

~~MTA0034 - BIOTECNOLOGIA COMPUTACIONAL (2026.2 - T01)~~

Bioinformática aplicada à medicina

ALLYSSON ALLAN

PROFESSOR VISITANTE

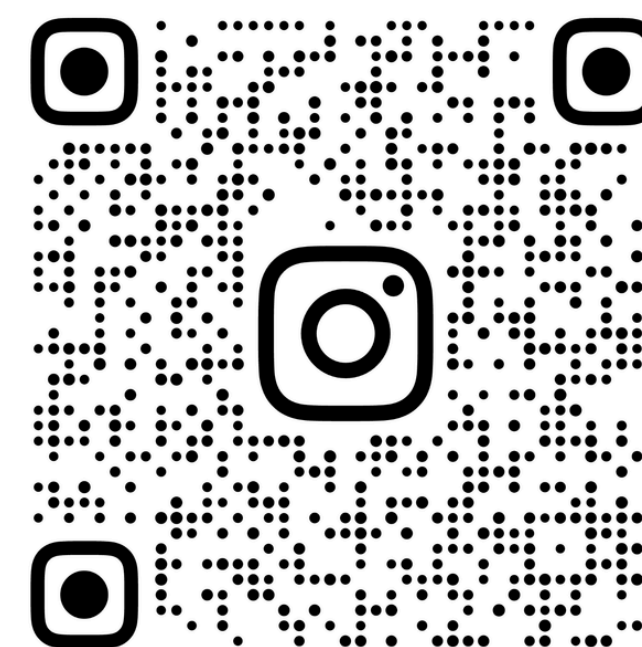
PPG EM MEDICINA TRANSLACIONAL

NPDM / UFC

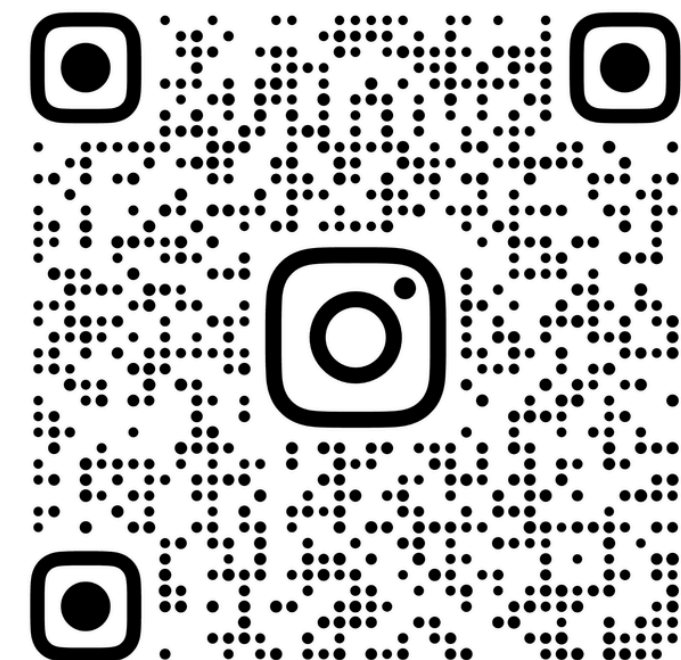
ALLYSSON@UFC.BR

FORTALEZA

AGOSTO 2026



LABBAT.NPDM.UFC



PPGMTUFC

TABELA DE HORÁRIOS

8H - APRESENTAÇÃO DA DISCIPLINA;

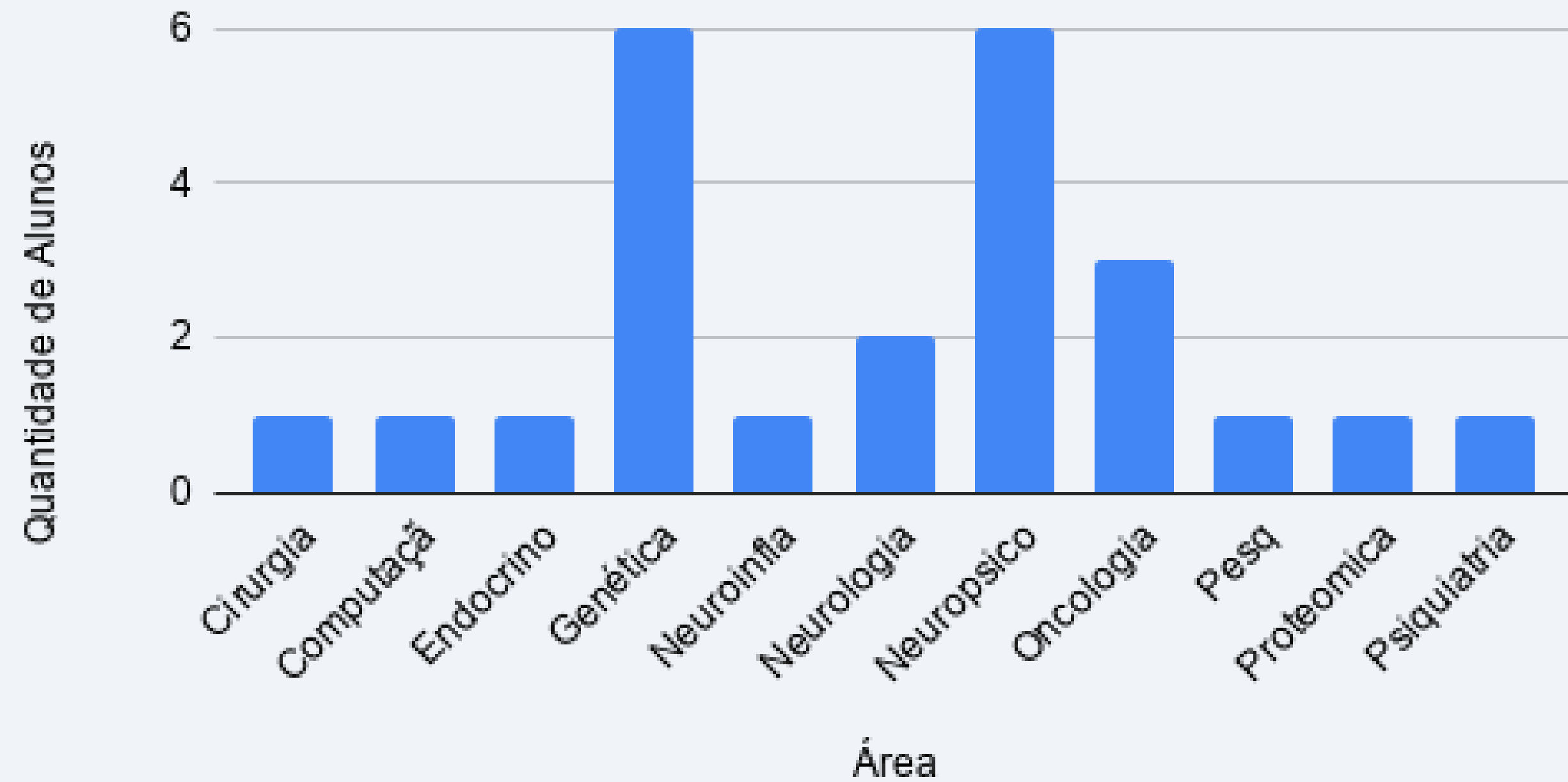
9H - ORGANIZAÇÃO DO GENOMA HUMANO;

10H - EXPLORANDO O GENOME BROWSER;

10H30 - EXERCÍCIO 1;

11H50 - CONSIDERAÇÕES FINAIS.

Distribuição de Alunos por Área



OBJETIVOS

- COMPREENDER E APLICAR CONCEITOS, MÉTODOS E FERRAMENTAS DE BIOINFORMÁTICA NA ANÁLISE DE DADOS GENÔMICOS, TRANSCRIPTÔMICOS E CLÍNICOS;
- UTILIZAR A LINGUAGEM R, BANCOS DE DADOS BIOMÉDICOS E FLUXOS COMPUTACIONAIS REPRODUTÍVEIS PARA INVESTIGAR DOENÇAS RARAS, CÂNCER E OUTRAS APLICAÇÕES DA MEDICINA DE PRECISÃO;
- INTERPRETAR CRITICAMENTE OS RESULTADOS E ORGANIZAR DADOS ÔMICOS SEGUNDO OS PRINCÍPIOS FAIR (*FINDABLE, ACCESSIBLE, INTEROPERABLE, REUSABLE*).

CRONOGRAMA 2026.2

24/08 - CONCEITOS BÁSICOS DE GENÉTICA;

26/08 - CONCEITOS BÁSICOS DE PROGRAMAÇÃO EM R;

31/08 - BIOINFORMÁTICA EM R;

02/09 - GENÔMICA E TRANSCRIPTÔMICA;

09/09 - BANCOS DE DADOS DE ÔMICAS E CLÍNICOS;

14/09 - ONCOLOGIA DE PRECISÃO;

16/09 - DOENÇAS RARAS;

21/09 - MEDICINA DE PRECISÃO E IMPLEMENTAÇÃO CLÍNICA.

AVALIAÇÃO

- PARTICIPAÇÃO NOS EXERCÍCIOS EM SALA;
- TRABALHO FINAL COM OS DADOS ÔMICOS PARA UMA ANÁLISE ESTATÍSTICA COM TESTE DE HIPÓTESES E GERAÇÃO DE GRÁFICOS NO R. OS CRITÉRIOS SERÃO BASEADOS NOS PRINCÍPIOS FAIR (*FINDABLE, ACCESSIBLE, INTEROPERABLE, REUSABLE*).



RECOMENDAÇÕES INICIAIS

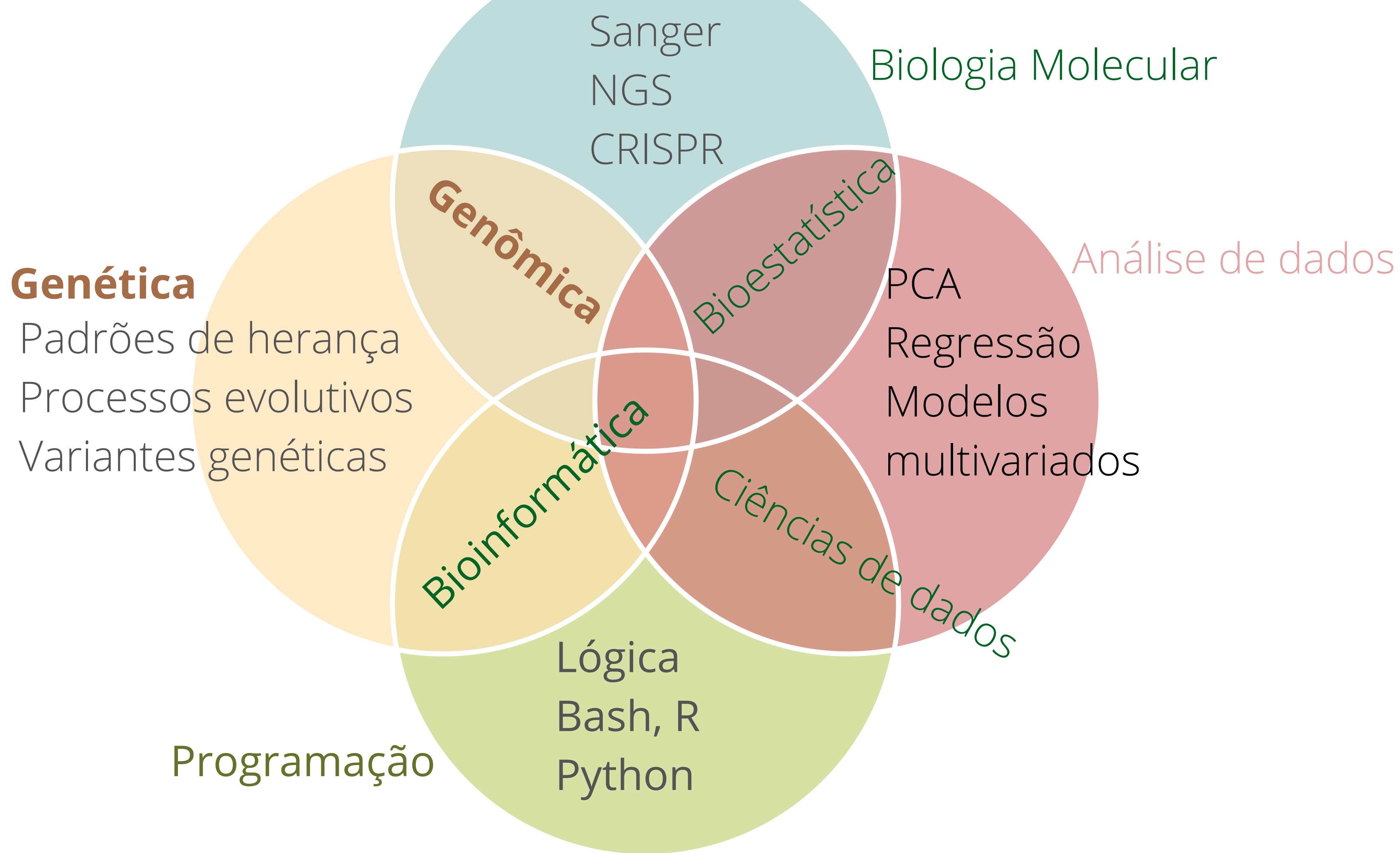
- TRAZER UM NOTEBOOK PESSOAL OU DO LABORATÓRIO PARA AS AULAS;
- OU CASO NÃO TENHA NOTEBOOK, UTILIZAR O GOOGLE COLAB PELO CELULAR OU TABLET (CONFIGURAÇÃO POR SUA CONTA);
- OU AINDA, UTILIZAR OS CRÉDITOS GRATUITOS DE PROVEDORES DE PROCESSAMENTO EM NUVEM - AWS, GCP, AZURE, DENTRE OUTROS (CONFIGURAÇÃO DAS MÁQUINA VIRTUAL POR SUA CONTA).

Perguntas de partida:

Por que bioinformática?

O que é bioinformática?

Qual a diferença de bioinformática para biologia
computacional?



AULA 1 – CONCEITOS BÁSICOS DE GENÉTICA (ORGANIZAÇÃO DO GENOMA HUMANO) – 24/08

- PRITCHARD, J. K. "AN OWNER'S GUIDE TO THE HUMAN GENOME: AN INTRODUCTION TO HUMAN POPULATION GENETICS, VARIATION AND DISEASE." (2023);
- STRACHAN T, READ AP – HUMAN MOLECULAR GENETICS, 5TH ED. CRC PRESS, TAYLOR & FRANCIS, BOCA RATÓN, FL, 2019.



1.1 Some Very Useful Numbers for human population genetics.

When you're thinking about genomes and genomic data it's often useful to have a sense of rates and scales¹.

Genome properties.

Genome size: 3.1 Gb (haploid size).

Number of chromosomes: 23 pairs

Number of coding genes: ~20,000

Exons per gene: 8 (median)

Number of genes per megabase: 6.5 (mean)

Total in protein-coding exons: 1% of genome

Total in genes (introns+exons): 40% of genome

Active chromatin (per cell type): 1% of genome

Active chromatin (all cell types): 13% of genome

Length scales. (Orders of magnitude.)

Transcription factor binding site: 10 bp

Enhancer: 100 bp – 1 Kb

Exon (coding): 150 bp

Coding length per gene: 1200 bp (median)

Intron: 1 – 50 Kb

Gene (pre-mRNA): 10 – 100 Kb

Extent of LD: 10 Kb – 1 Mb (varies by locus & population)

Enhancer–promoter interactions: 1 Kb – 1 Mb

Chromatin topological domains (TADs): ~1 Mb

Chromosome lengths: 47 Mb – 250 Mb

Genetic variation.

Heterozygosity: $0.5\text{--}1.0 \times 10^{-3}$ (varies by population)

Human-Chimpanzee divergence: 1.4×10^{-2}

Number of common SNPs: ~ 8 million (at $> 5\%$ MAF in global sample)

Number of SNPs for genome-wide SNP tagging: ~ 1 million

Fst between populations: $\sim 0.10\text{--}0.15$ between continents

Population genetic parameters.

Mutation rate per generation: 1.3×10^{-8} per basepair (at parental age 30)

Mutation rate per year: 4.0×10^{-10} per basepair

Number of mutations per child: ~ 70

Recombination rate: 1.2 centiMorgans per Mb (mean, sex-averaged)

Cross-overs per egg: ~ 42

Cross-overs per sperm: ~ 26

Effective population size: 10,000–20,000 ($H/4\mu$)

Timescales of population divergence. (Take with grain of salt)

Human to chimpanzee: 6.5 MY

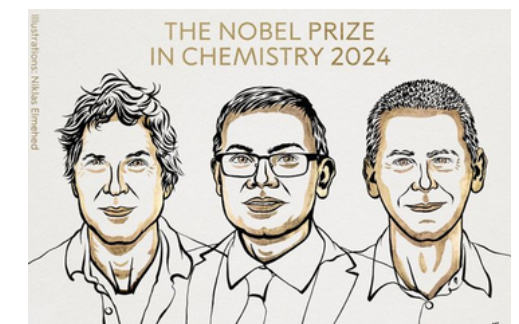
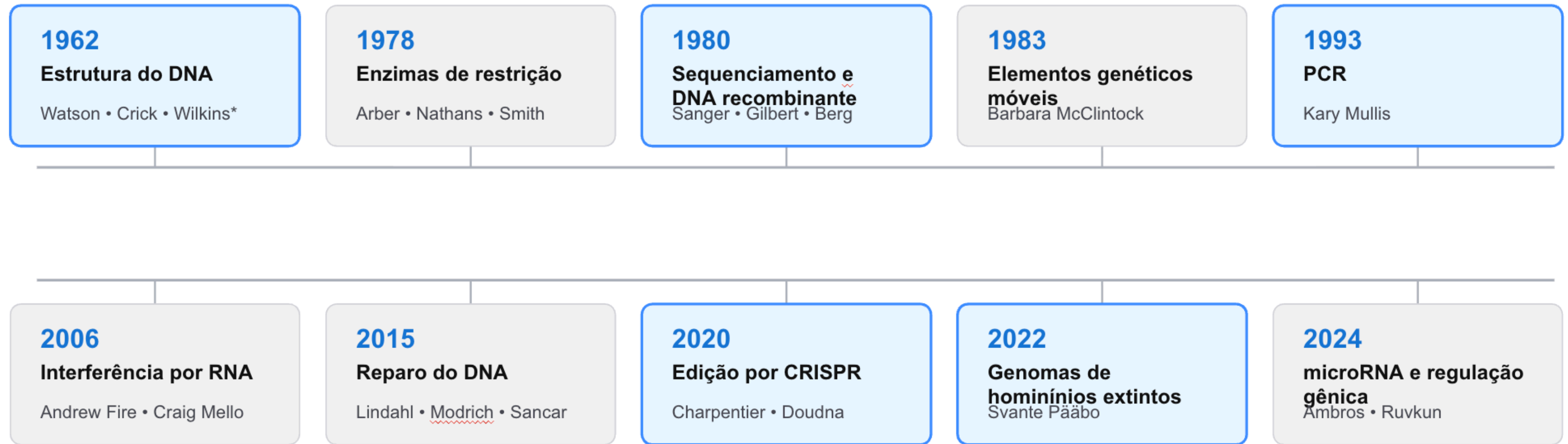
Human to Neanderthal/Denisovan: 600 KY

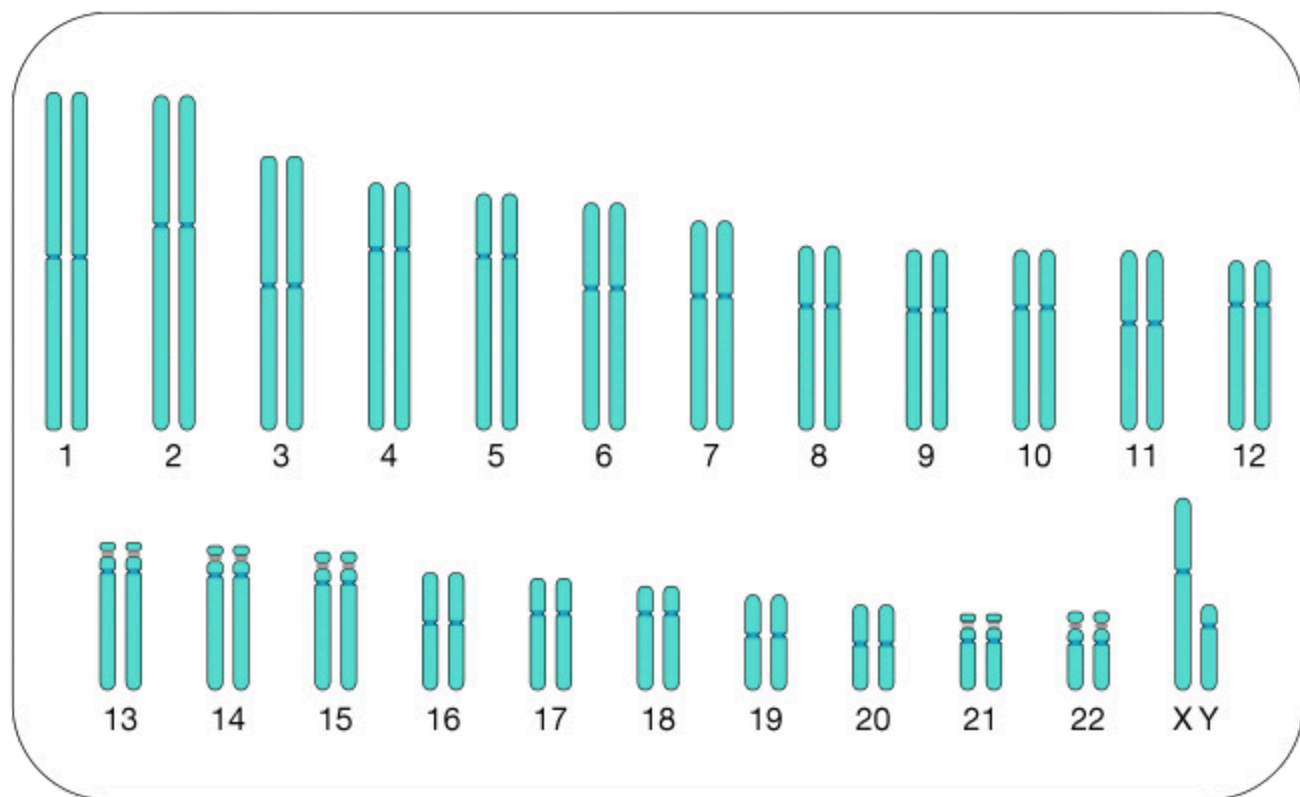
Deepest human population splits: ~ 200 KY

Out-of-Africa migration: 65–100 KY

Deepest non-African splits: 65 KY

Descobertas laureadas após o descobrimento da estrutura do DNA por **Rosalind Franklin**

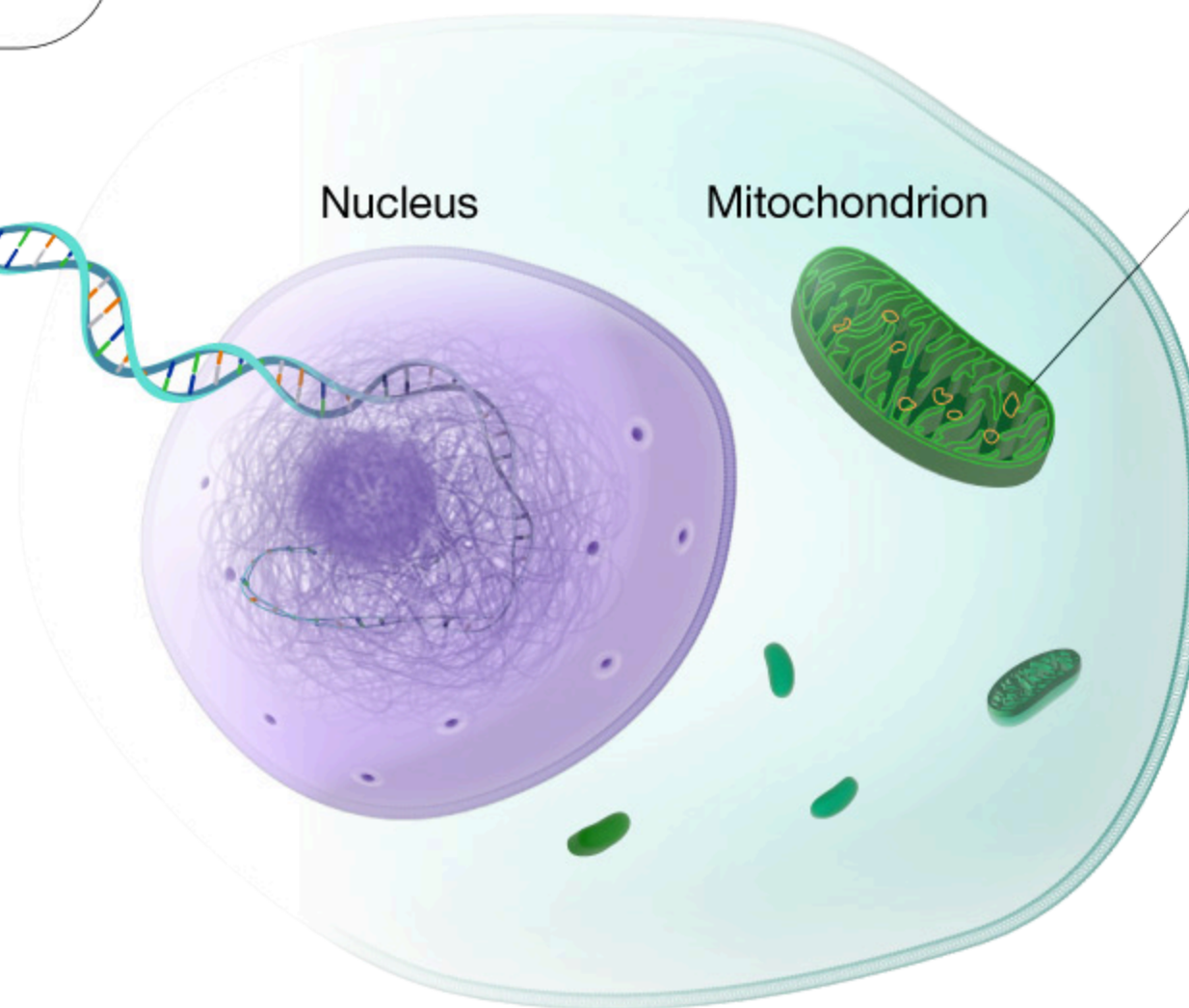




Shown here are human chromosomes

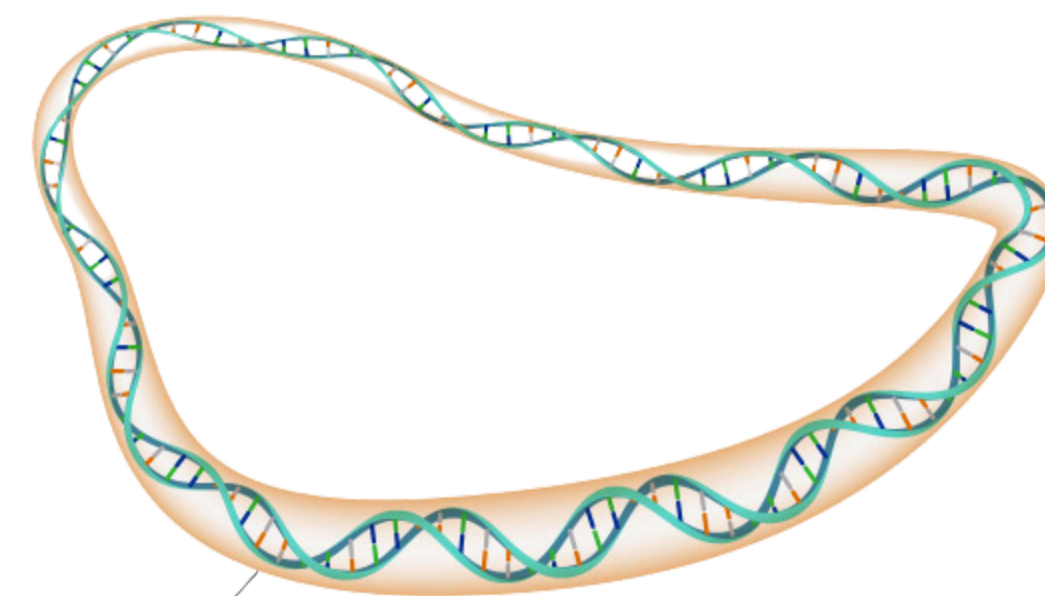


Nuclear DNA



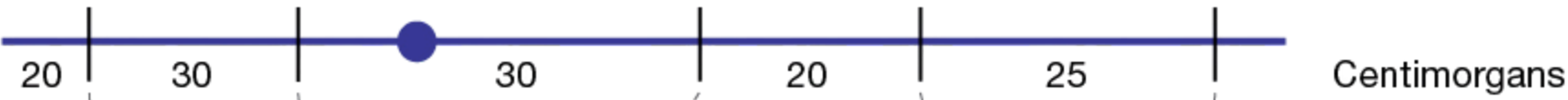
Nucleus

Mitochondrion

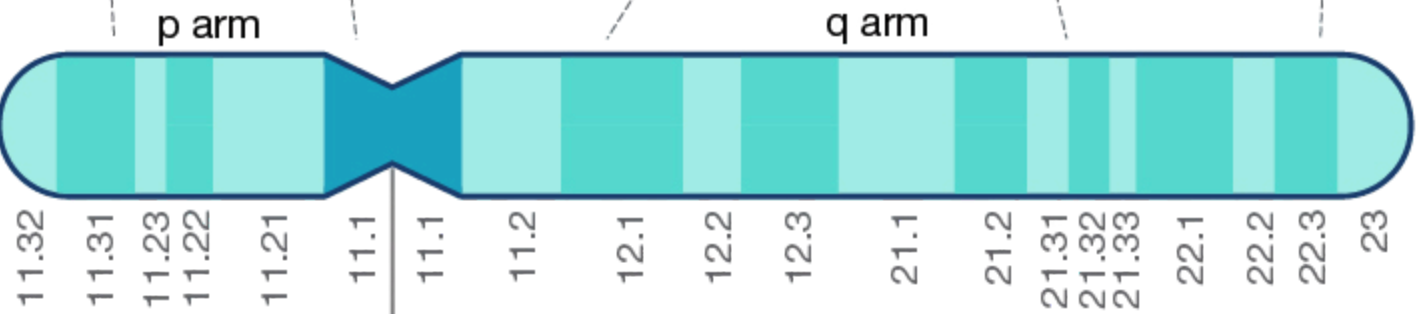


Mitochondrial DNA

Genetic map



Cytogenetic map



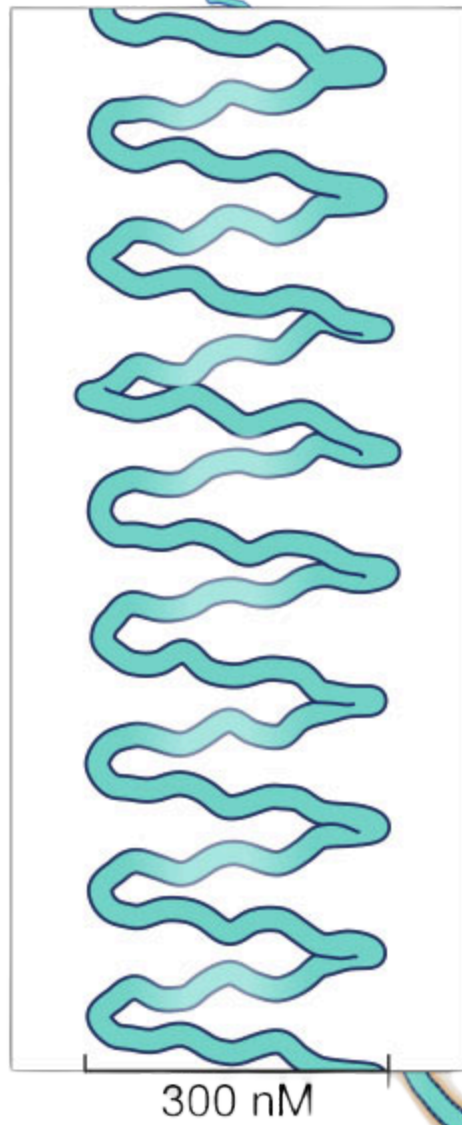
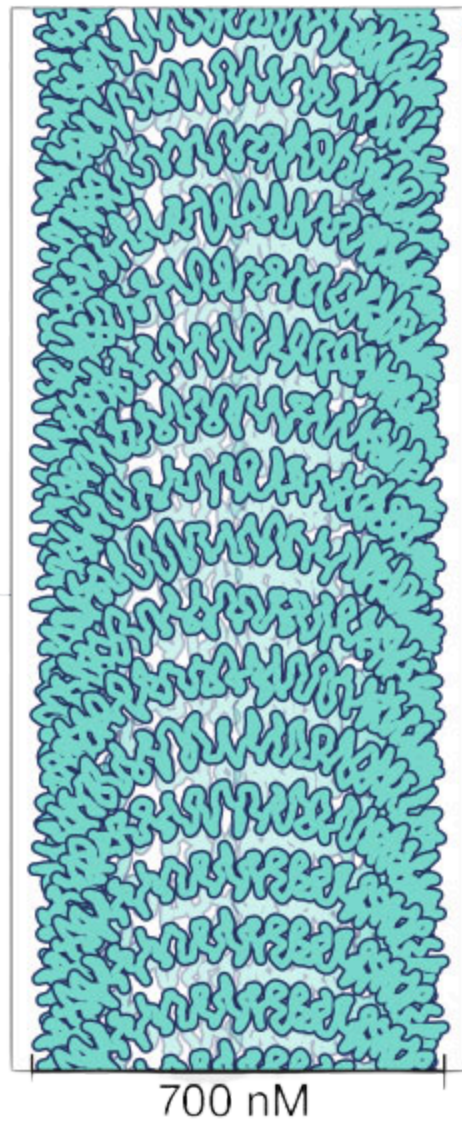
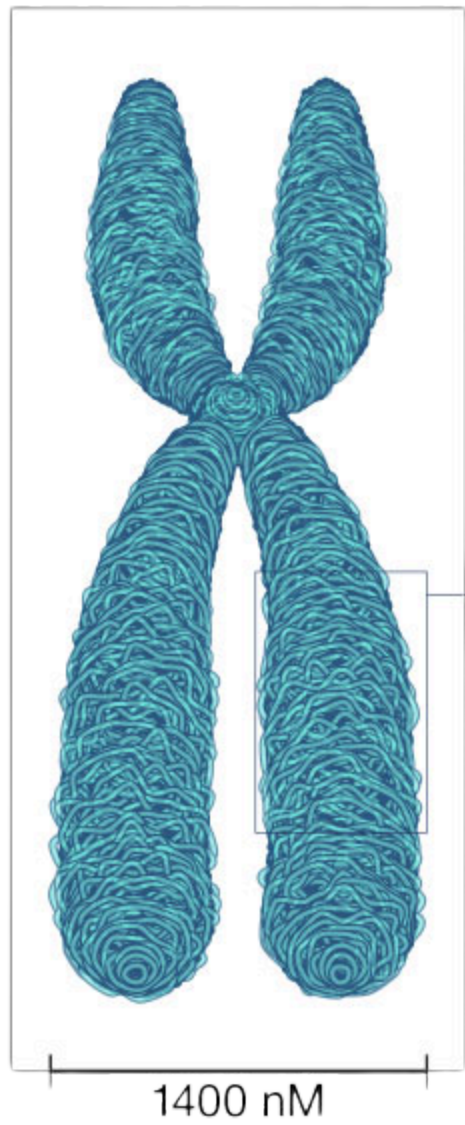
Physical map



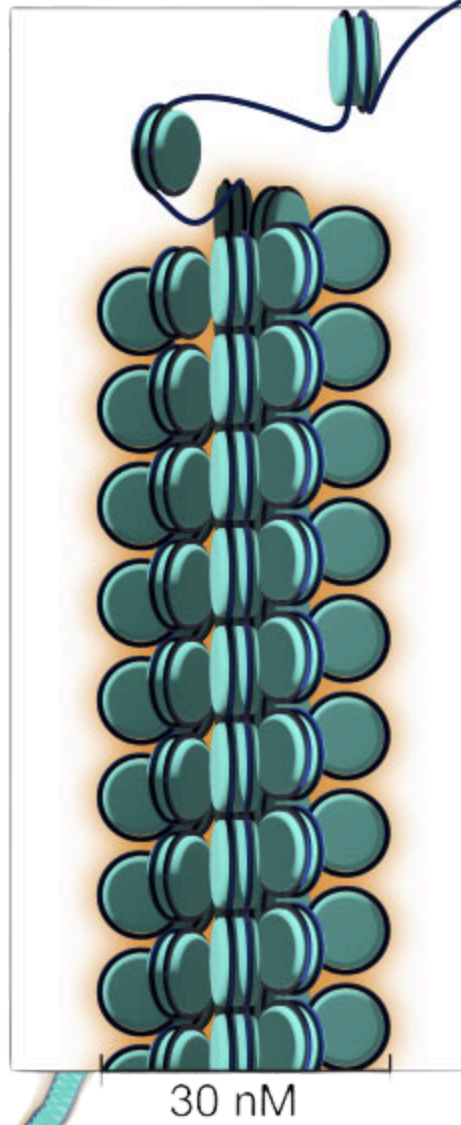
DNA sequence

TCGATCGATCGCTATCAGTAGCATGCATGCATGCATGC

Metaphase chromosome



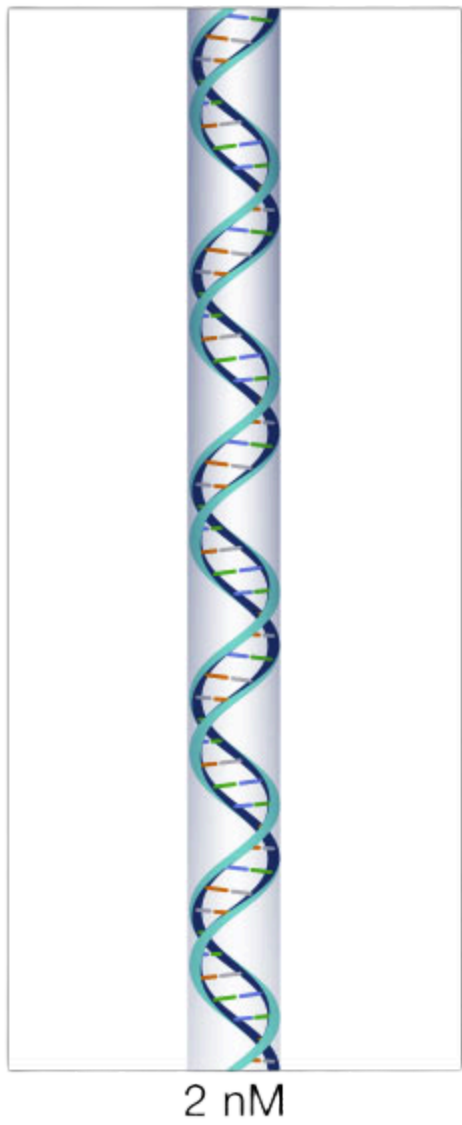
Chromatin fiber



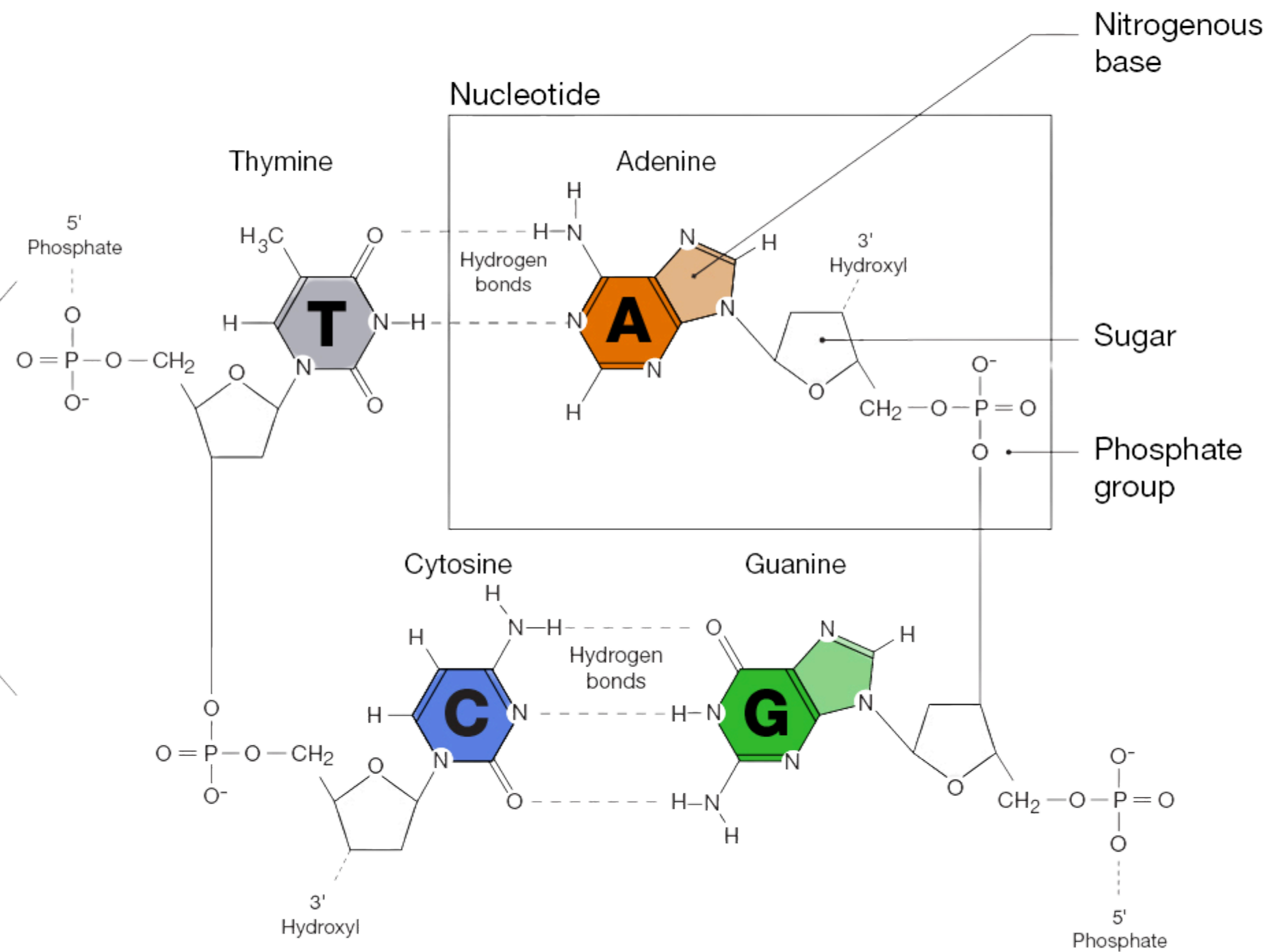
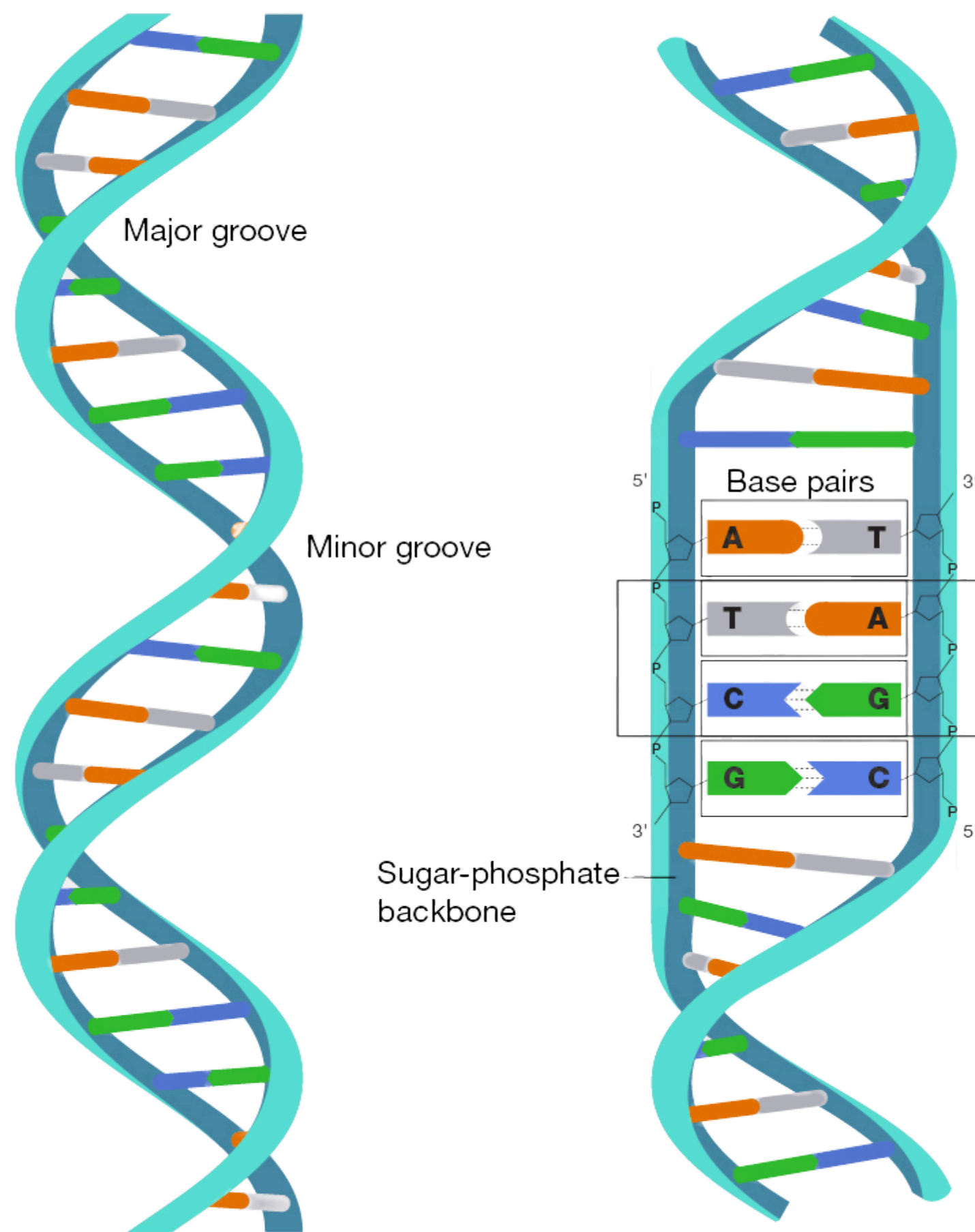
Nucleosomes

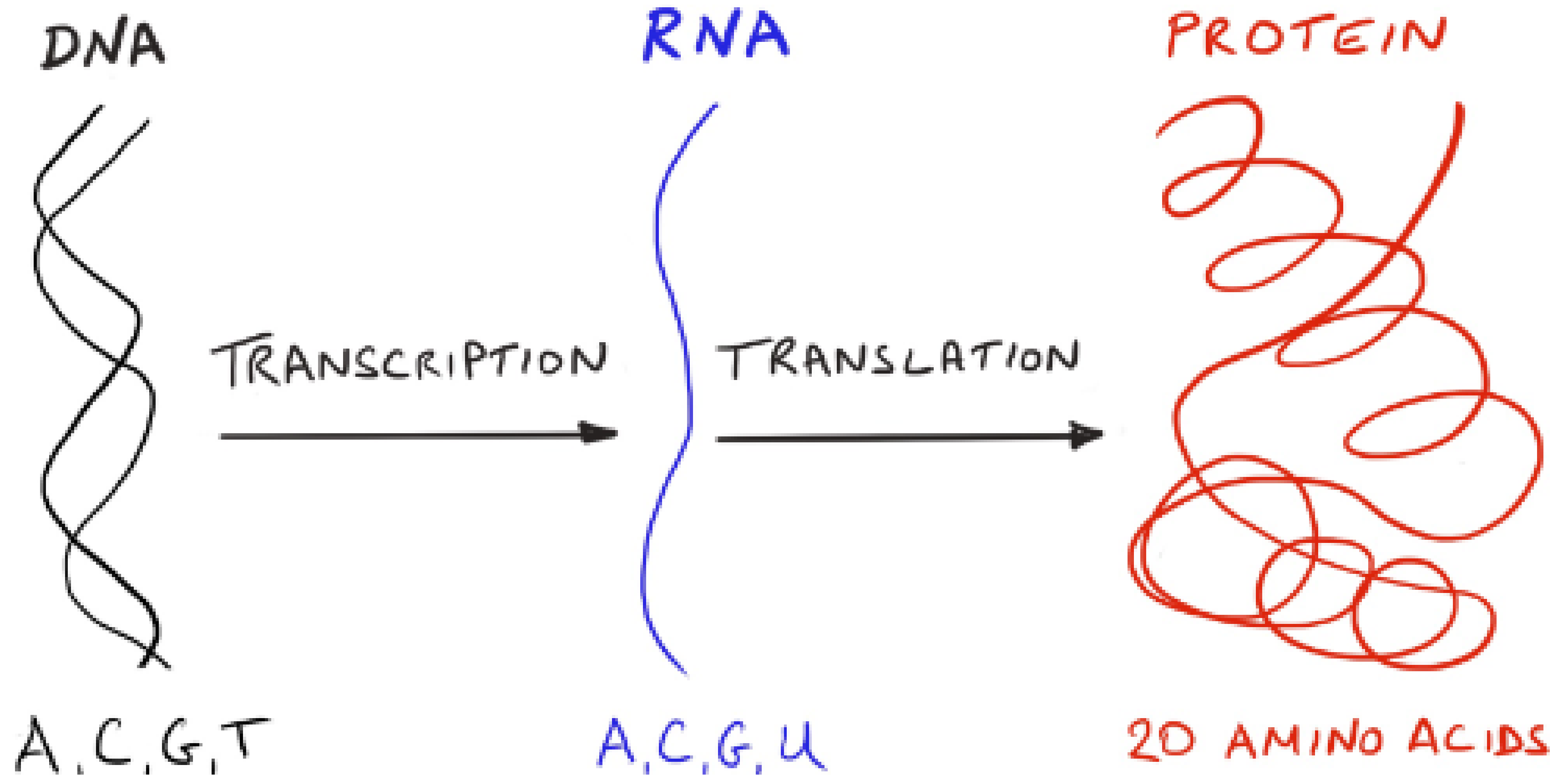


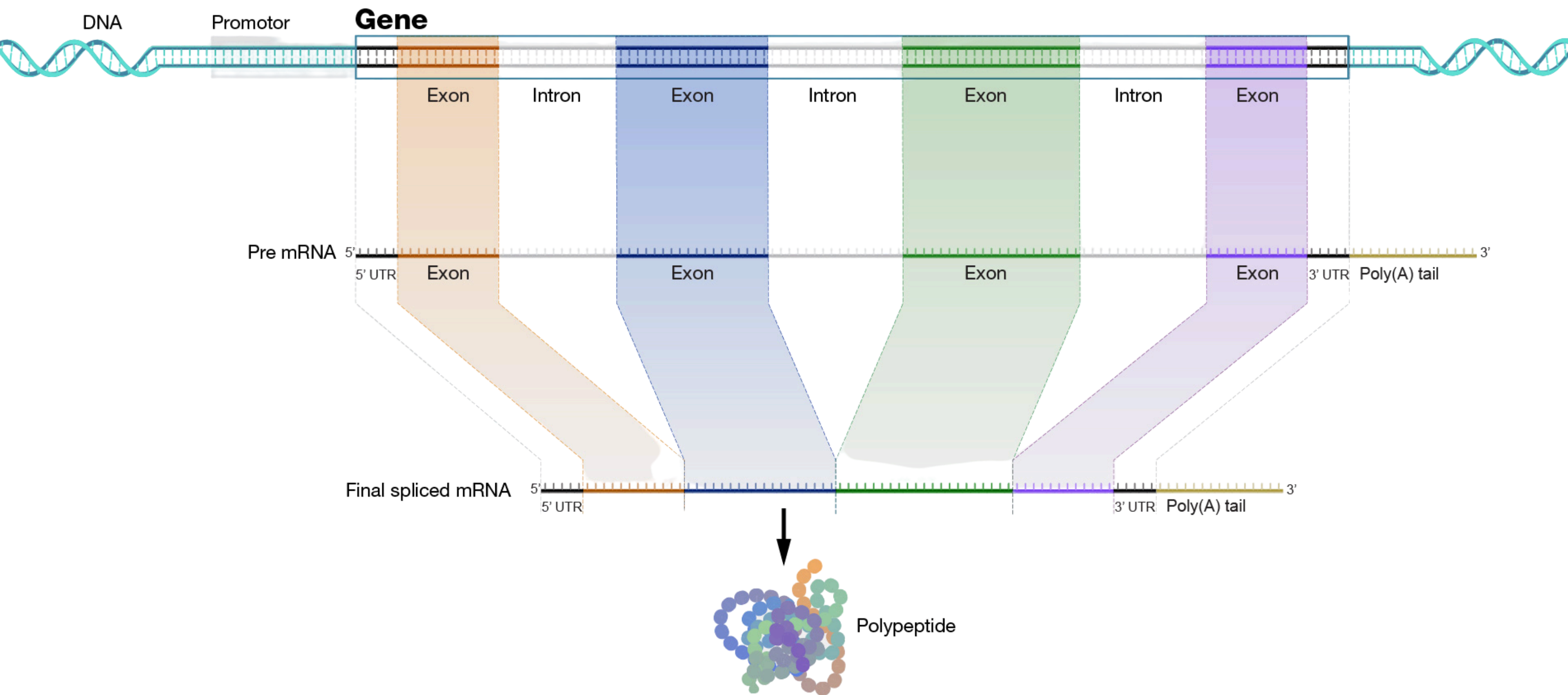
DNA

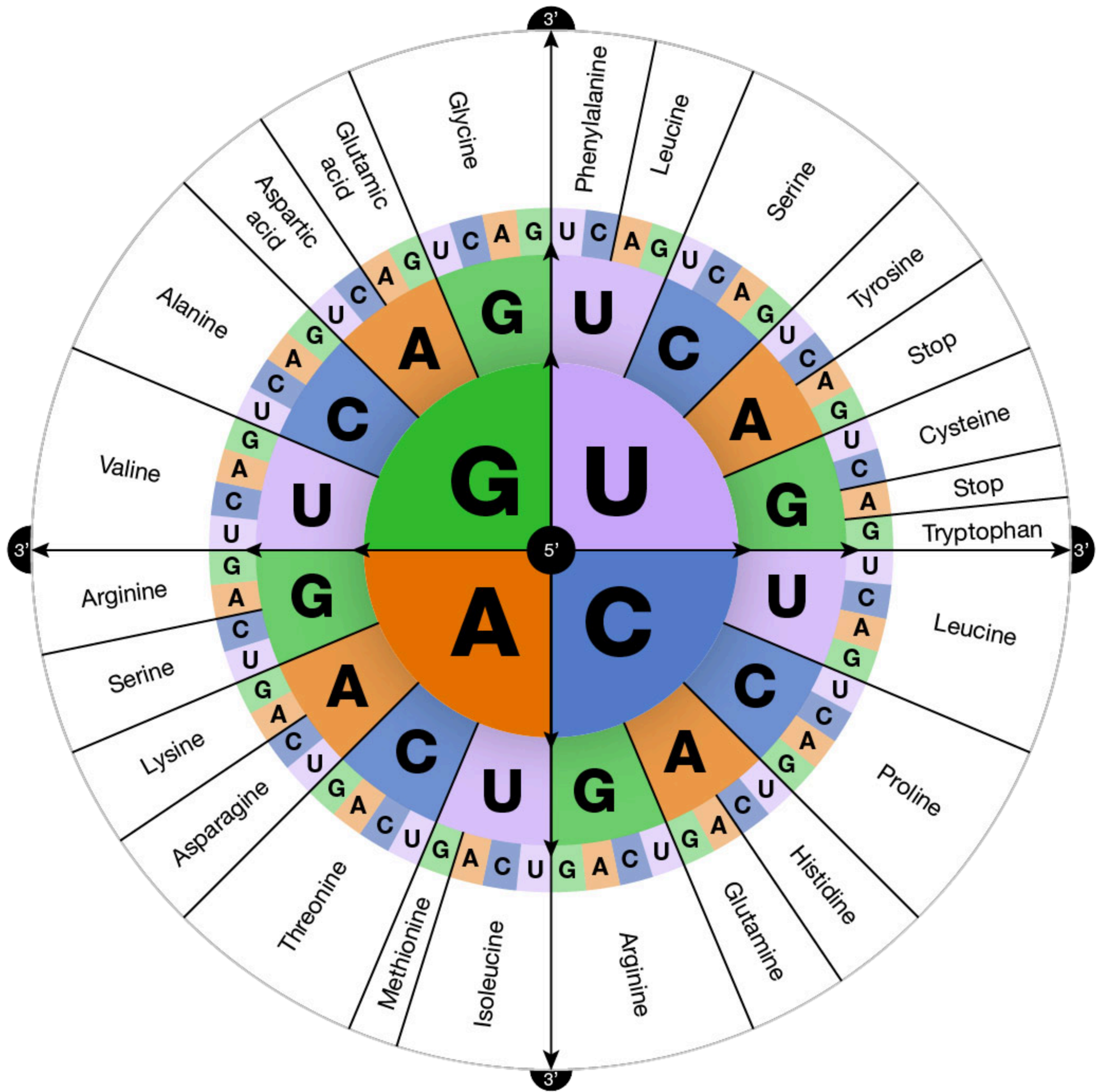


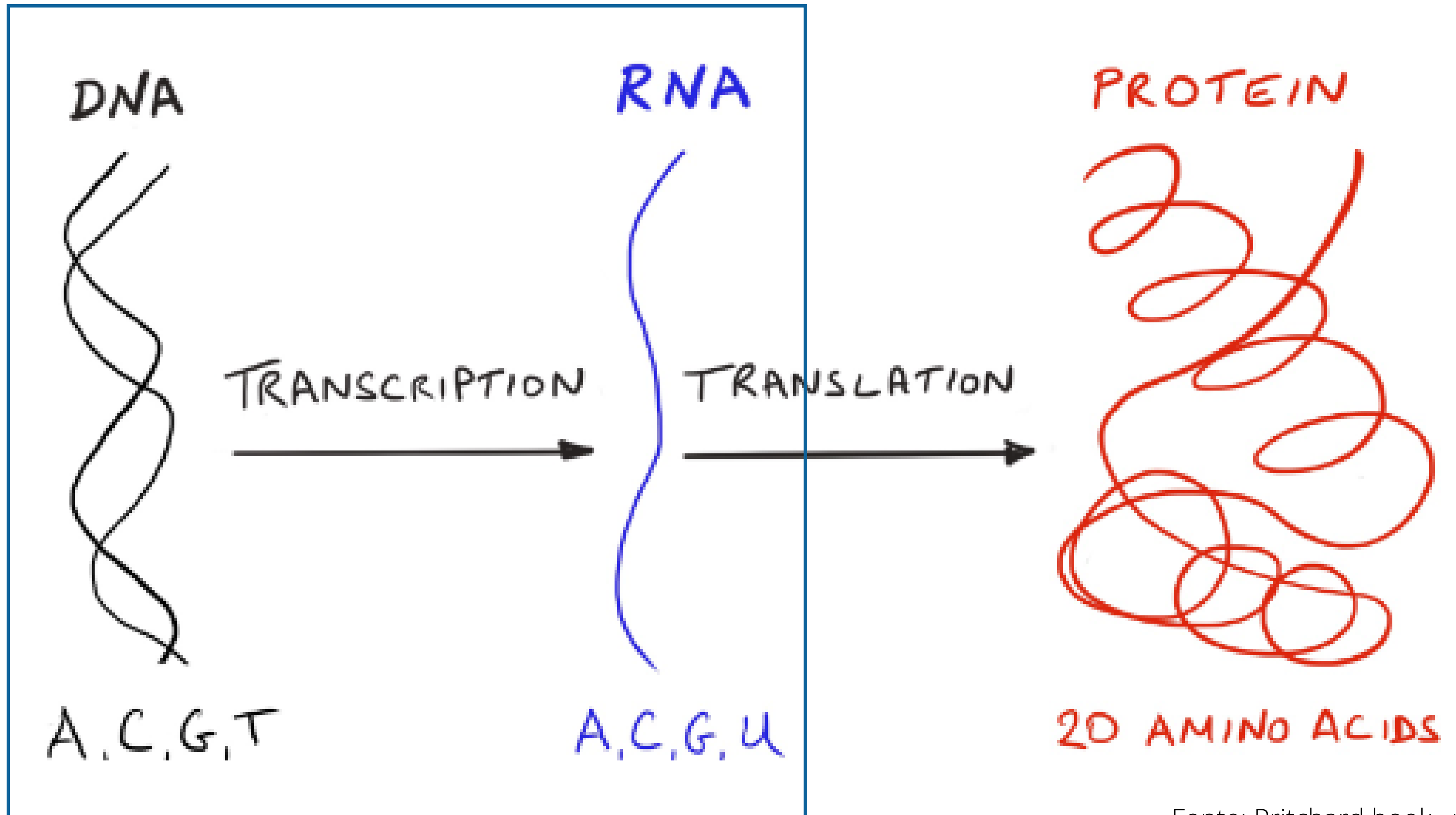
Histones



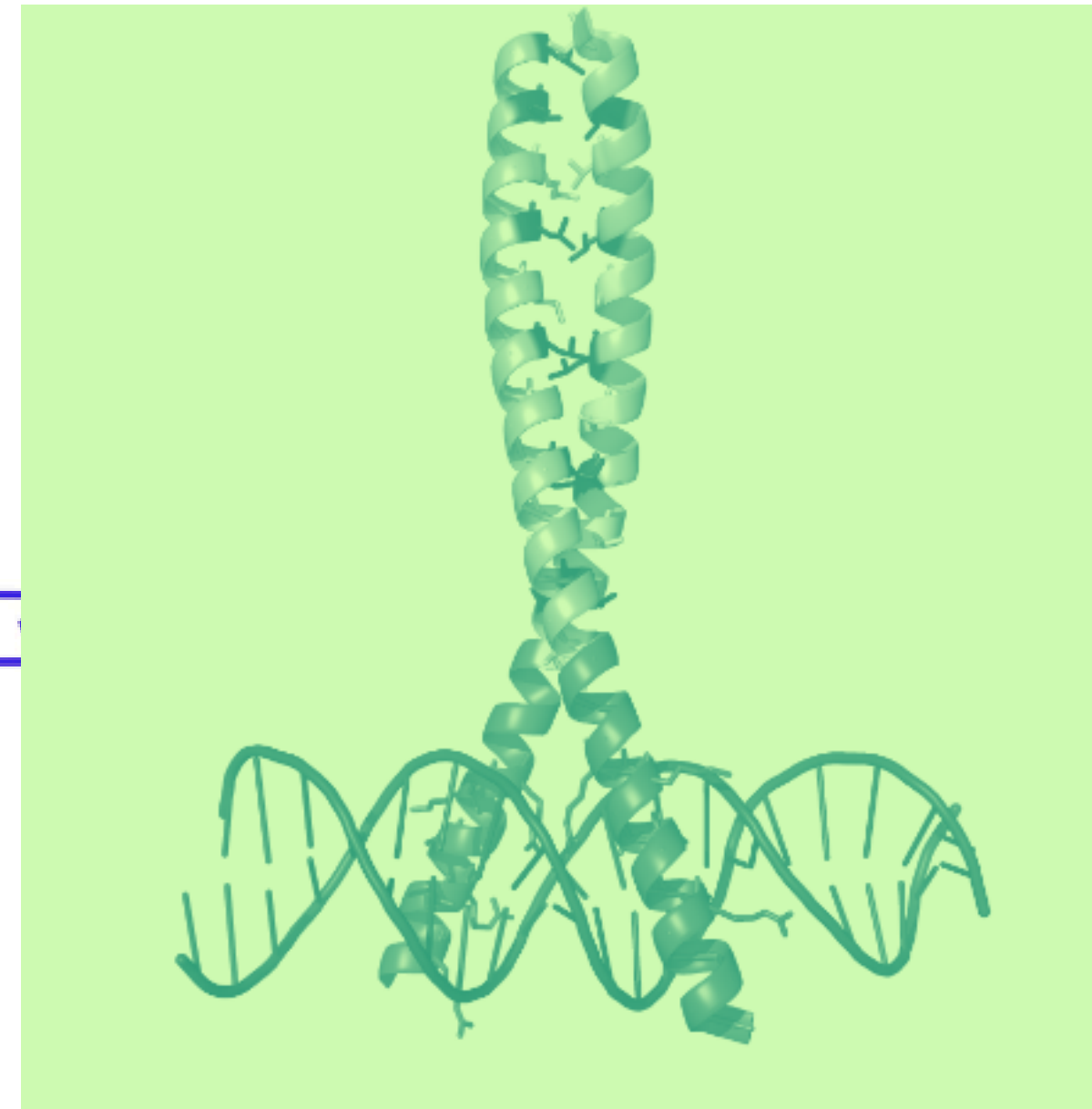
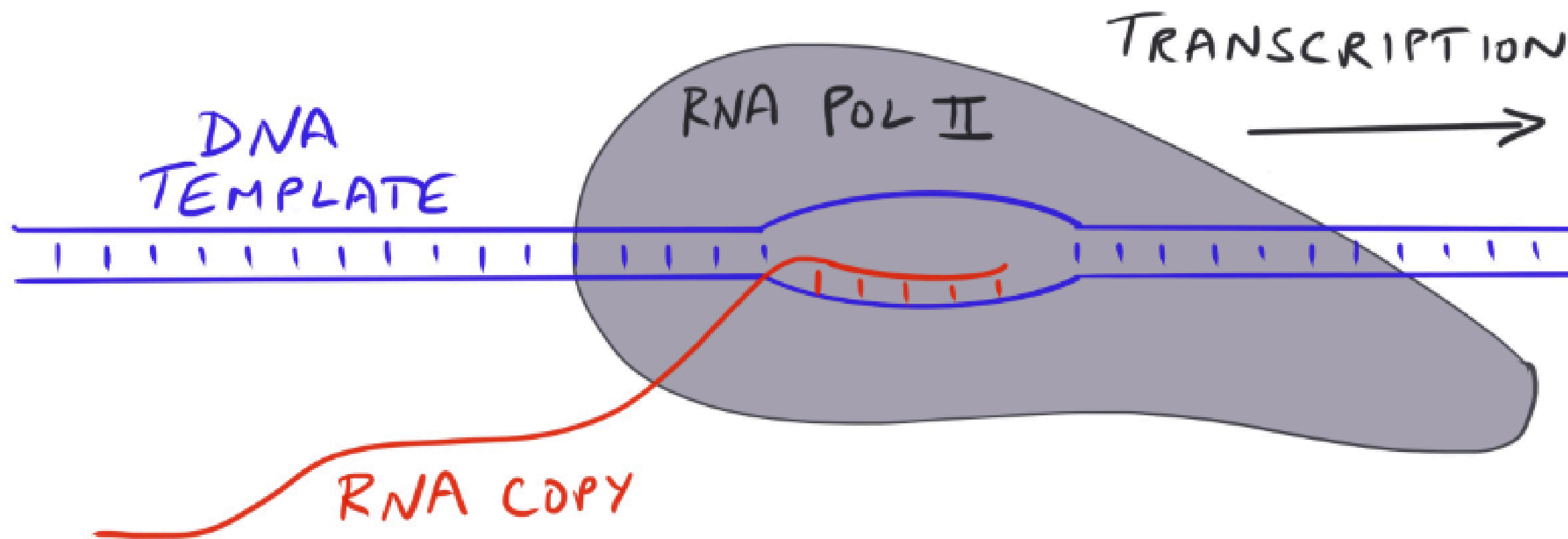




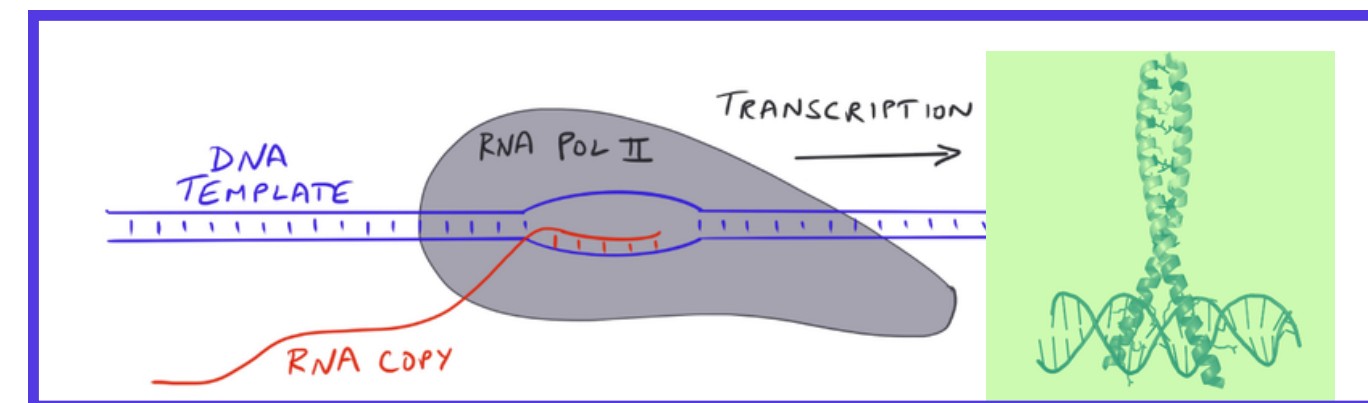
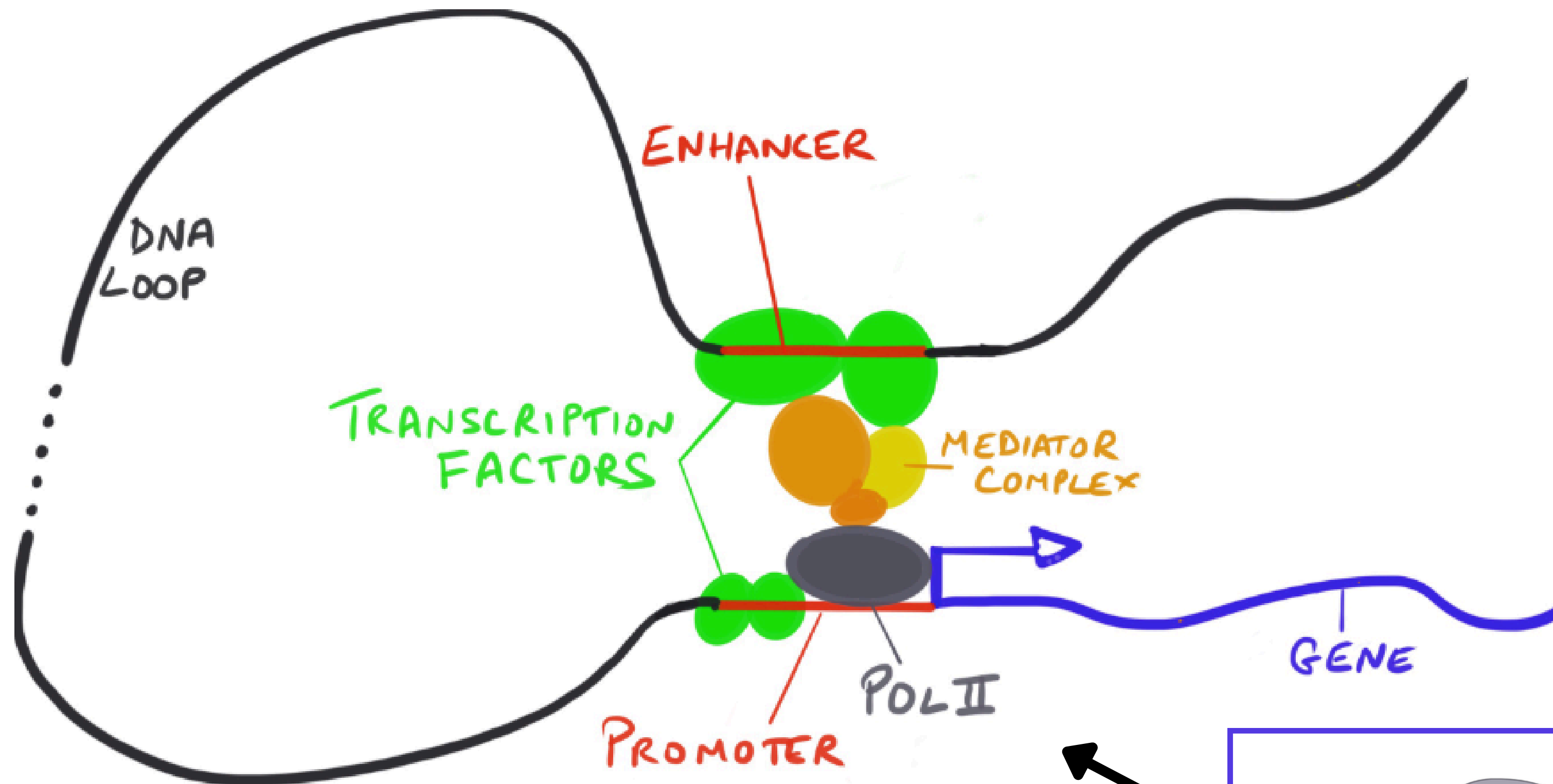




Transcrição e fatores de transcrição



Complexo mediador



Explorando o genome browser

UCSC Genome Browser Home

Explore, visualize, and compare genomic data with the UCSC Genome Browser: free, open-access tools for the human genome and beyond.



Exercício I

- Abra o google scholar e busque o gene mais estudado para o seu fenotipo de interesse (pode usar a IA);
- Abra o genome browser (escolha o hg38) e procure quais fatores de transcrição aparecem próximos ao gene mais estudado (olhar os tracks na base da página e clicar em refresh depois de escolher o track relacionado à regulação);
- Use zoom in e zoom out do próprio browser para encontrar os fatores de transcrição;
- Anote o gene, informe a quantidade de éxons para a referência MANE, cite 10 fatores de transcrição e o fenótipo.