Smart Restriction Enzyme Selector (SRES)

Background:
Molecular cloning is foundational to genetic research, allowing scientists to amplify, recombine, and insert DNA sequences for downstream analysis. However, one important step in the cloning pipeline is the selection of restriction enzymes that cleave the ends of a DNA insert without cutting internally. This requires a careful search for enzymes that can:

1. Recognize specific flanking sequences,
2. Do not have additional recognition sites inside the insert, and
3. Are compatible in terms of ligation efficiency and buffer conditions.

Currently, researchers are forced to consult multiple databases (like NEB and REBASE), spreadsheets, and web tools, making enzyme selection pretty inefficient and error-prone. This project proposes a Smart Restriction Enzyme Selector, a centralized tool to streamline this step. Users can paste a DNA sequence into a web interface, and the tool will recommend compatible 5′ and 3′ restriction enzymes that do not cut internally and are suitable for cloning.

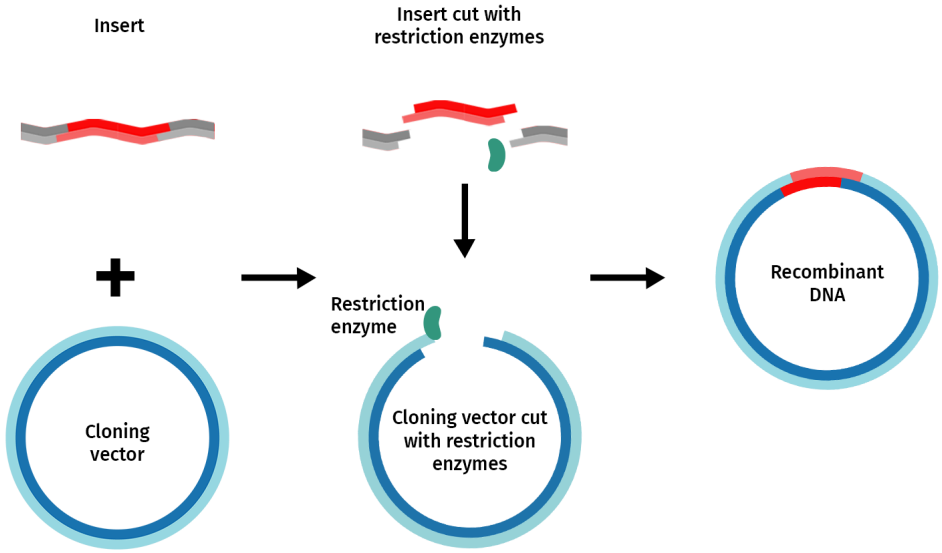

Figure 1. Restriction enzyme cloning process (image from Snapgene)

Functionality:

The tool will require a single input: a DNA sequence (in plain text or FASTA format). Optional settings will allow the user to specify whether they prefer sticky ends, blunt ends, or only enzymes that cut at the sequence boundaries.

The tool will output:

1. A list of enzyme pairs that do not cut internally.
2. Overhang information (5', 3', or blunt).
3. Buffer compatibility (based on NEB buffers or a simplified version).
4. Overhang sequence visualizations.
5. Suggested combinations for cloning with minimal effort.

By scanning for enzyme recognition sites through consulting a database of enzyme cut patterns and buffer conditions and then applying logical filters, the backend will produce a curated result for each request. This will save researchers from the painful trial-and-error enzyme screening process and speed up the overall cloning design workflow.

Tool Description:

The tool will use the following software components:

1. **SQL Relational Database** – Contains information on common restriction enzymes, including:
    - Recognition sequence (e.g., GAATTC)
    - Cut site position
      Overhang type and sequence
    - Buffer compatibility
    - Supplier reference links (e.g., NEB)

2. **Python-based CGI Backend** – Handles user input, scans the DNA sequence for internal recognition sites using string pattern matching, filters out incompatible enzymes, and identifies compatible pairs with matching buffer conditions and ligation-friendly overhangs. It queries the SQL database using mysql.connector and returns the selected enzyme pairs with necessary metadata.

3. **HTML/CSS/JavaScript GUI** – Provides a clean, interactive user interface for:
    - Sequence input
    - Output table with filterable columns (buffer compatibility, overhang type)
    - Links to additional information
    - Visual highlights of where enzymes cut in the input sequence (if applicable)
    - All HTML and JavaScript will be separated from CGI logic to follow the content separation principles we learned in class, with external files for styling and interactivity.

References:

Restriction enzyme cloning - Snapgene. (n.d.).

https://www.snapgene.com/guides/restriction-enzyme-cloning