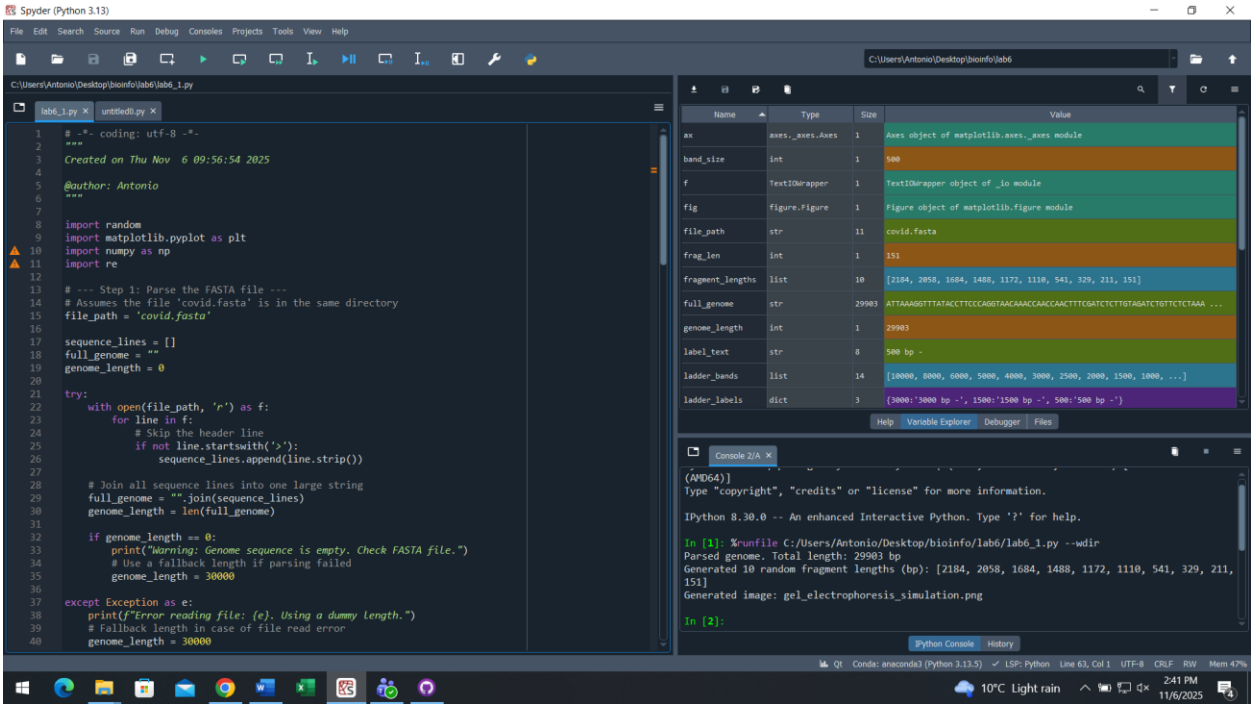
LABORATORY REPORT #6

GEORGESCU Mihai-Antonio, 1242EA
Bioinformatics, 4th year 1st semester, 2025-2026

lab6_1.py



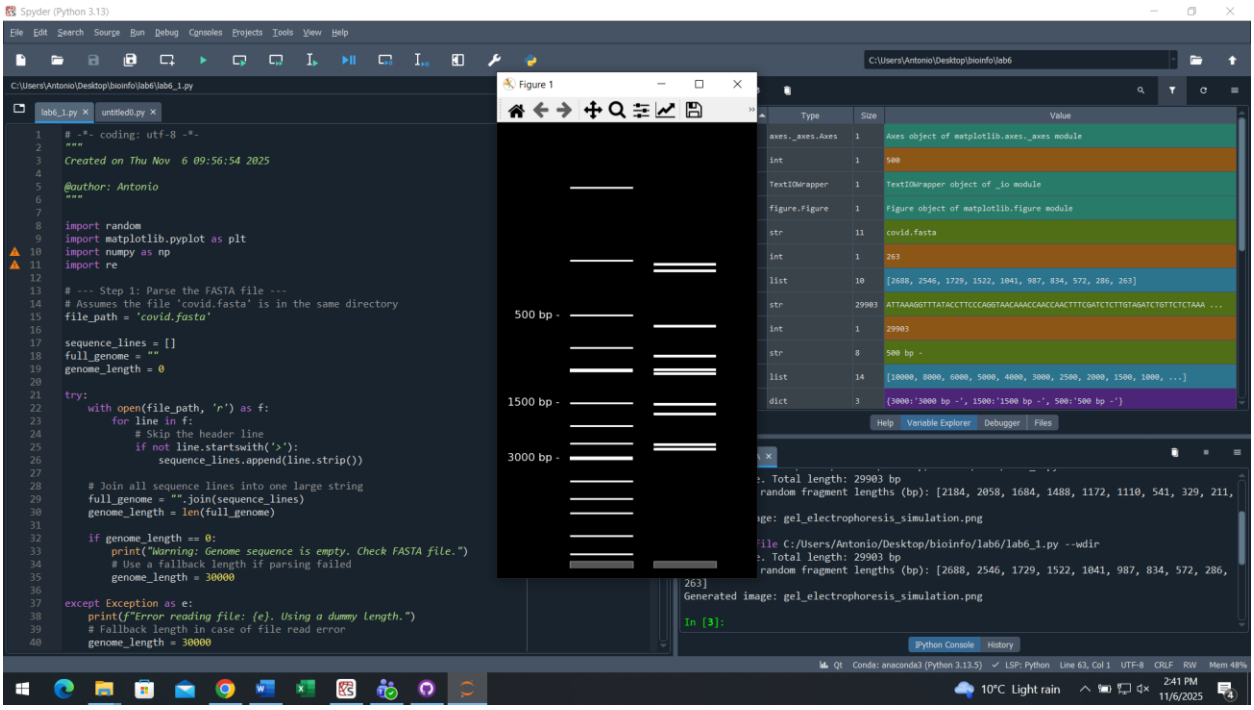


Figure 1
500 bp -
1500 bp -
3000 bp -

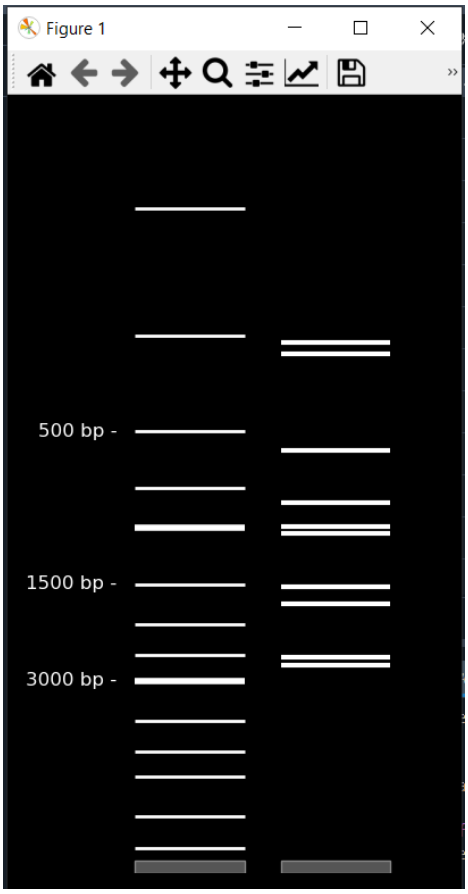
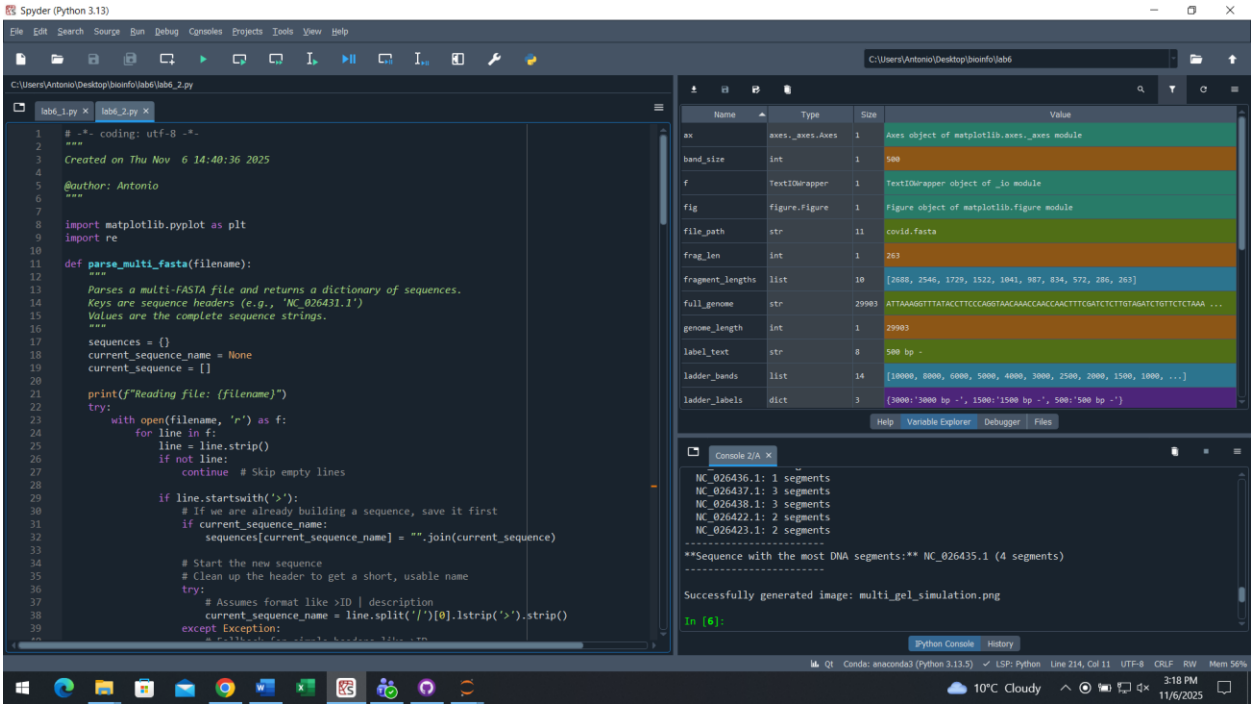


Figure 1
500 bp -
1500 bp -
3000 bp -

lab6_2.py


Spyder (Python 3.13)
-*- coding: utf-8 -*-
"""
Created on Thu Nov 6 14:40:36 2025
@author: Antonio
"""
import matplotlib.pyplot as plt
import re
def parse_multi_fasta(filename):
"""
Parses a multi-FASTA file and returns a dictionary of sequences.
Keys are sequence headers (e.g., 'NC_026431.1')
Values are the complete sequence strings.
"""
sequences = {}
current_sequence_name = None
current_sequence = []
print(f"Reading file: {filename}")
try:
with open(filename, 'r') as f:
for line in f:
line = line.strip()
if not line:
continue # Skip empty lines
if line.startswith('>'):
# If we are already building a sequence, save it first
if current_sequence_name:
sequences[current_sequence_name] = "".join(current_sequence)
# Start the new sequence
# Clean up the header to get a short, usable name
try:
# Assumes format like >ID | description
current_sequence_name = line.split('|')[0].lstrip('>').strip()
except Exception:
NC_026436.1: 1 segments
NC_026437.1: 3 segments
NC_026438.1: 3 segments
NC_026422.1: 2 segments
NC_026423.1: 2 segments
Sequence with the most DNA segments: NC_026435.1 (4 segments)
Successfully generated image: multi_gel_simulation.png
In [6]:
10°C Cloudy
3:18 PM
11/6/2025

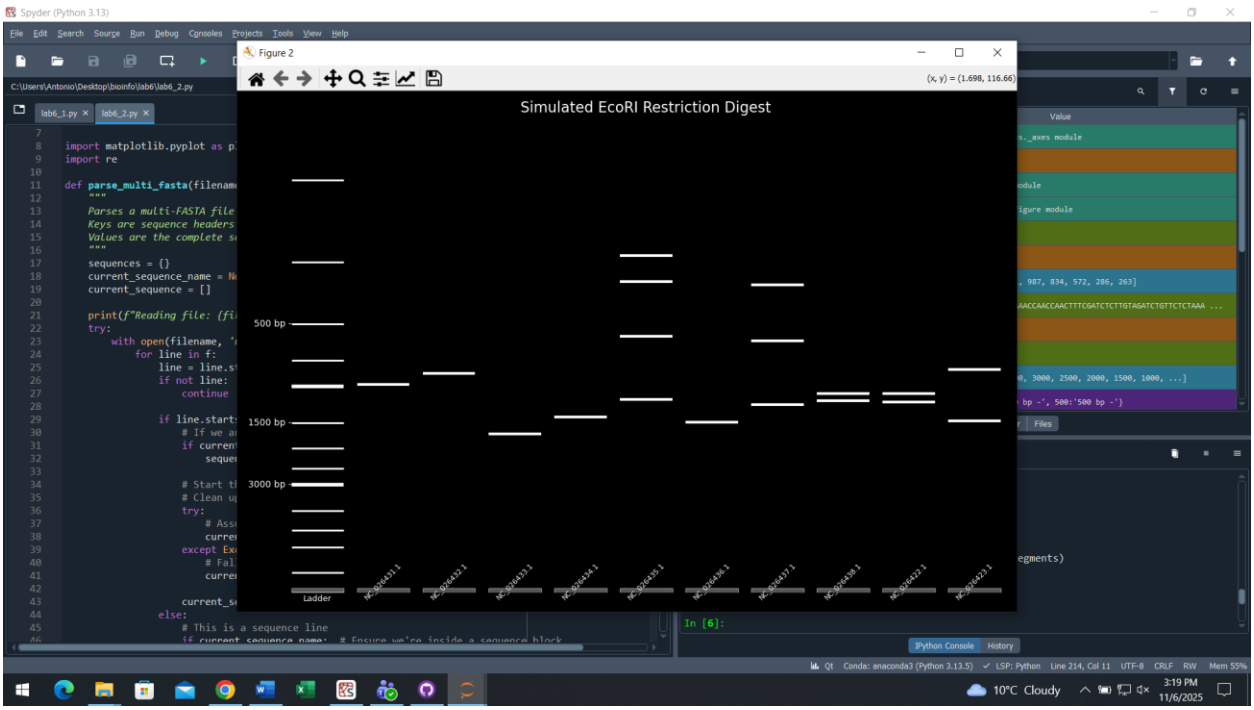
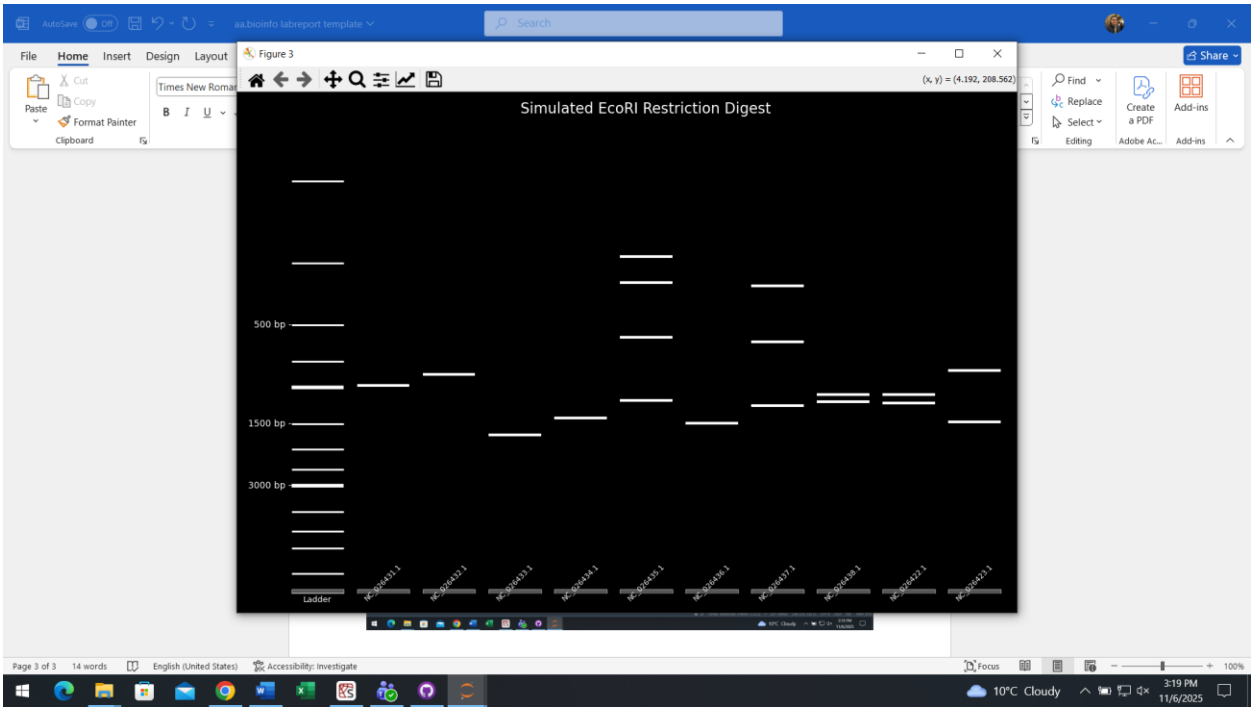


Figure 2
Simulated EcoRI Restriction Digest
500 bp
1500 bp
3000 bp
Ladder
In [6]:
10°C Cloudy
3:19 PM
11/6/2025



NC_026435.1: 3 bands

NC_026437.1: 2 bands

NC_026438.1: 2 bands

NC_026422.1: 2 bands

NC_026423.1: 2 bands

NC_026431.1: 1 band

NC_026432.1: 1 band

NC_026433.1: 1 band

NC_026434.1: 1 band

NC_026436.1: 1 band