Pyrms Data Analysis User Manual

Introduction

This Python script analyzes Pyrms sample data, focusing on sample distribution across locations. The sample used was obtained from a research study entitled "*Sex at the origin: an Asian population of the rice blast fungus Magnaporthe oryzae reproduces sexually*".

Background

- The capacity of organisms to adjust to their surroundings is intricately tied to their reproductive mechanisms
  - reshuffling existing genetic material
  - creating selectable genetic variation
- Segregation and recombination during sexual reproduction
  - eliminate detrimental mutations
  - fix advantageous mutations
  - enhance effectiveness of natural selection
- Sexual reproduction
  - crucial for species navigating dynamic envrionments
    - pathogens adjusting to new hosts or evolving agricultural practices
- Lack of recombination
  - preservation of beneficial associations between loci
  - facilitates retention of local adaptations once they have emerged
- Fungal pathogens adopt a **dual strategy**
  - alternate between **recombination and vegetative reproduction**
    - allows adaptation to changes
    - poses a threat to longevity of control measures

Magnaporthe oryzae

- Fungus causing blast disease in rice and other grasses.
- Considered a model phytopathogenic species (Wilson & Talbot 2009).
- **Genetic Structure and Reproductio**n
  - Genetic structure suggests predominantly clonal reproduction in most rice-growing regions.

- Observed in various countries (Colombia, USA, Europe, Korea, Japan: Zeigler 1998; Morocco: El Guilli et al. 2005; Madagascar: Andriantsimialona & Tharreau 2008).
- **Absence of Meiosis Organs**
  - Despite thorough investigations, perithecia (organs for meiosis) have never been reported in the field (Zeigler 1998).
- **Sexual Reproduction *in vitro***
  - In vitro studies have demonstrated sexual reproduction between strains collected from natural environments (Silue´ & Notte´ghem 1990; Hayashi et al. 1997).
- **Sexual Cycle Requirements**
  - Sexual cycle of this fungus requires two strains of opposite mating types (MAT1-MAT2).
  - At least one strain must be female-fertile, capable of producing perithecia.
- **Significance of Female Fertility**
  - Female fertility is a crucial biological characteristic influencing *M. oryzae*'s mode of reproduction.
- **Rare Collection of Female-Fertile Strains**
  - Female-fertile strains of Magnaporthe oryzae are infrequently found in rice.
  - Primarily collected in the Yunnan province of China, northern Thailand, and northern India (Zeigler 1998; Kumar et al. 1999; Mekwatanakarn et al. 1999).
- **Domestication Centers**
  - Southern China and northern India are recognized as the two domestication and diversification centers for Asian cultivated rice, *Oryza sativa* (Londo et al. 2006)
- **Probable Center of Origin**
  - Genetic diversity analyses suggest the Himalayan foothills as the probable center of origin for *M. oryzae* populations causing rice pathology (Zeigler 1998; Tharreau et al. 2009).

Significance

- Fungal pathogens in cultivated plants offer a unique opportunity for studying sex evolution.
  - A single species can utilize multiple reproductive strategies.
  - Evolutionary changes occur more rapidly in agro-ecosystems than in natural ecosystems.
    - The accelerated evolution is attributed to high reciprocal selection pressures

Installation Requirements

Python Installation:
- Install Python (version 3.x recommended) from the official Python website.

Required Libraries Installation
- Install the necessary libraries:
    - Pandas: pip install pandas
    - NumPy: pip install numpy

The program was developed and run in a laptop that has the following specifications:
- System Manufacturer: Dell Inc.
- System Model: Inspiron 5405
- BIOS: 1.12.0
- Processor: AMD Ryzen 5 4500U with Radeon Graphics
- Memory: 16384MB RAM
- Page file: 16190MB used, 3016MB available
- DirectX Version: DirectX 12

Getting Started

Dataset Preparation
- Prepare dataset in CSV format.
- Include columns like country, location, collected by, populations, individuals, and Pyrms sample columns (e.g., Pyrms233-234, Pyrms319-320, etc.).

Install Dependencies

- Ensure you have Python (version 3.6 or later) installed.
- Use pip to install the required packages:lib tk

```
pip install pandas matplotlib tk
```

Usage

Run the Application
- Retrieve the code through this link:
  https://github.com/andreapongco/CMSC191_Project.git

- To launch the application, simply navigate to the directory where all the python files are located. This may be done by navigating the file explorer or using the following command

```
cd <file_path>
```

- Launch the application by executing the Python script

```
python3 Plot.py
```

Using the Application

- The application window opens with a dropdown menu.
- Select Pyrms from the dropdown list.
- Click the "Plot" button to generate the stacked bar graph displaying Pyrms type counts per country.
- The graph will visualize the counts of different Pyrms types across countries for the selected Pyrm.

Troubleshooting

- Missing Data Label: Any Pyrms types with a value of 0 in the dataset will be labeled as "Missing Data" in the generated graph.

Notes

- Replace 'Dataset-Saleh-et-al.csv' with the correct file path to your dataset containing Pyrm information.
- Modify the code to suit your dataset structure if necessary.