Pan-Genomic Graph Building and Visualization

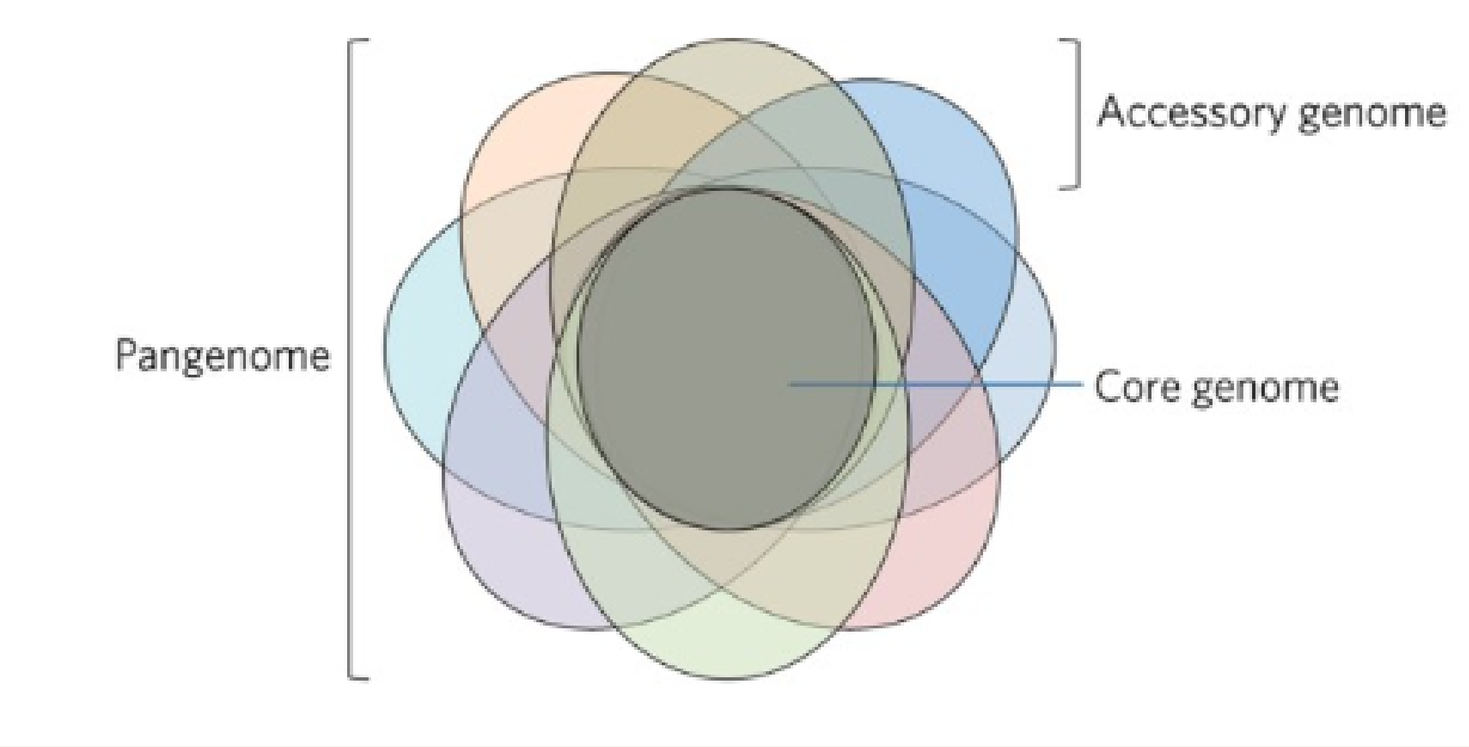
UNCOVERING GENETIC DIVERSITY

AGENDA

- Introduction to Pan-Genome
- Classification of Pan-Genome
- Types and Levels of Pan-Genome
- Tools and Methods to Create Pan-Genome
- Pan-Genome Indexing Tools
- Pan-Genome Formats and Visualization
- Uses and Applications of Pan-Genome
- Pan-Genome Graph Construction and Visualization Case Study

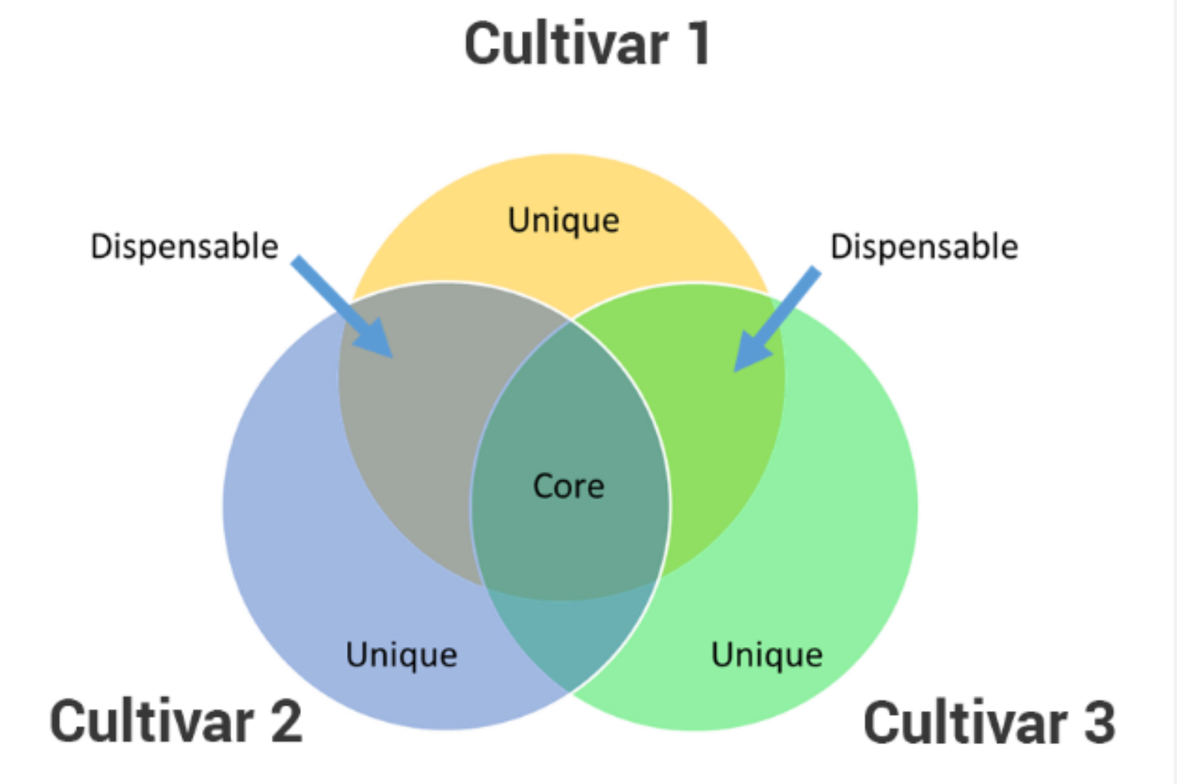
INTRODUCTION TO PAN-GENOME

- Pangenome or Supragenome is the entire set of genes from all strains within a clade
- The field of study of pangenome is called pangenomics



CLASSFICATION OF PAN-GENOME

- Core Genome - shared by all genomes
- Shell Genome - shared by majority genomes
- Dispensable Genome - shared by minimal genomes

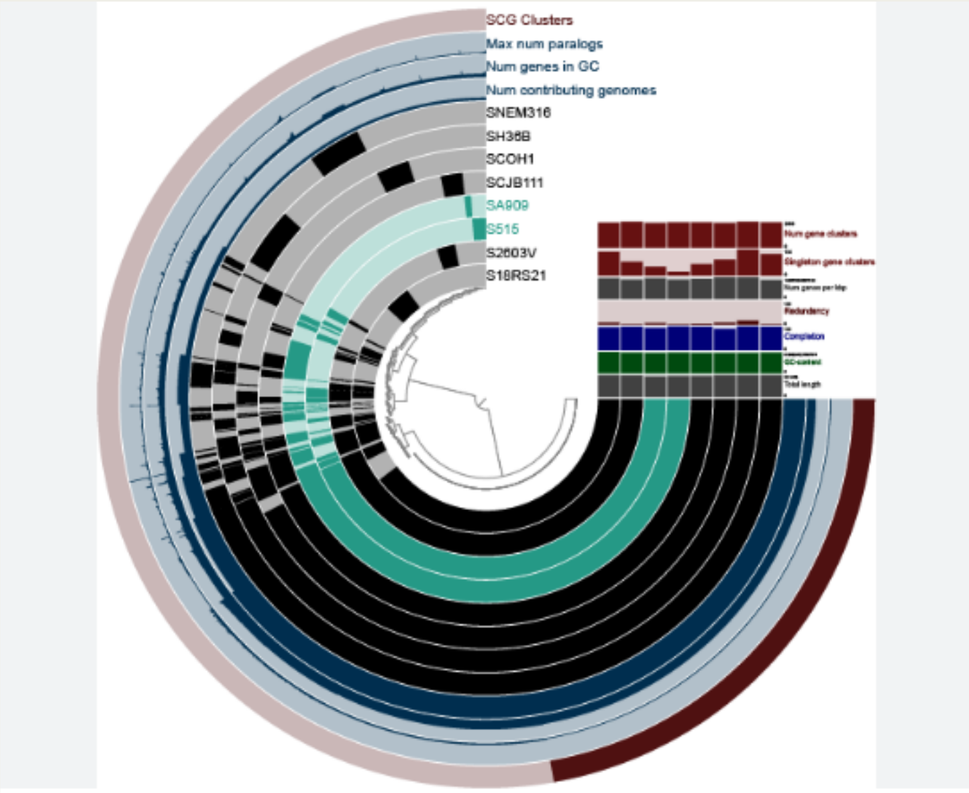


TYPES AND LEVELS OF PAN-GENOME GRAPHS

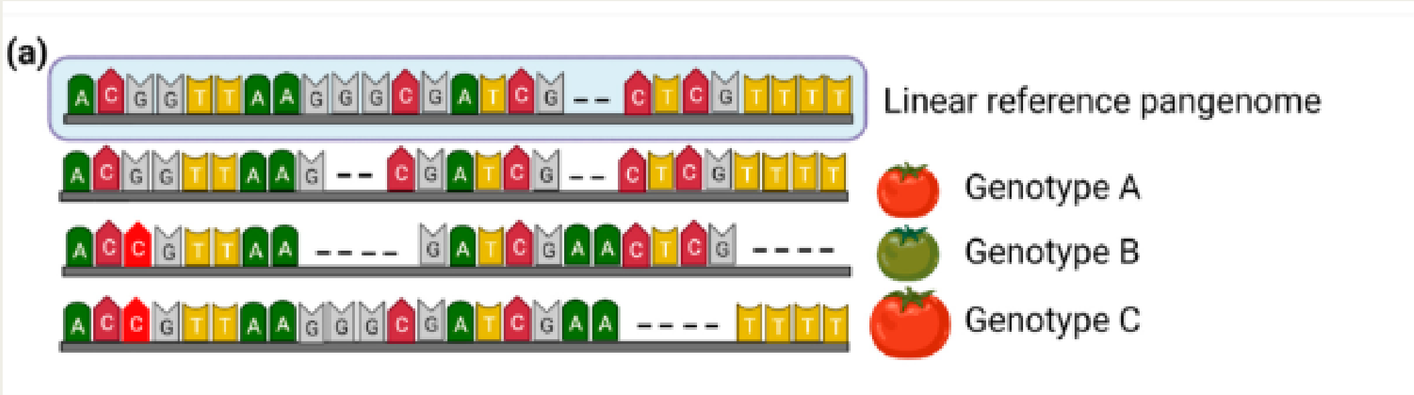
- Types of Pangenomes :-
  1) Circular Pan-Genome Graph
  2) Linear Pan-Genome Graph
  3) Graph-Based Pan-Genome

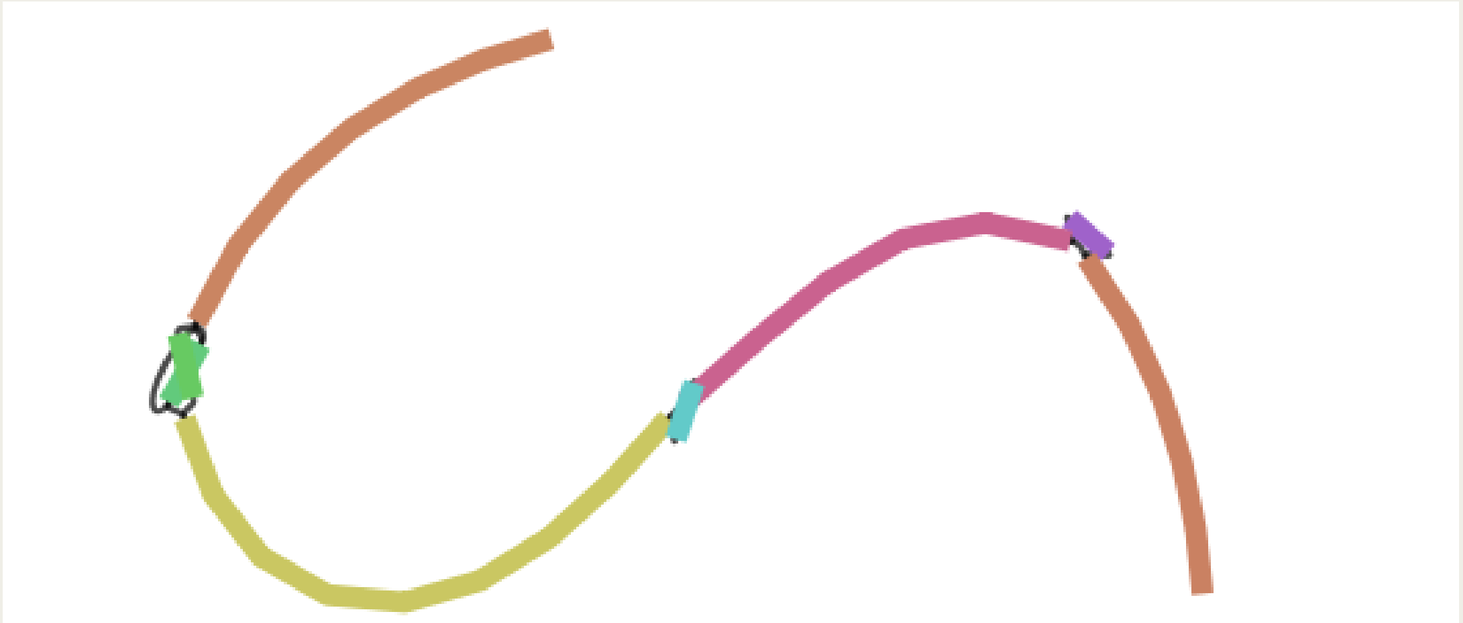
- Levels of Pangenomes :-
  1) Species-specific
  2) Genus-specific
  3) Family-specific

TYPES AND LEVELS OF PAN-GENOME GRAPHS



CIRCULAR PANGENOME





GRAPH BASED PANGENOME

TYPES OF PANGENOME GRAPHS

- Pan-Genome Graphs :-
    - Variation graphs
    - De Bruijn Graphs
    - Sequence Graphs
    - Genome Feature and Variation Graphs (GFA)
    - Interval Graphs

TOOLS AND METHODS TO CREATE PAN-GENOME GRAPHS

- Tools to create Pan-Genome Graphs :-
  - Basic Pan-Genome Graphs :- minigraph
  - Variation graphs :- vg , GFAKluge
  - Sequence graphs :- SeqAn, vg
  - Genome Feature and Variation Graphs (GFA) :- GFAKluge
  - Interval graphs :- BED Tools

- Other common tools :- PanOct, PGAP, PanGenomeTool

PAN-GENOME INDEXING TOOLS

- Pan-Genome Visualization ways and tools :-
  - GCSA2: Used with variation graphs, which represent genetic variation across multiple genomes or individuals.
  - SeqAn: Provides functionalities for working with various types of sequence graphs, including pangenome graphs.
  - BWBBLE: Utilizes VCF-based pangenome graphs, where variations are encoded using IUPAC codes for substitution and additional sequences for insertions and deletions.
  - GFAKluge: Designed for working with Genome Feature and Variation Graphs (GFA), which are commonly used to represent pangenome graphs.

PAN-GENOME FORMATS AND VISUALIZATION

- Pan-Genome Formats :-
  - FASTA format
  - Variant Call Format (VCF)
  - Graphical Fragment Assembly (GFA)
  - JSON (JavaScript Object Notation
  - Binary Alignment/Map (BAM)

PAN-GENOME FORMATS AND VISUALIZATION

- Pan-Genome Visualization ways and tools :-

    - Circular Plots :- Circos, PanGP, Gview
    - Linear Plots :- GenomeDiagram, Gview
    - Heatmaps :- Python libraries, Heatmapper
    - Network Visualization :- Bandage, Gephi, igraph
    - Interactive Web-Based Tools :- PanWeb, PanX, PanGIA

USES AND APPLICATIONS OF PAN-GENOME GRAPHS

- Genome Annotation and Gene Discovery
- Understanding Genetic Variation and Diversity
- Comparative Genomics
- Functional Genomic
- Medicinal and crop improvement

PAN-GENOME GRAPH CONSTRUCTION AND VISUALIZATION CASE STUDY

- Pan-Genome Graph Construction :- using Minigraph
- Pangenome Data Representation :- using gfa format
- Pangenome Data Visualization :- using Bandage

RESULT