

Pan WSGA Web Server Tutorial

A Screenshot-Based Guide for Single Genome Analysis to identify unique features of a Bacterium

Section: I

Overview of PanWSGA

PanWSGA

Reference-Guided Pangenome Analysis for Rapid Bacterial Genome Interpretation

PanWSGA enables rapid interpretation of newly sequenced bacterial genomes using curated species-level reference pangenomes. Project your genome to isolate core, accessory, and unique genes with dynamic visualizations.

[Get Started](#)[Learn More](#)

The navigation menu provides access to all major sections of Pan WSGA



Key Features



1. Single-Genome Pangenome Projection

Classifies genes from a single genome as core, accessory, cloud, or unique by mapping them to a curated species pangenome. Full context without multi-genome builds.



2. Unique Gene Finder

Strict validation thresholds ($\geq 90\%$ identity and $\geq 90\%$ coverage) to find unique genes, generating clean FASTA collections and structured logs.



3. GeneEye Visualization Module

Publication-ready summary dashboards combining gene classification and EggNOG annotations with distribution graphs, bar charts, and tabular outputs.

Tell User live Job status of the Server

Analysis Platform



Comparative Analysis Dashboard

Upload assemblies, configure parameters, and execute genomic comparison pipelines
against curated reference databases.

System Job Overview

[Live Status](#)

TOTAL ANALYSES

83



ACTIVE JOBS

1



QUEUED JOBS

0



SYSTEM HEALTH

Optimal



Quick Resume Analysis



 Retrieve the status or details of any previously completed comparative job.

Upload Genome FASTA

Upload your genome sequence file in FASTA format to run comparisons against reference databases.

Input File Type

Section II

**Species cataloug
(Referance Pangenome Database)**

To access Reference Database

Species Catalog

Browse all available species in the database with their gene statistics and genome information.



Search Species

Sort By

Total Genomes (High to Low) ▼



Acinetobacter pittii

42,695

Total Genes

1,002

Genomes

2,022

Core Genes

40,673

Accessory

Added: Dec 6, 2025

[Information file](#)

Klebsiella pneumoniae

37,050

Total Genes

1,000

Genomes

3,818

Core Genes

33,232

Accessory

Added: Jul 9, 2026

[Information file](#)

Salmonella enterica

32,819

Total Genes

1,000

Genomes

3,106

Core Genes

29,713

Accessory

Added: Jul 9, 2026

[Information file](#)

Section III

Demo Datasets

Demo Data

Test our tools with pre-loaded curated genomes. Explore the full pipeline without uploading your own files.

1 Select Genome



2 Choose Species



3 Configure & Run

Quick Testing Guide

Browse and select a pre-loaded demo genome from the cards below.
Choose the target species and its corresponding reference pan-genome.
Configure settings and hit "Run Pipeline Test".

Pre-loaded Demo Genomes



Treponema paraluisccuniculi

GCF_000217655.1

Demo dataset for testing

• 982 predicted genes

Select for Testing

Download Dataset

View Results



Mycoplasmoides genitalium

GCF_040556925.1

Demo dataset for testing

• 524 predicted genes

Select for Testing

Download Dataset

View Results



Fructilactobacillus florum

GCF_000512295.1

Demo dataset for testing

• 1,259 predicted genes

Select for Testing

Download Dataset

View Results



Chlamydia pneumoniae

GCF_000007205.1

Demo dataset for testing

• 1,049 predicted genes

Select for Testing

Download Dataset

View Results

Section IV

Help & Frequently ask Questions

Frequently Asked Questions

Find answers to common questions about uploading genomes, running pipelines, and interpreting results.

Input Data Options

What file formats are accepted for genome upload?

PanWSGA accepts standard nucleotide FASTA formats: .fasta, .fa, and .fna. Files must contain valid DNA sequences. The maximum allowed file size is 200 MB. If you have pre-predicted coding sequences, use the CDS input mode.

Should I upload genomic FASTA or CDS FASTA?

Is there a size limit for uploaded genomes?

Pipeline & Settings

How long does an analysis take?

What is the EggNOG annotation stage?

Can I re-run or resume a past analysis?

Results & Interpretation

What do Core, Accessory, and Unique genes mean?

Why are results different from my own BLAST search?

How do I download my results?

Still have questions?

Check the User Guide or reach out to the support team.

[View User Guide](#)[Email Support](#)

Help & Documentation

Step-by-step instructions for using the Pan-Web Single Genome Analyzer platform.

Quick Start Guide

Step 1 — Upload Your Genome

Navigate to the **Dashboard** and go to the "**Pan-Genome Comparison**" tab.
Click "**Choose File**" and select your FASTA file.
Optionally fill in organism name, strain, and description.
Click "**Upload Genome**" and wait for the confirmation message.

 **Supported formats:** .fasta, .fa, .fna — Maximum file size: 200 MB

Step 2 — Select Reference Database

Step 3 — Run the Pipeline

Step 4 — View & Download Results

Supported File Formats

Input Formats

.fasta — Standard FASTA format
.fa — FASTA short form
.fna — Nucleotide FASTA

Maximum file size: 200 MB

Output Formats

.fasta — Gene sequences (Core, Accessory, Unique)
.xlsx — Excel annotations with COG categories
.png — COG classification charts
.txt — Summary statistics and mapping logs

Tutorial Guide (PDF)

Download the comprehensive tutorial with screenshots and step-by-step walkthroughs for every feature.

 **Download PDF**

Troubleshooting

Section V

How to use PanWSGA?

User can access the tool by both Dashboard and tools section

Genomic Calculation Toolkit

Run isolated, advanced genomic calculations and comparative workflows with publication-ready results.



Unique Gene Finder

Primary workflow (upload → compare → annotate)

Runs the core pipeline that classifies genes into core, accessory, and unique buckets using the selected pan-genome reference. Outputs FASTA files, summaries, and download bundles.

Same Celery comparison job exposed on the dashboard workspace.

Stores every run under your account for later downloads.

Works with all registered database pan-genomes.

[Open Workspace](#)

Click here

Pipeline Checklist

- ☐ **Include annotation (eggNOG-mapper)**
Runs functional annotations after gene classifications.
- ☐ **Visualise annotation results**
Generates interactive GO/COG chart visualizations.

ⓘ Annotation is skipped (only FASTA outputs).
Visualization remains disabled.

Upload Genome FASTA

Upload your genome sequence file in FASTA format to run comparisons against reference databases.

Input File Type



CDS (Coding Sequences)

Pre-predicted coding sequences (FASTA format containing genes/proteins).



Genome Assembly

Recommended

Raw draft assembly FASTA file. Will automatically execute Prokka for gene calling.

FASTA File



Drag & drop your FASTA file here, or click to browse
Supports .fasta, .fa, .fna (Max: 20MB)



Upload Sequence

Organism Name (optional)

e.g., Acinetobacter baumannii

Strain (optional)

e.g., AB5075

Description (optional)

e.g., Clinical isolate

Select Species and Pan-Genome

Choose the species and pan-genome reference database for comparative analysis. Select a species first to see available pan-genomes.

Species Selection

Type or select species...



Reference Pan-Genome

-- Select a species first --



Selecting a species automatically loads its pan-genome references.

Upload Query
genome here

Upload it

Write query
genome specie
name here

Reference pan genome of that
species will be selected here
automatically

Run Comparison

Start the pan-genome comparison using your query genome and selected database. This compares your sequence against the references to identify core, accessory, and unique genes.

Default
parameters
recommended

 **Prokka Parameters (Genome Assembly Annotation)**



 **Unique Gene Finder Parameters**



Optional Processing Pipelines



Run eggNOG-mapper Annotation

Perform functional annotation of identified genes against orthology databases.



Generate COG/GO Classification Charts

Visualize functional profiles (requires eggNOG-mapper enabled).

Run eggNOG-mapper on output of:



Stage 2 (Nucleotide Validation)



Stage 3 (Protein Validation)

Recommended

Select Chart Visualizations to Generate



Pie



Vertical Bar



Horizontal Bar



Radar



Heatmap



Start Comparison Pipeline



Secure sandboxed biological calculation

Click here to start
analysis

Run Comparison

Start the pan-genome comparison using your query genome and selected database. This compares your sequence against the references to identify core, accessory, and unique genes.

Prokka Parameters (Genome Assembly Annotation)

CPUs

4

Gene Calling Mode

Prodigal Single



Locus Tag Prefix (optional)

e.g., ABAU



Enable Min Contig Length

Min Contig Length

200

Optimised for standard bacterial CDS generation: fast mode, no rRNA/tRNA search, centre X, increment 1.

Unique Gene Finder Parameters

Optional Processing Pipelines



Run eggNOG-mapper Annotation

Perform functional annotation of identified genes against orthology databases.



Generate COG/GO Classification Charts

Visualize functional profiles (requires eggNOG-mapper enabled).



Start Comparison Pipeline



Secure sandboxed biological calculation

Default parameters recommended but user can adjust according to requirement

Only gene fasta files can be obtained with out there classification or. Visual if disabled these two option



Run Comparison

Start the pan-genome comparison using your query genome and selected database. This compares your sequence against the references to identify core, accessory, and unique genes.

[Prokka Parameters \(Genome Assembly Annotation\)](#)

Unique Gene Finder Parameters

Stage 1: K-mer Filtering

Identity Threshold

K-mer Size

Min Shared K-mers

Strict Identity

Strict Coverage

Relaxed Identity

Relaxed Coverage

Fallback Min K-mers

Stage 2: Nucleotide Alignment Validation

Align Identity

Align Coverage

Align E-value

Align Threads

Fallback Identity

Fallback Coverage

Stage 3: Protein-Level Validation

Protein Identity

Protein Coverage

Protein E-value

Protein Threads

Default parameters
recommended but
user can adjust
according to
requirements

Run Comparison

Start the pan-genome comparison using your query genome and selected database. This compares your sequence against the references to identify core, accessory, and unique genes.

 **Prokka Parameters (Genome Assembly Annotation)**

 **Unique Gene Finder Parameters**

Optional Processing Pipelines

☒ Run eggNOG-mapper Annotation

Perform functional annotation of identified genes against orthology databases.

☒ Generate COG/GO Classification Charts

Visualize functional profiles (requires eggNOG-mapper enabled).

Run eggNOG-mapper on output of:

☐ Stage 2 (Nucleotide Validation) ☒ Stage 3 (Protein Validation) **Recommended**

Select Chart Visualizations to Generate

☒ Pie

☒ Vertical Bar

☒ Horizontal Bar

☒ Radar

☒ Heatmap

 **Running Comparison...**

 Secure sandboxed biological calculation

**Click here to start
analysis**

Comparative Pipeline Running

RUNNING

Job ID: `0fb47094-90a0-4cec-b810-ebf806856c12`

Elapsed Time: **1m 22s**

Track & Share Job URL:

`https://www.panwsa.org/dashboard?id=0fb`

 **Copy Link**

 These results will be automatically deleted after 7 days.

**Running Job apper
like that**

**Analysis Results**

Comparison ID: 38629bdb-e1f4-431d-884e-b77acedfb48a

Pipeline Completed Successfully!

Total Run Time: 0m 0s

Saved Comparison Result

This demo result is stored permanently. Use this URL to view or share:

View on Demo Page:<https://www.panwsga.org/demo?id=38629bdb-e1f4-431d-884e-b77acedfb48a>[Copy Link](#)**Completed job can be
accessed by using this link****Test Results****855**

Total Genes

**789**

Core Genes

**57**

Accessory Genes

**9**

Unique Genes

**Gene Distribution**



Analysis Results

TOTAL GENES
1,090

CORE GENES
692

ACCESSORY GENES
359

UNIQUE GENES
39

Gene Distribution Percentages



Gene Preview

Preview of the first 20 genes classified in each category. You can click on the gene IDs or products to query NCBI or UniProt databases.

- Core
- Accessory
- Unique

#	Gene ID	Product / Description	Length
1	MIADNNAI_00001	I core_gene:APDOBHHF_00459 identity:strict	1,194 bp
2	MIADNNAI_00002	I core_gene:APDOBHHF_00458 identity:strict	114 bp
3	MIADNNAI_00003	I core_gene:APDOBHHF_00457 identity:strict	201 bp
4	MIADNNAI_00004	I core_gene:APDOBHHF_00456 identity:strict	942 bp
5	MIADNNAI_00007	I core_gene:APDOBHHF_00453 identity:relaxed	1,080 bp

Share Analysis Results

Copy this permalink to share the results directly or access this dashboard in the future.

View on Dashboard:

<https://www.panwsga.org/dashboard?id=f4d99df2-c538-4148-90f7-beadc4b8803e>

Copy Link

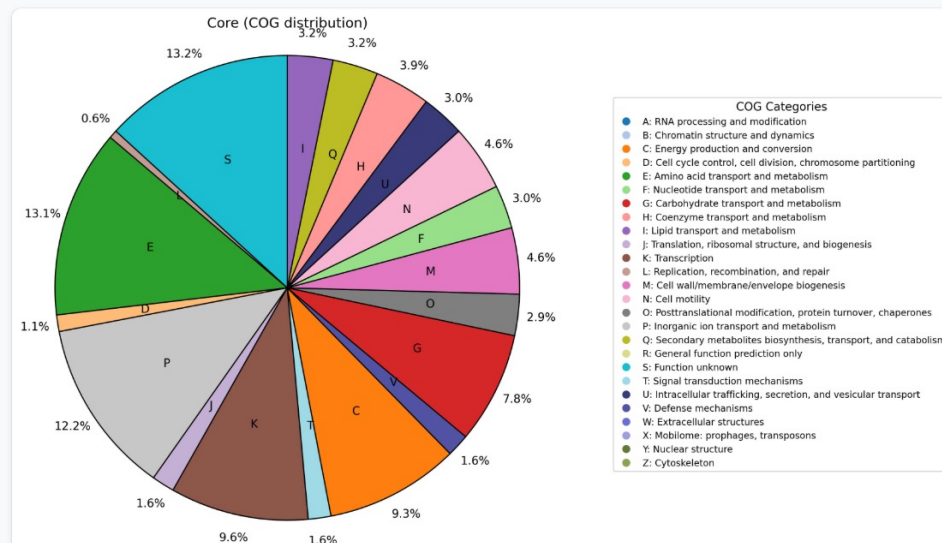
Note: These results will be automatically deleted after 7 days.

Completed job can be
accessed by using this link

Functional Classification Chart (COG)

Select Classification Category:

Pie Chart - Core



Download Graphic Chart



Gene Preview

Core (1,999) Accessory (801) Unique (46)

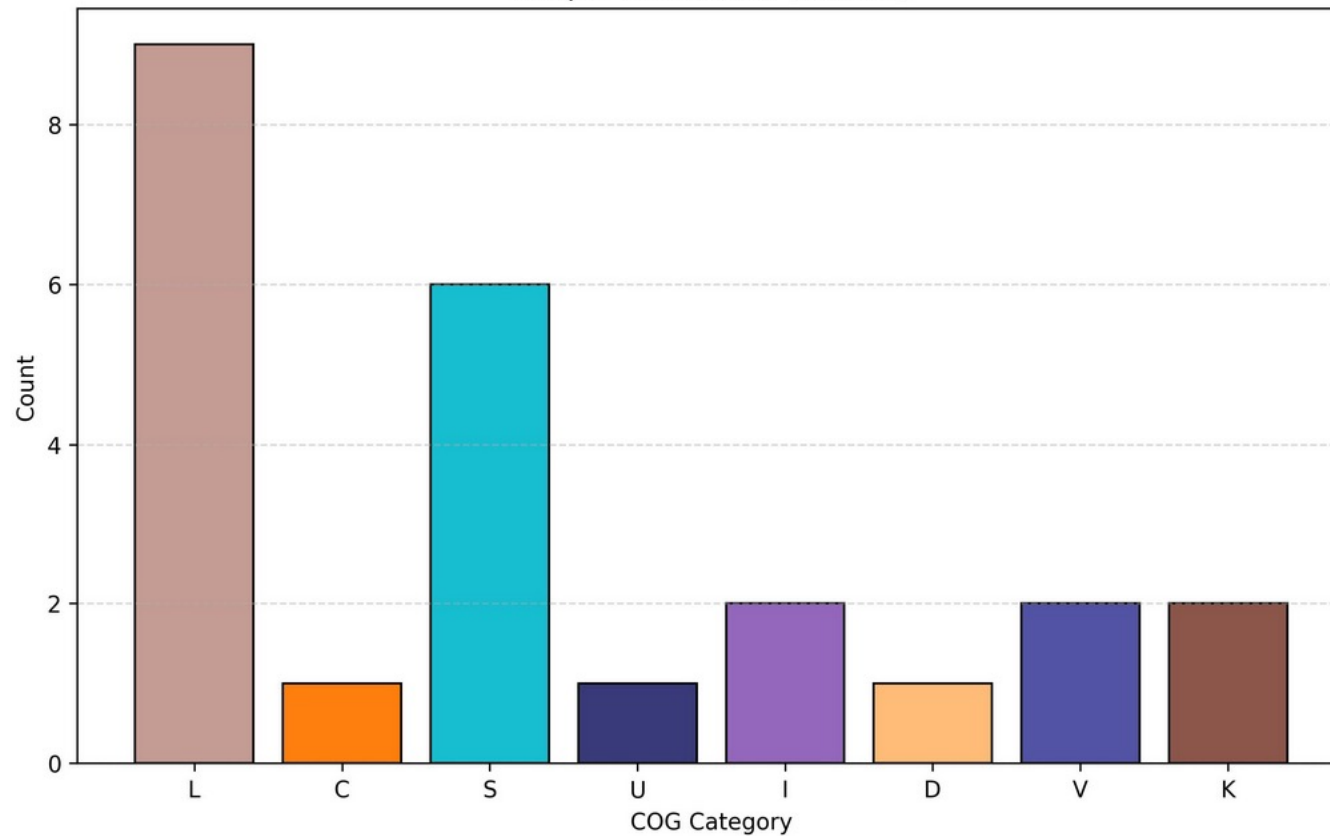
GO Classification Chart

Select Chart Type:

Vertical Bar Chart - Unique

Unique (COG distribution)

Unique COG counts (vertical)



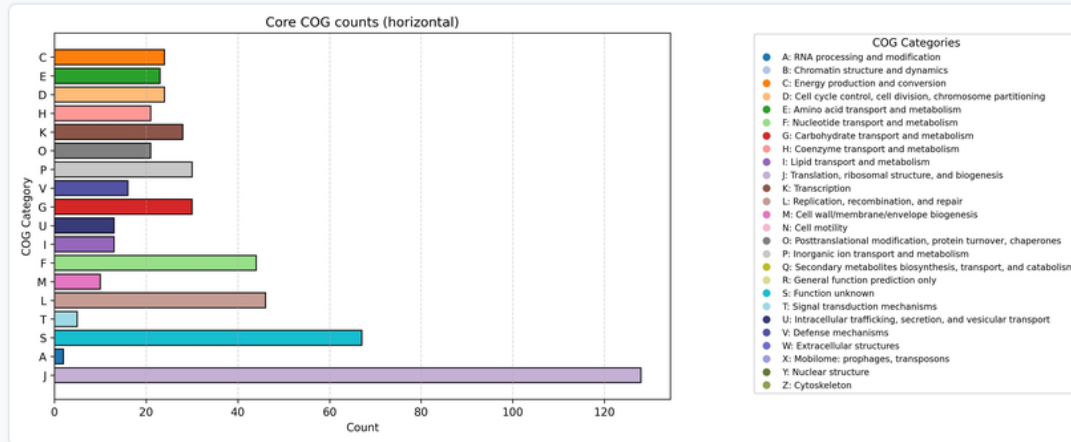
COG Categories

- A: RNA processing and modification
- B: Chromatin structure and dynamics
- C: Energy production and conversion
- D: Cell cycle control, cell division, chromosome partitioning
- E: Amino acid transport and metabolism
- F: Nucleotide transport and metabolism
- G: Carbohydrate transport and metabolism
- H: Coenzyme transport and metabolism
- I: Lipid transport and metabolism
- J: Translation, ribosomal structure, and biogenesis
- K: Transcription
- L: Replication, recombination, and repair
- M: Cell wall/membrane/envelope biogenesis
- N: Cell motility
- O: Posttranslational modification, protein turnover, chaperones
- P: Inorganic ion transport and metabolism
- Q: Secondary metabolites biosynthesis, transport, and catabolism
- R: General function prediction only
- S: Function unknown
- T: Signal transduction mechanisms
- U: Intracellular trafficking, secretion, and vesicular transport
- V: Defense mechanisms
- W: Extracellular structures
- X: Mobilome: prophages, transposons
- Y: Nuclear structure
- Z: Cytoskeleton

COG Classification Chart

Select Chart Type:

Horizontal Bar Chart - Core



Download Chart

Downloaded from here

Download Results

FASTA Sequences:

Initial Results (Stage 1: K-mer Filtering)

Core Genes Accessory Genes Unique Genes

Refined Results (Stage 2: Nucleotide Alignment)

Core Genes Accessory Genes Unique Genes

Final Results (Stage 3: Protein Validation)

Core Genes Accessory Genes Unique Genes

Annotations (Excel):

Core Annotations Accessory Annotations Unique Annotations

All results can be downloaded from here

Export Results

FASTA Sequences (Genes & Proteins)

Initial Results (Stage 1: K-mer Filtering)

Core Genes FASTA

Contains all highly conserved core genes found across target genomes.

[Download FASTA](#)

Accessory Genes FASTA

Contains accessory genes shared by some but not all genomes.

[Download FASTA](#)

Unique Genes FASTA

Contains genes completely unique to your query genome assembly.

[Download FASTA](#)

Refined Results (Stage 2: Nucleotide Alignment)

Core Genes FASTA

Contains all highly conserved core genes found across target genomes.

[Download FASTA](#)

Accessory Genes FASTA

Contains accessory genes shared by some but not all genomes.

[Download FASTA](#)

Unique Genes FASTA

Contains genes completely unique to your query genome assembly.

[Download FASTA](#)

Final Results (Stage 3: Protein Validation)

Core Genes FASTA

Contains all highly conserved core genes found across target genomes.

[Download FASTA](#)

Accessory Genes FASTA

Contains accessory genes shared by some but not all genomes.

[Download FASTA](#)

Unique Genes FASTA

Contains genes completely unique to your query genome assembly.

[Download FASTA](#)

Functional Annotations (Excel Spreadsheets)

Core Annotations

Functional descriptions, GO terms, and COG categories for core genes.

[Download Excel](#)

Accessory Annotations

Functional descriptions, GO terms, and COG categories for accessory genes.

[Download Excel](#)

Unique Annotations

Functional descriptions, GO terms, and COG categories for unique genes.

[Download Excel](#)

Job Configuration & Log

Job Parameters & Log

Download a structured text report containing all parameters and the job execution log.

[Download Parameters & Log](#)

Parameters and log file for the job run can be downloaded from here