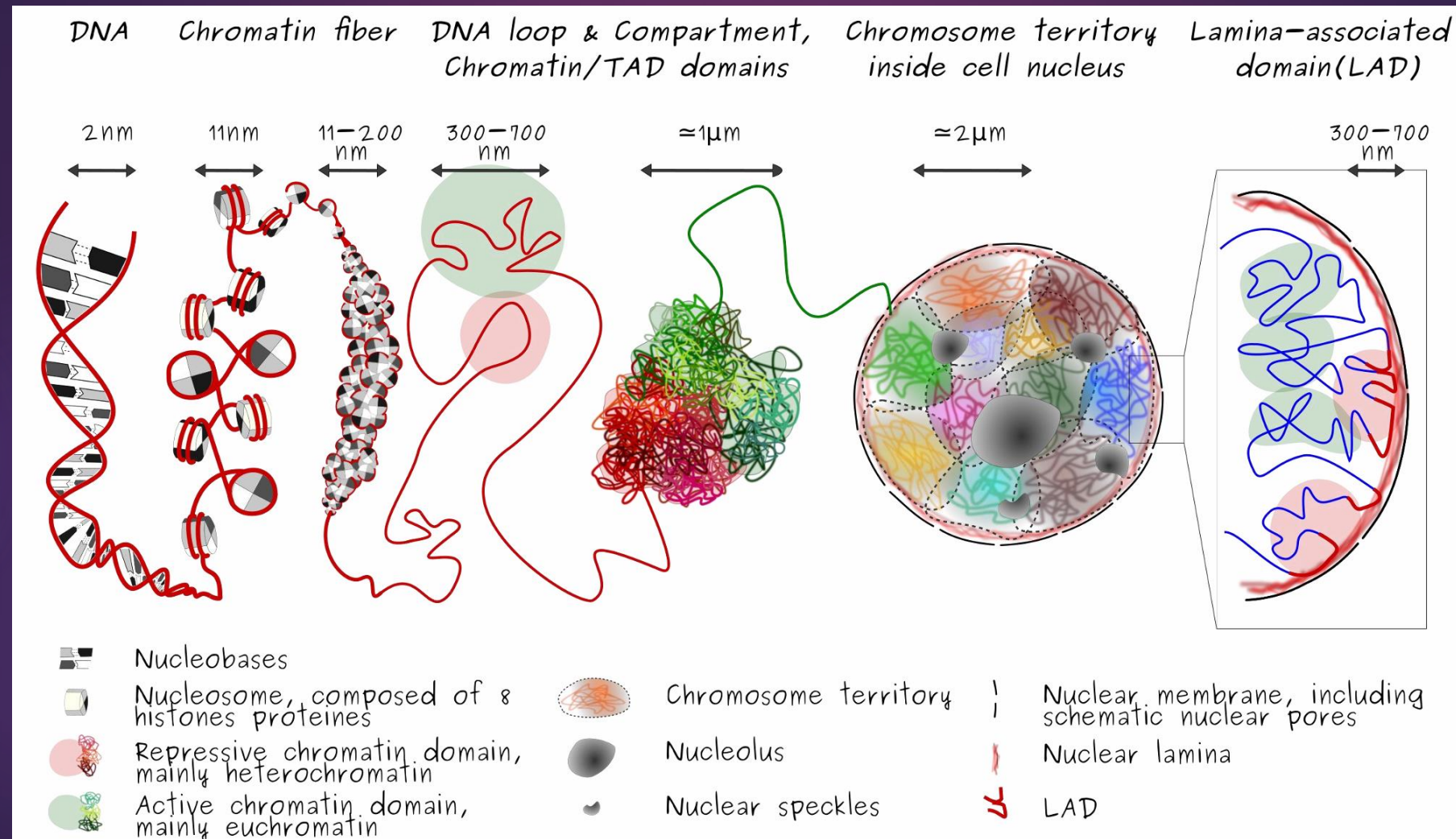
Video tutorials

- Browser Basics, Part One: [Getting around in the Browser](#). [\[transcript\]](#)
- Browser Basics, Part Two: [Configuring the Browser](#). [\[transcript\]](#)
- Browser Basics, Part Three: [Configuration + DNA navigation](#). [\[transcript\]](#)
- Making Links, Part One: [Understanding the URL](#). [\[transcript\]](#)
- Making Links, Part Two: [Jump into genes](#). [\[transcript\]](#)
- Making Links, Part Three: [Composites, custom tracks, spreadsheets](#). [\[transcript\]](#)
- Coronavirus Basics: [Coronavirus Browser SARS-CoV-2](#). [\[transcript\]](#)
- [Saving and sharing sessions](#) in the Browser. [\[transcript\]](#)
- [Controlling visibility of data tracks](#) in the Browser. [\[transcript\]](#)
- Using the isPCR tool ([isPCR](#)) in the UCSC Genome Browser. [\[transcript\]](#)
- [dbSNP resources](#) in the UCSC Genome Browser database. [\[transcript\]](#)
- Using the UCSC Genome Browser [Data Integrator](#). [\[transcript\]](#)

Video tutorials

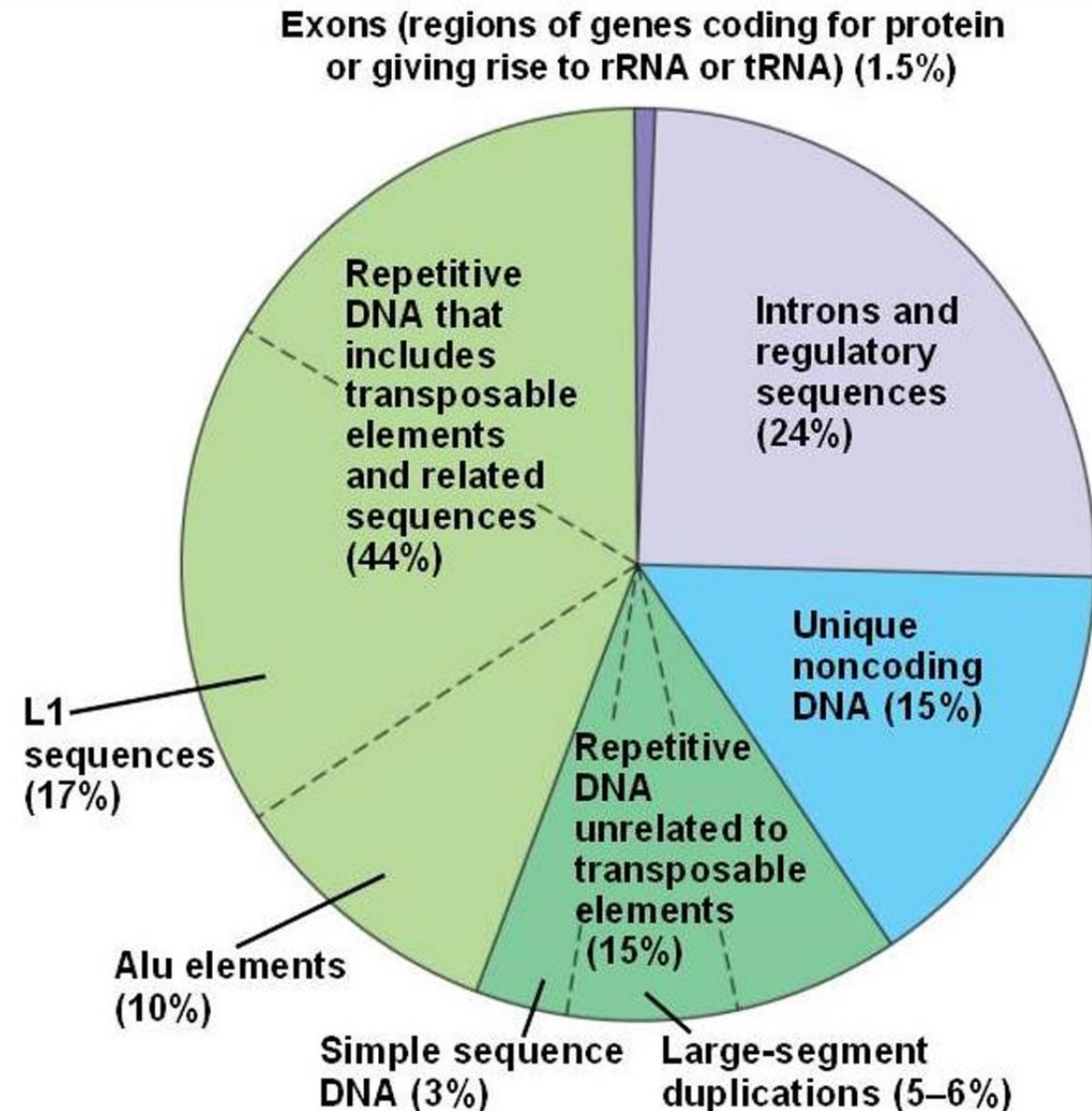
- Finding a [list of genes](#) in a region. [\[transcript\]](#)
- Finding [exon numbers](#). [\[transcript\]](#)
- Finding all [SNPs in a gene](#). [\[transcript\]](#)
- Finding [SNPs upstream](#) from a gene. [\[transcript\]](#)
- Find [which tables belong to a data track](#). [\[transcript\]](#)
- Identifying [codon numbers](#) in a gene. [\[transcript\]](#)
- Obtaining [exon coordinates and sequences](#). [\[transcript\]](#)
- Multi-Region View: [Exon-only display mode](#). [\[transcript\]](#)
- Multi-Region View: [Alternate haplotypes](#). [\[transcript\]](#)
- Multi-Region View: [Discontinuous regions](#). [\[transcript\]](#)
- How-to: Genome Browser [in the Cloud](#).
- How-to: Genome Browser [Gateway](#).

Human Genome Organization



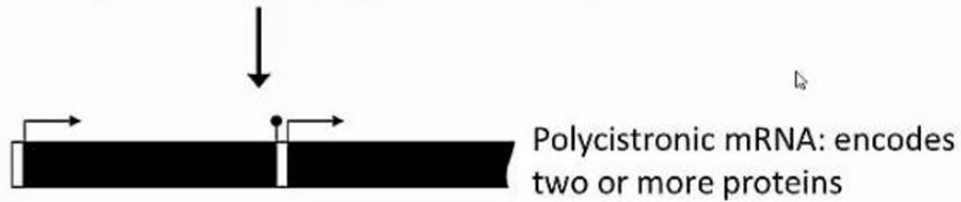
Genome components

- The human genome contains less than 2% coding exons within genes.
- The remaining DNA consist of:
 - introns and regulatory sequences such as enhancers
 - unique noncoding DNA contains many pseudogenes (genes that have accumulated mutations and became nonfunctional)
 - repetitive DNA are sequences that are repeated many times; much of these are
 - transposable elements that can "jump" between chromosomes, leading to transpositions.

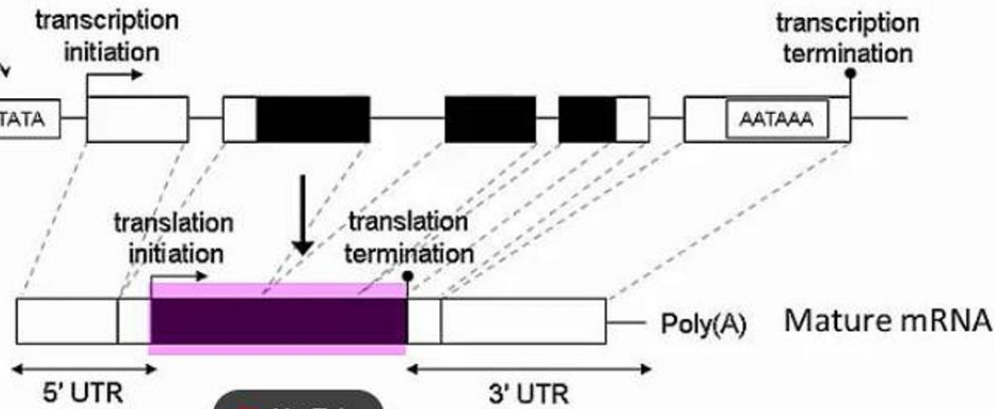


Gene structure

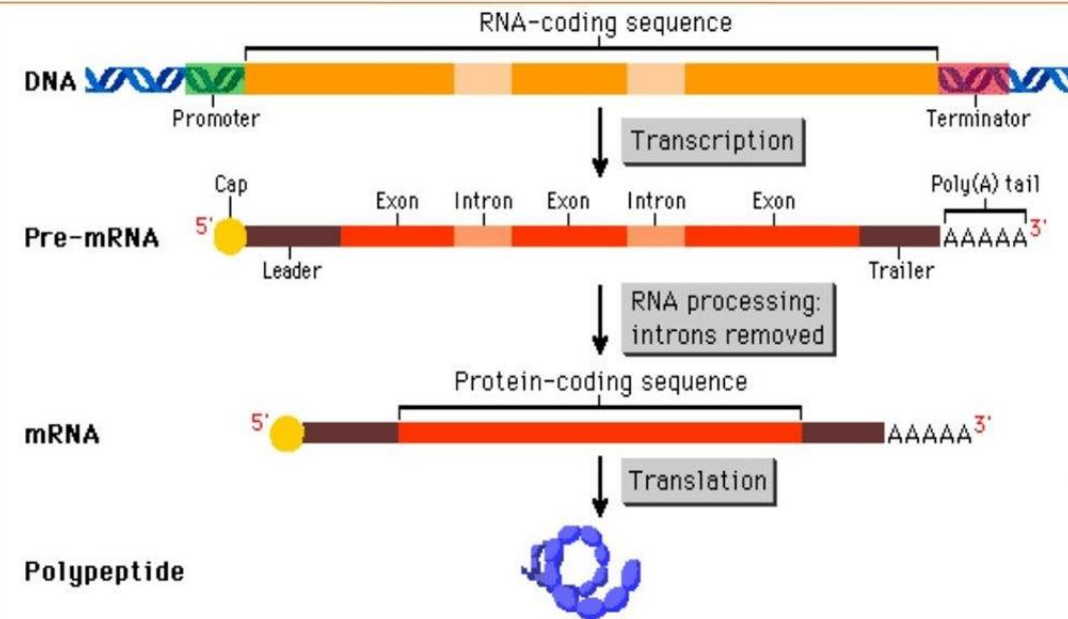
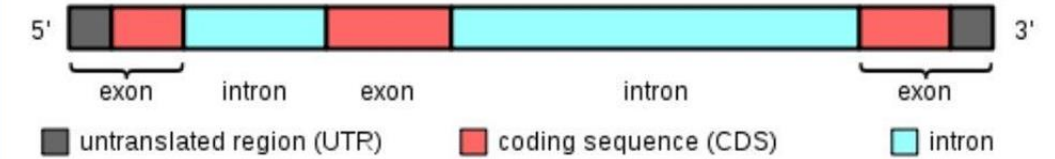
Prokaryotes:



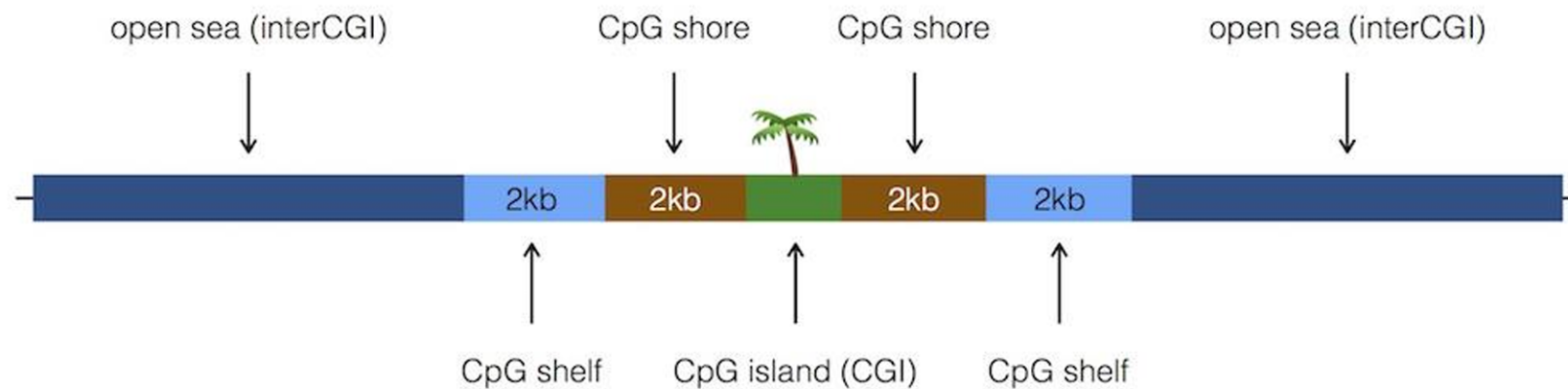
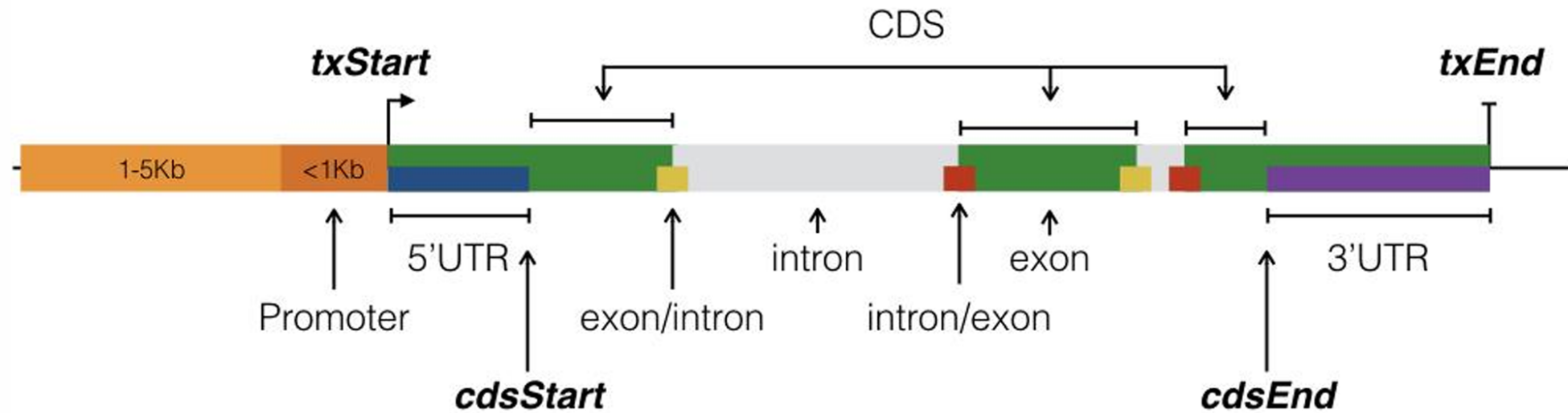
Eukaryotes:



Eukaryotic gene.




Gene structure



Understanding Gene Annotation through GENCODE


(<https://www.gencodegenes.org/>)



Human Mouse How to access data FAQ Documentation About us


HUMAN

GENCODE 47 (October 2024)



MOUSE

GENCODE M36 (October 2024)



The goal of the GENCODE project is to identify and classify all gene features in the human and mouse genomes with high accuracy based on biological evidence, and to release these annotations for the benefit of biomedical research and genome interpretation.

The GENCODE human and mouse lncRNA annotations are significantly expanding as we integrate models from our [Capture Long-read Sequencing project](#).



Xin chân thành cảm ơn!

LUU PHUC LOI, PHD

ZALO: 0901802182

LUU.P.LOI@GOOGLEMAIL.COM