

# CS102: Presenting Projects

Oliver BONHAM-CARTER

30 November 2022

# Presentations

CS102:  
Presenting  
Projects

Oliver  
BONHAM-  
CARTER

Presentations

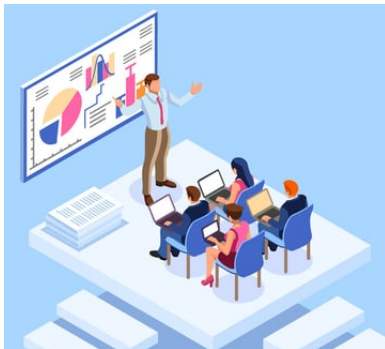
Let's Discuss

Structural  
Areas

CDS

Research  
Question

References



- **First Presentation:** during class, Wednesday 5<sup>th</sup> December 2022
- **Last Presentation:** Friday's class 7<sup>th</sup>
- Lab will not meet

# Imagine A Book

Intelligible text mixed with gibberish

CS102:  
Presenting  
Projects

Oliver  
BONHAM-  
CARTER

Presentations

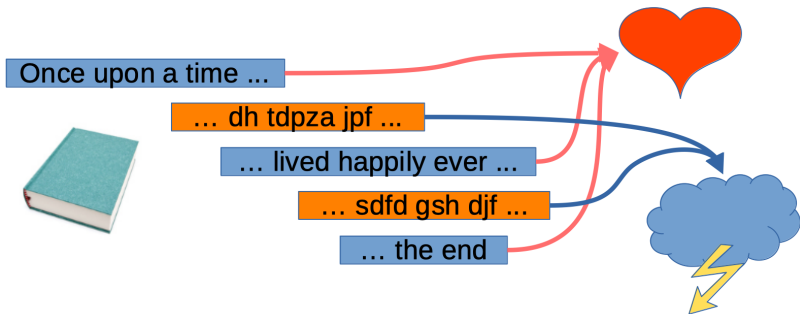
Let's Discuss

Structural  
Areas

CDS

Research  
Question

References



- Written in alphabetic characters
- The book's intelligible areas of language found by reading

# Structural Areas In DNA For Protein Synthesis

CS102:  
Presenting  
Projects

Oliver  
BONHAM-  
CARTER

Presentations

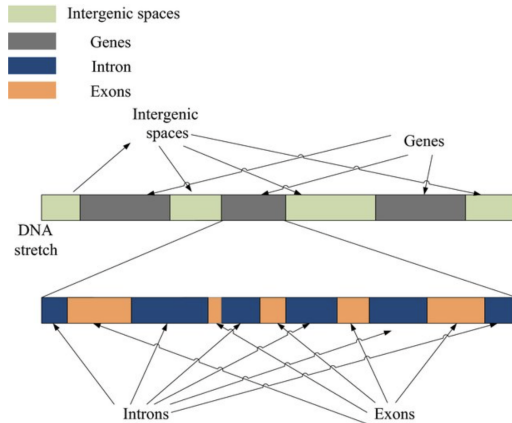
Let's Discuss

Structural  
Areas

CDS

Research  
Question

References



- **Exons** - Contains information to make protein product
- **Introns** - No information for protein product

# Data: Gene Records

Homo sapiens clone DNA71277 PINK1 (UNQ740) mRNA, complete cds

CS102:  
Presenting  
Projects

Oliver  
BONHAM-  
CARTER

Presentations

Let's Discuss

Structural  
Areas

CDS

Research  
Question

References

LOCUS AY358957 1768 bp mRNA linear PRI 03-OCT-2003  
DEFINITION Homo sapiens clone DNA71277 PINK1 (UNQ740) mRNA, complete cds.  
ACCESSION AY358957  
VERSION AY358957.1  
KEYWORDS FLI\_CDNA.  
SOURCE Homo sapiens (human)  
ORGANISM [Homo sapiens](#)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;  
Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1768)

<https://www.ncbi.nlm.nih.gov/nuccore/AY358957.1>

# Data: Gene Records contain CDS

## CDS (Coding Regions)

CS102:  
Presenting  
Projects

Oliver  
BONHAM-  
CARTER

Presentations

Let's Discuss

Structural  
Areas

CDS

Research  
Question

References

FEATURES	Location/Qualifiers
<b>source</b>	1..1768 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon: <a href="#">9606</a> " /clone="DNA71277"
<b>gene</b> .....	1..1768 /locus_tag="UNQ740"
<b>CDS</b> .....	152..481 /locus_tag="UNQ740" /note="PRO1446" /codon_start=1 /product="PINK1" /protein_id=" <a href="#">AA089316.1</a> " /translation="MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHS GTRTEVSTHTVPSKPGTASPCWPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGP CPGRRRD"

- What information is in a coding region in DNA?
  - Structure: landmarks, special words, scattered combinations of bases, etc.

# Coding Regions Have Unusual Compositions

CS102:  
Presenting  
Projects

Oliver  
BONHAM-  
CARTER

Presentations

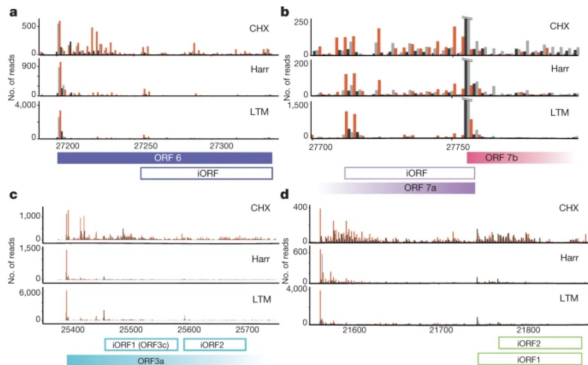
Let's Discuss

Structural  
Areas

CDS

Research  
Question

References



- Plots of composition describe activity
  - High count of certain words, bases, etc.
  - All visual perception ...

# Use Sound to Study Base Compositions in Genes?

CS102:  
Presenting  
Projects

Oliver  
BONHAM-  
CARTER

Presentations

Let's Discuss

Structural  
Areas

CDS

Research  
Question

References

- Sound contains many patterns (Refs 3 and 4)
  - That song you have never heard before that you recognize!



## Research Question

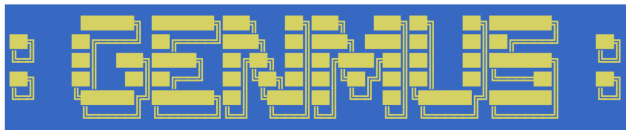
- Can coding regions of genes, when converted into “music”, be recognized by the **ear** for their structure?
  - Is there a *music* that implies genetic coding regions from non-coding regions?



# GenMus: Genetic Music

## Live Demonstration

- Genmus: software to convert DNA to piano music
- GitHub Repository:  
<https://github.com/developmentAC/geneticMusic>



- 1 Singh, Amit Kumar, and Vinay Kumar Srivastava. **"A tri-nucleotide mapping scheme based on residual volume of amino acids for short length exon prediction using sliding window DFT method."** Network Modeling Analysis in Health Informatics and Bioinformatics 9.1 (2020): 1-13.
- 2 Finkel, Yaara, *et al.* **"The coding capacity of SARS-CoV-2."** Nature 589.7840 (2021): 125-130.
- 3 Gupta, Shantanu Sen, Shifat Hossain, and Ki-Doo Kim. **"Recognize the surrounding: Development and evaluation of convolutional deep networks using gammatone spectrograms and raw audio signals."** Expert Systems with Applications 200 (2022): 116998.
- 4 Raponi, Simone, Gabriele Oligeri, and Isra Mohamed Ali. **"Sound of guns: digital forensics of gun audio samples meets artificial intelligence."** Multimedia tools and applications (2022): 1-26.