DNA Barcode Processor User Manual

1.   Overview
The DNA Barcode Processor is a versatile tool designed for researchers in the field of genetics, specifically those working with DNA barcoding. This manual provides comprehensive guidance on how to effectively use the software to process DNA sequences, calculate specificity and recall rates, create visual representations of barcode sequences, and design primers for targeted DNA amplification.

2.   Installation
Before running the DNA Barcode Processor, ensure you have Python installed on your system (version 3.6 or later is recommended). This application relies on Tkinter for the GUI, Pandas for data handling, and python-docx for document manipulation, which are included in standard Python installations or can be easily installed via pip.

3.   Features and Functions
3.1. File and Folder Browsing
- **Browse File**: Allows the user to select a single file (e.g., FASTA, TXT, DOCX) from their file system.
- **Browse Folder**: Enables the selection of a directory, facilitating operations on multiple files or storing output files.

3.2. Simplified Base Processing (DeleteDBs)
- **Input**: Requires a sequence file path in FASTA, FAS, or TXT format.
- **Output**: Specifies the path where the processed file, with simplified bases, will be saved.

3.3. Specificity and Recall Rate Preprocessing
- **Input**: Accepts a path to a sequence file containing both the sequences and barcode fragments.
- **Output Directory**: Designates a folder path to save the processed sequences, preparing them for specificity and recall rate calculation.

3.4. Recall Rate Processing
- **Barcode File**: Input file containing barcode sequences.
- **Species Sequence Folder**: Directory with species sequence files.
- **Output File**: Location to save the calculated recall rates.

3.5. Specificity Processing
Similar to recall rate processing, but focused on calculating the specificity of the barcodes against the provided species sequences.

3.6. Barcode Visualization (Barcodemaking)
- **Input File**: Specifies the path to a document containing nucleotide sequences for

visualization.
- **Output File**: Path to save the visually formatted barcode sequence document.

3.7. Primer Design (Primermaking)
- **Input File**: Text file containing nucleotide sequences for which primers are to be designed.
- **Output File**: Excel file path to save the designed primers and their properties.

4.   Running the Application
Start the application by executing the script in a Python environment. The graphical user interface (GUI) allows easy navigation between different functionalities, enabling users to process data efficiently.

4.1. Select the desired operation from the menu.
4.2. Fill in the required fields for each function.
4.3. Click the corresponding "Browse..." buttons to select files or folders.
4.4. Hit the "Start Process" button to execute the operation.

5.   Troubleshooting
- Ensure all input files are in the correct format as specified.
- Check the output paths to ensure they point to valid directories or file names.
- For large datasets, processing times may be extended; ensure the application is not interrupted during these operations.

6.   Support
For technical support or to report bugs, please contact the software development team at support@dnabarcodeprocessor.org.

7.   Citation
If you use the DNA Barcode Processor in your research, please cite it as follows:
You C, Jiang S, Ding Y, Ye S, Zou X, Zhang H, Li Z, Chen F, Li Y, Ge X, Guo X. RNA barcode segments for SARS-CoV-2 identification from HCoVs and SARSr-CoV-2 lineages. Virol Sin. 2024 Feb;39(1):156-168. doi: 10.1016/j.virs.2024.01.006. Epub 2024 Jan 20. PMID: 38253258; PMCID: PMC10877444.

This manual is intended to assist users in the effective utilization of the DNA Barcode Processor for their research purposes. For further details or updates, please visit the official website or contact the support team.