Microbial Genomics and Metagenomics

From data to discovery

The role of bioinformatics in Microbiology

Microbiome profiling is used to detect microbes in metagenomic samples, and to explore links between taxonomic or functional profiles and external factors like patient health or plant yield.

Microbial typing of isolates at the whole genome level is useful to characterize pathogens, or to provide quality control for valuable strains.

Outbreak analysis leverages whole genome information for pathogen typing and epidemiological outbreak investigation.

Sample to Insight

Common challenges

A lack of integrated analytics leaves organizations with the burden of integrating and maintaining all the bioinformatics-, statistics- and visualization tools required for their microbial research.

High sample volumes are common thanks to the throughput capacity of modern NGS instruments. But sample load introduces a computational burden and makes managing data and metadata challenging.

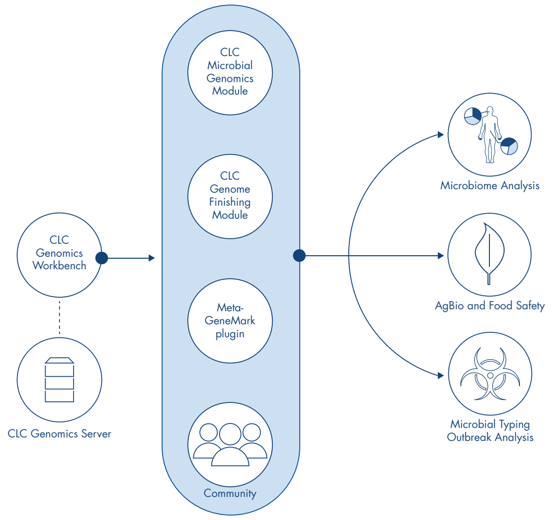
QIAGEN Microbial Genomics Pro Suite

Integrated analytics deliver research continuity

All analytics for microbial genomics and metagenomics come fully integrate into one scalable and enterprise-ready solution. Data and sample-metadata management is included.

Focus on what matters

We shift the focus from data to discovery with preconfigured workflows that streamline the processing of raw NGS data. Researchers can focus on the interpretation of results in the context of other samples and metadata.



High performance algorithms

Designed to save time and compute resources, and to be accessible to bioinformatics experts and non-bioinformaticians alike.

Microbial genomics powered by quality components

The QIAGEN Microbial Genomics Pro Suite is built on CLC Genomics Workbench, the industry standard platform for bioinformatics computing. Plugins and Modules add a layer of specialized tools and workflows to CLC Genomics Workbench, creating a comprehensive solution for microbial genomics and metagenomics data analysis.
There is more: Experts in the field of microbial genomics are using the QIAGEN Microbial Genomics Pro Suite to bring additional tools and services to the user community.

Features

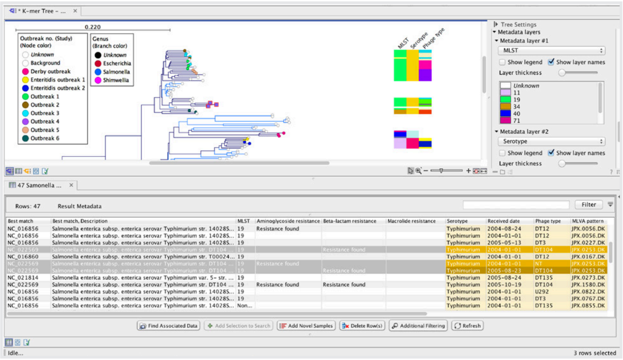
Pathogen typing and Outbreak analysis

Infectious diseases impact public health globally and come with a personal burden for patients as well as with a high economical burden. Detecting and controlling pathogen outbreaks early is key to protecting public health. Whole genome analysis offers unique advantages when used for typing and characterizing microbial isolates.

Isolate management – Samples, metadata and results can be accessed from a convenient Analysis Dashboard.

Typing – Preconfigured but customizable workflows streamline typing accepting whole genome isolate data as input. In one analysis workflow you can confirm the identity and characteristics of microbes, revealing:

- NGS-MLST (Multi locus sequence typing)
- Taxonomy
- Closest known reference genome
- Antimicrobial resistance genes



Rows: 47 Result Metadata

Best match	Best match, Description	MLST	Aminoglycoside resistance	Beta-lactam resistance	Macrolide resistance	Serotype	Received date	Phage type	MLVA pattern
NC_016856	Salmonella enterica subsp. enterica serovar Typhimurium str. 14028S...	19	Resistance found			Typhimurium	2004-08-24	DT12	JPX.0056.DK
NC_016856	Salmonella enterica subsp. enterica serovar Typhimurium str. 14028S...	19				Typhimurium	2004-01-01	DT12	JPX.0056.DK
NC_016856	Salmonella enterica subsp. enterica serovar Typhimurium str. 14028S...	19				Typhimurium	2005-05-13	DT3	JPX.0227.DK
NC_022569	Salmonella enterica subsp. enterica serovar Typhimurium str. DT104 ...	19		Resistance found		Typhimurium	2004-01-01	DT104	JPX.0253.DK
NC_016860	Salmonella enterica subsp. enterica serovar Typhimurium str. T00024...	19				Typhimurium	2004-01-01	DT12	JPX.0167.DK
NC_022569	Salmonella enterica subsp. enterica serovar Typhimurium str. DT104 ...	19	Resistance found			Typhimurium	2004-01-01	NT	JPX.0253.DK
NC_022569	Salmonella enterica subsp. enterica serovar Typhimurium str. DT104 ...	19		Resistance found		Typhimurium	2005-08-23	DT104	JPX.0253.DK
NC_021814	Salmonella enterica subsp. enterica serovar Typhimurium var. 5- str. ...	19				Typhimurium	2005-08-24	DT135	JPX.0273.DK
NC_022569	Salmonella enterica subsp. enterica serovar Typhimurium str. DT104 ...	19	Resistance found	Resistance found		Typhimurium	2005-10-19	DT104	JPX.1580.DK
NC_016856	Salmonella enterica subsp. enterica serovar Typhimurium str. 14028S...	19				Typhimurium	2004-01-01	U292	JPX.0822.DK
NC_016856	Salmonella enterica subsp. enterica serovar Typhimurium str. 14028S...	19				Typhimurium	2004-01-01	DT3	JPX.0767.DK
NC_016856	Salmonella enterica subsp. enterica serovar Typhimurium str. 14028S...	Non...				Typhimurium	2004-01-01	DT135	JPX.0855.DK

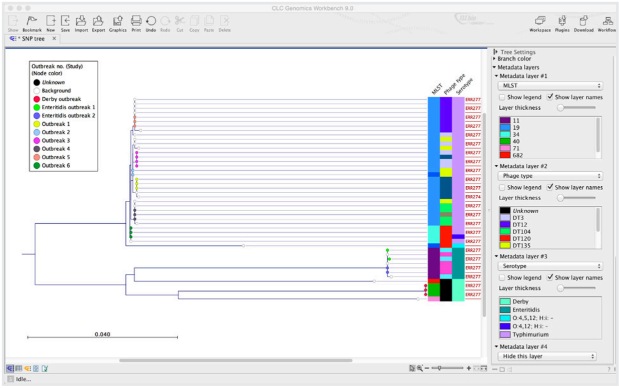
Idle... 3 rows selected

Phylogenetic analysis – Preconfigured workflows streamline phylogenetic analysis:

- K-mer based trees from raw NGS reads, assembled contigs, and closed genomes.
- Whole-genome SNP trees for accurate outbreak investigation and source tracking of pathogen isolates.

Outbreak trees can help monitor outbreaks and are essential for source tracking of pathogens. Useful features:

- Trees are interactively linked to Analysis Dashboards.
- Trees can be explored in the context of typing results, epidemiological metadata, or other sample information.



Assembly of microbial reference genomes

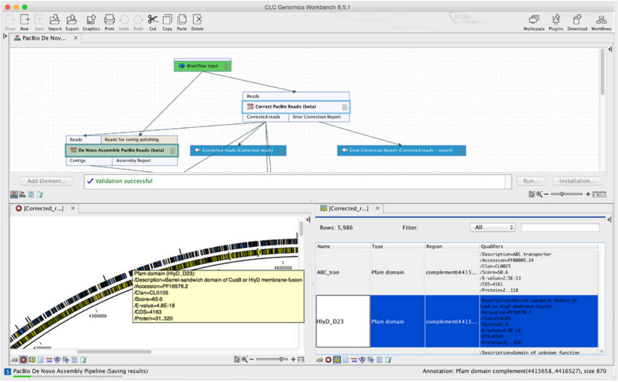
- Short read and hybrid assembly
- Scaffolding, contig joining and genome finishing

PacBio *de novo* assembly made fast and easy

- Raw PacBio reads are error-corrected and *de novo* assembled using a novel approach that can quickly convert reads to high quality reference genomes.

Genome Annotation

- The MetaGeneMark plugins carries out gene finding.
- Annotate gene models with Pfam- and GO information, or with top BLAST hits.



✓ Validation successful

Pfam domain (HlyD_D23):
/Description=Barrel-sandwich domain of CusB or HlyD membrane-fusion
/Accession=PF16576.2
/Clan=CL0105
/Score=65.0
/E-value=4.8E-18
/CDS=4163
/Protein=31..320

Rows: 5,986

Name	Type	Region	Qualifiers
ABC_tran	Pfam domain	complement(4415...	/Description=ABC transporter /Accession=PF00005.24 /Clan=CL0023 /Score=50.6 /E-value=2.5E-13 /CDS=4161 /Protein=2..118
HlyD_D23	Pfam domain	complement(4415...	/Description=Barrel-sandwich domain of CusB or HlyD membrane-fusion /Accession=PF16576.2 /Clan=CL0105 /Score=65.0 /E-value=4.8E-18 /CDS=4163 /Protein=31..320
			/Description=Domain of unknown function

PacBio De Novo Assembly Pipeline (Saving results)

Annotation: Pfam domain complement(4415658..4416527), size 870

Microbiome Profiling

All animals and plants live in close association with microbial communities. Our microbiome regulates our immune system and metabolism, it offers protection against pathogenic microbes, or produces essential vitamins.

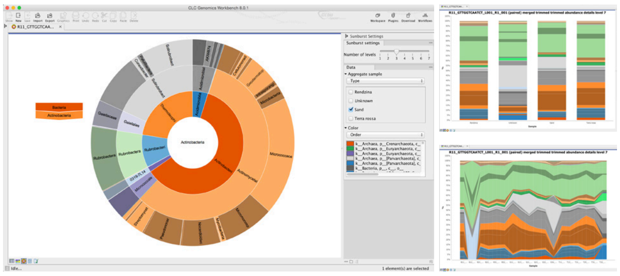
Explore the taxonomic and functional profiles of microbiomes, and associate microbiome profiles with factors like:

- Health and diseases
- Changing yields of agricultural crops or livestock
- The emergence of public health threats

Taxonomic Profiling of Microbiomes

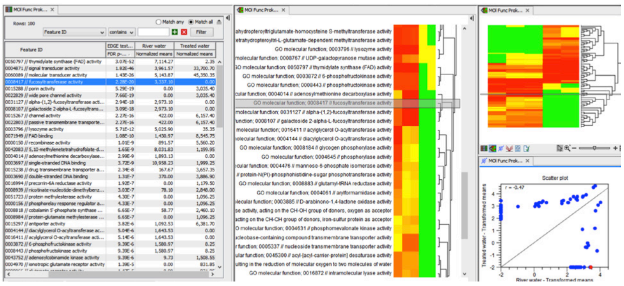
- Analysis of 16S rRNA or other amplicon data.
- *De novo*- or reference based OTU-clustering.

- Support for common taxonomy databases Greengenes, Silva, and UNITE, as well as custom databases.
- Compute relative abundance of microbes in a microbial community or study the differential abundance between different microbiomes. Data can be explored at different taxonomic levels and in the context of metadata.
- From raw reads to interactive visualization of results in four steps through preconfigured but customizable workflows.



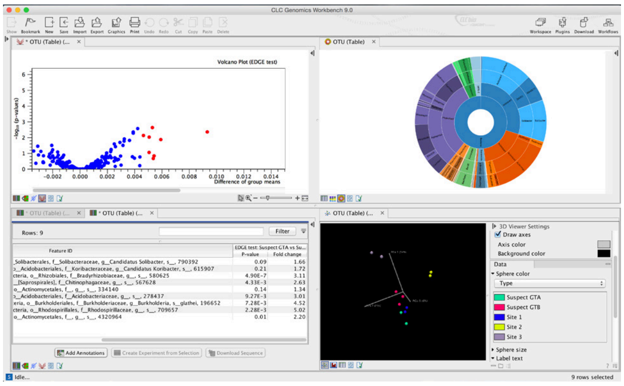
Functional Profiling of Microbiomes

- Assembly and annotation of whole metagenome data sets
- Functional profiling using PFAM, Gene Ontology and BLAST
- Stacked bar charts, area charts and zoomable sunburst diagrams to explore and compare the functional composition of metagenomic samples, or sample groups.



Statistical Analysis

- Principal Coordinate Analysis (PCoA) to cluster samples of similar functional composition.
- PERMANOVA analysis to assess statistical support for clustering.
- Testing for differential abundance of taxonomy or gene function;
  - Measure normalized fold change.
  - Estimate statistical significance.
  - Explore correlations between sample categories.
- Statistical testing and visualization in the context of the sample metadata.



Rows: 9

Feature ID	EDGE test: Suspect GTA vs Su... P-value	Fold change
Solibacterales, f__Solibacteraceae, g__Candidatus Solibacter, s__, 790392	0.09	1.66
o__Acidobacteriales, f__Koribacteraceae, g__Candidatus Koribacter, s__, 615907	0.21	1.72
cteria, o__Rhizobiales, f__Bradyrhizobiaceae, g__, s__, 580625	4.90E-7	3.11
__[Saprospirales], f__Chitinophagaceae, g__, s__, 567628	4.33E-3	2.63
o__Actinomycetales, f__, g__, s__, 334140	0.14	1.34
o__Acidobacteriales, f__Acidobacteriaceae, g__, s__, 278437	9.27E-3	3.01
eria, o__Burkholderiales, f__Burkholderiaceae, g__Burkholderia, s__glathei, 196652	7.28E-3	4.52
cteria, o__Rhodospirillales, f__Rhodospirillaceae, g__, s__, 709657	2.28E-3	5.02
o__Actinomycetales, f__, g__, s__, 4320964	0.01	2.20

Get your trial at **www.qiagenbioinformatics.com**

QIAGEN Bioinformatics

EMEA
Silkeborgvej 2 · Prismet
8000 Aarhus C
Denmark
Phone: +45 7022 5509

Americas
1700 Seaport Boulevard #3
Redwood City · CA 94063
USA
Phone: +1 (617) 945 0178

Sample to Insight

QIAGEN