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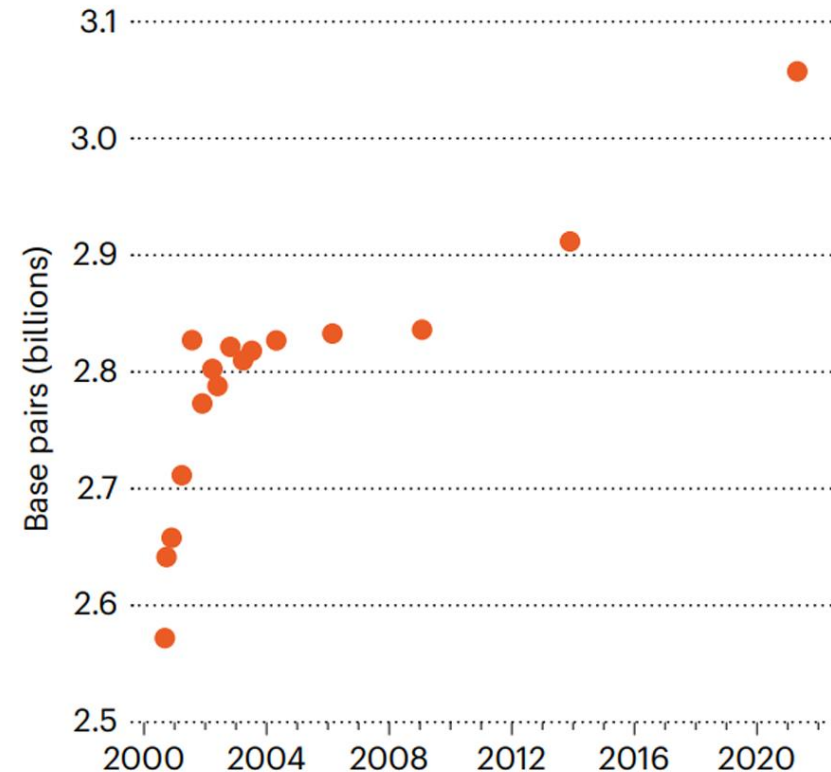
0.3% of sequence might still have errors. Includes X but not Y chromosome. Count excludes mitochondrial DNA.

Introduction to UCSC Genome Browser

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

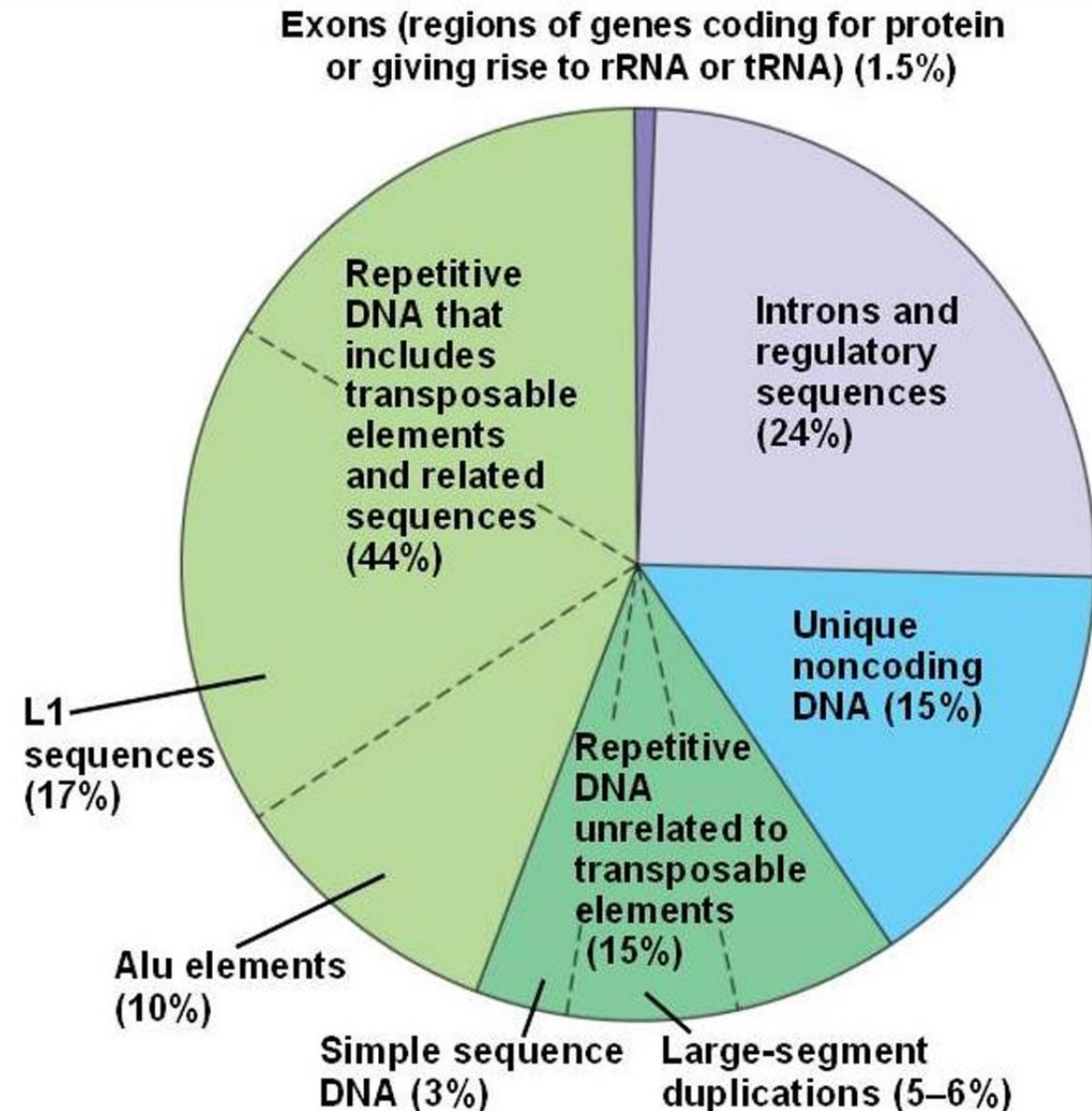- Finding a list of genes in a region. [transcript]
- Finding exon numbers. [transcript]
- Finding all SNPs in a gene. [transcript]
- Finding SNPs upstream from a gene. [transcript]
- Find which tables belong to a data track. [transcript]
- Identifying codon numbers in a gene. [transcript]
- Obtaining 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.

https://genome.ucsc.edu/training/
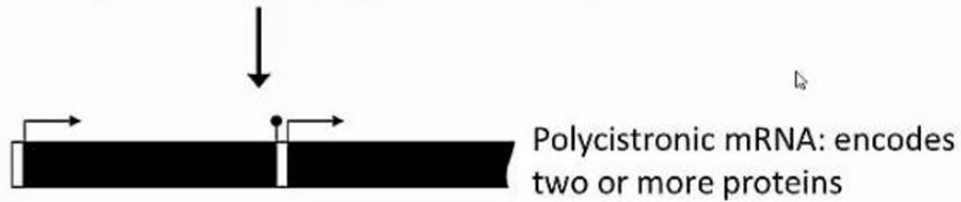
Human Genome Organization
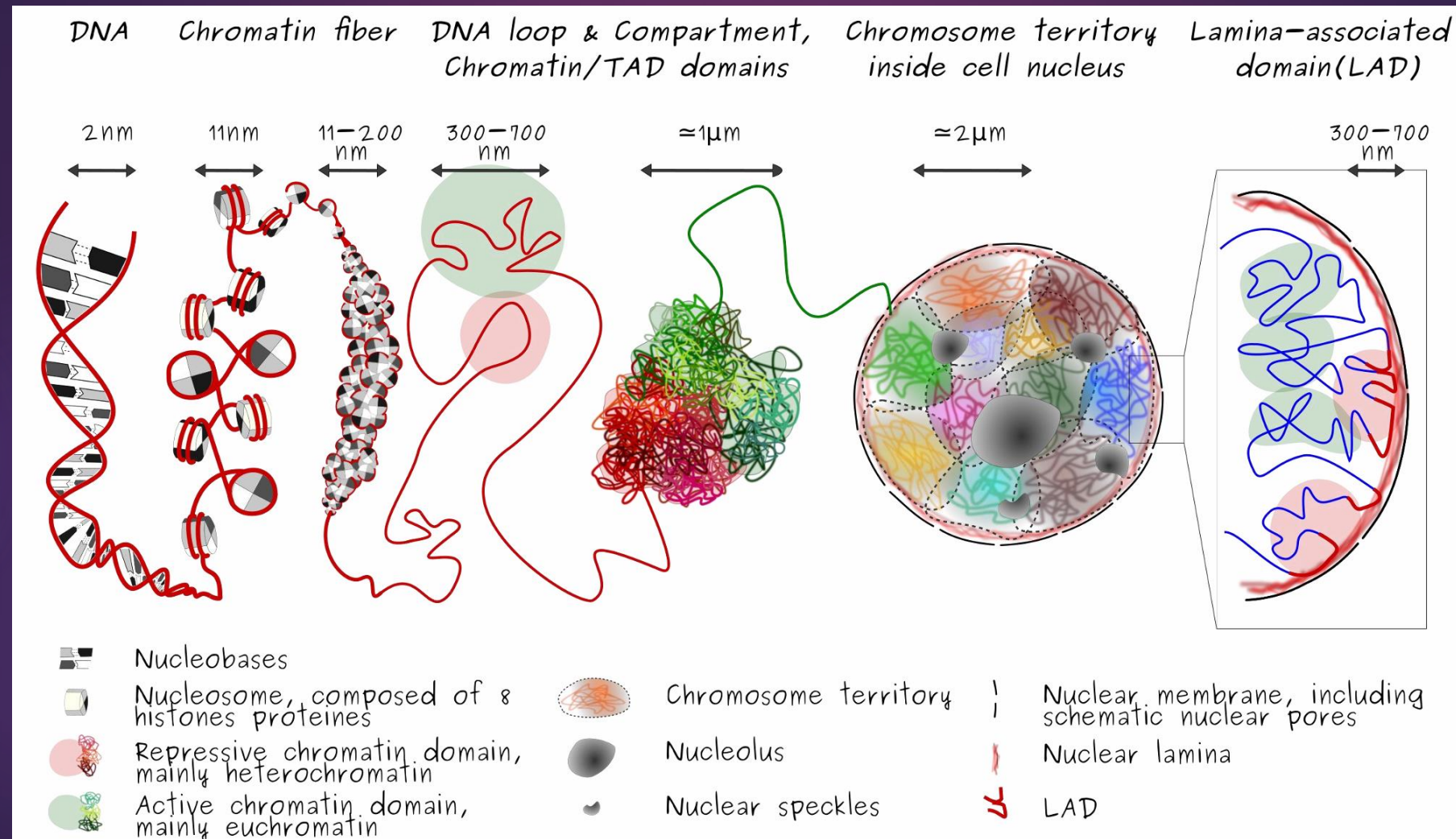
Genome components

- The human genome contains less than 2% coding exons within genes.

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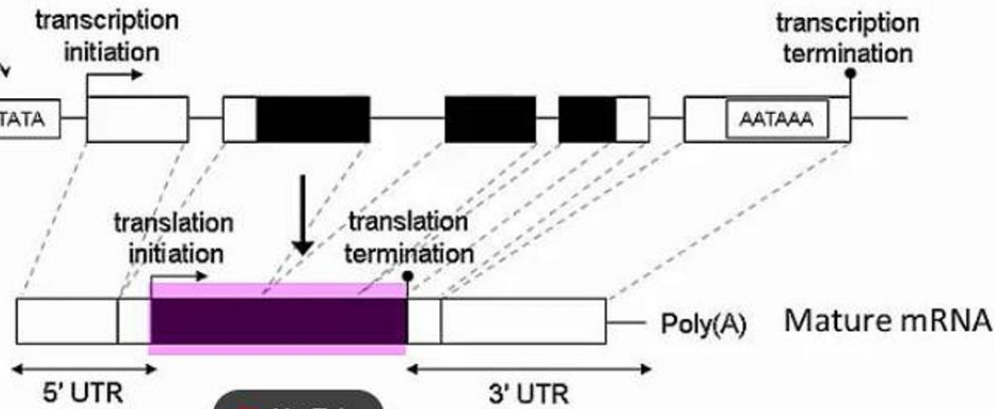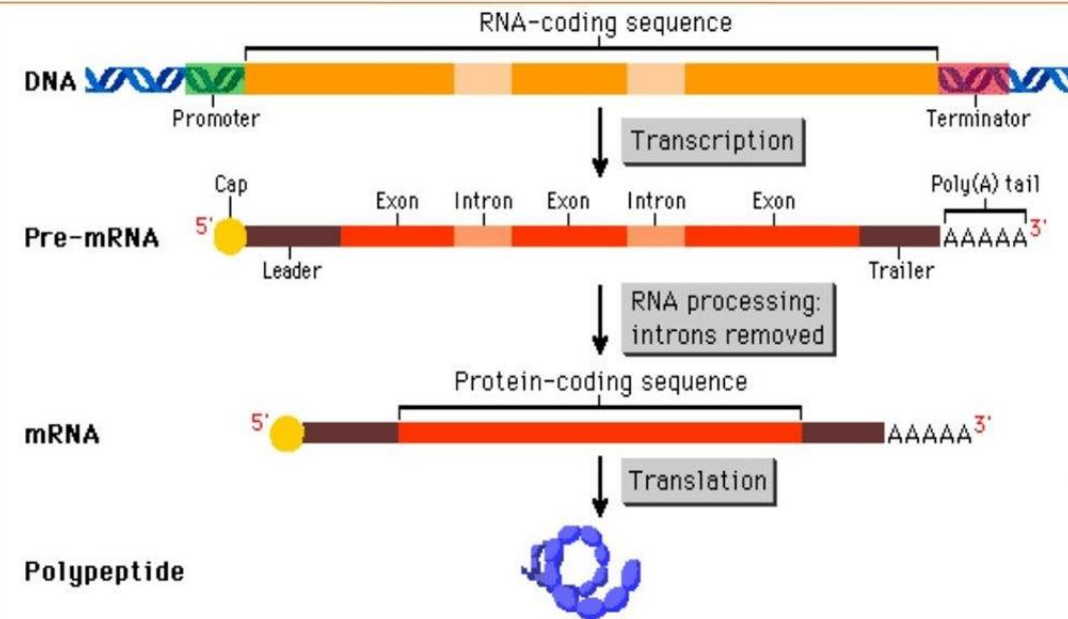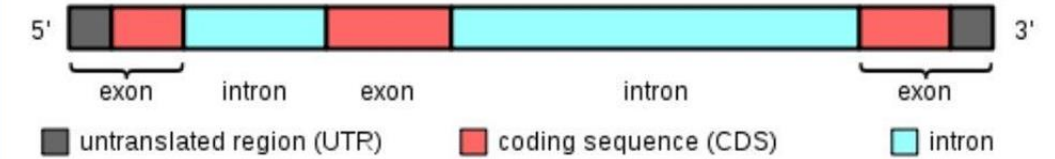 - o transposable elements that can "jump" between chromosomes, leading to transpositions.

Exons (regions of genes coding for protein or giving rise to rRNA or tRNA) (1.5%)

Repetitive DNA that includes transposable elements and related sequences (44%)

Introns and regulatory sequences (24%)

Unique noncoding DNA (15%)

L1 sequences (17%)

Repetitive DNA unrelated to transposable elements (15%)

Alu elements (10%)

Simple sequence DNA (3%)

Large-segment duplications (5–6%)

Gene structure

Gene structure

https://www.bioconductor.org/packages/devel/bioc/vignettes/annotatr/inst/doc/annotatr-vignette.html
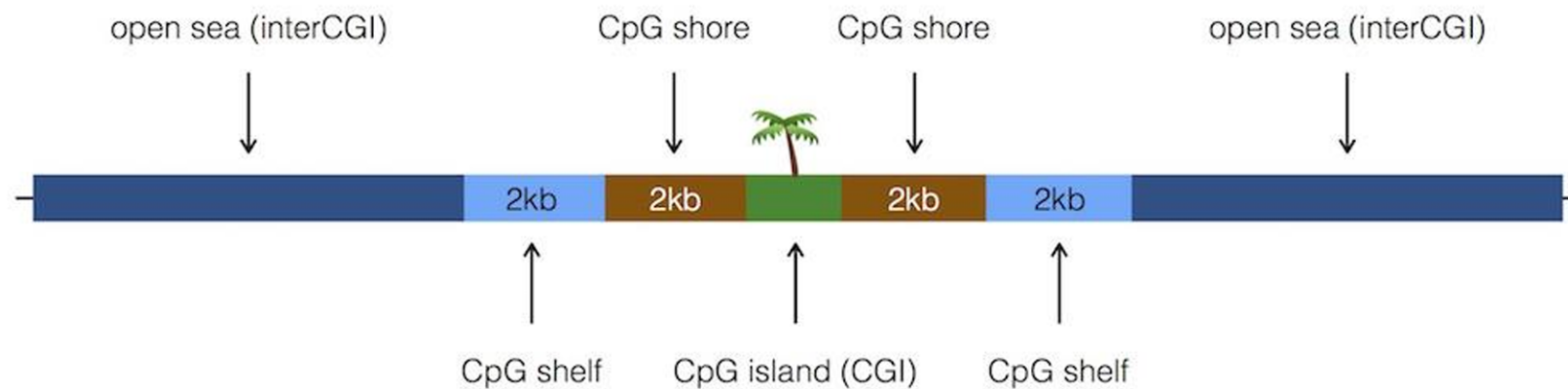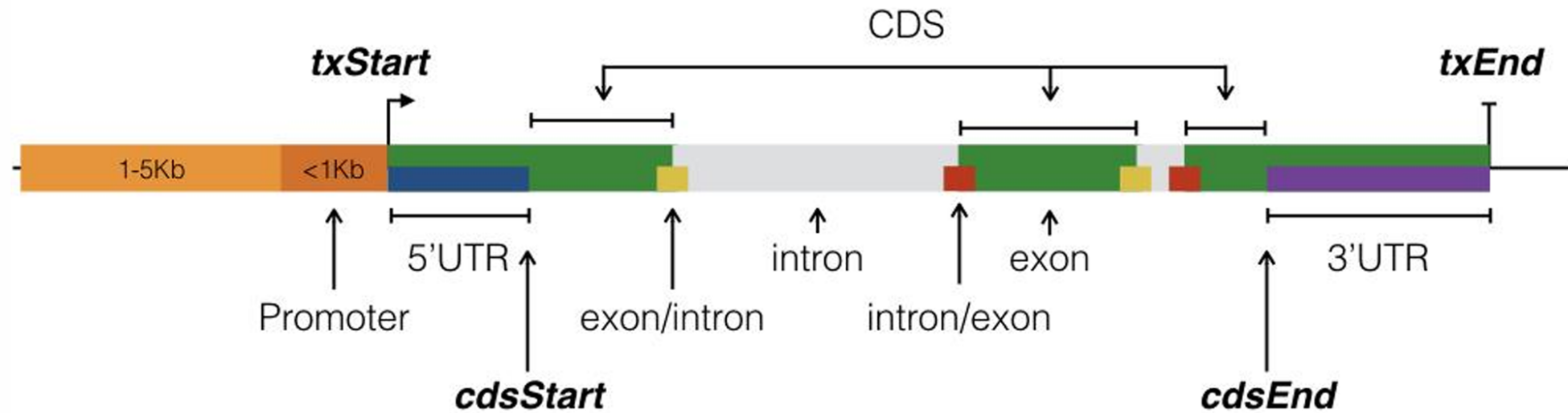
Understanding Gene Annotation through GENCODE (https://www.gencodegenes.org/)

Xin chân thành cảm ơn!

LUU PHUC LOI, PHD

ZALO: 0901802182

LUU.P.LOI@GOOGLEMAIL.COM