

IRIDA: A Genomic Epidemiology Platform Built on top of Galaxy

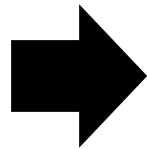
Aaron Petkau
Public Health Agency of Canada
Presented at GCC2015
Norwich, UK

Background

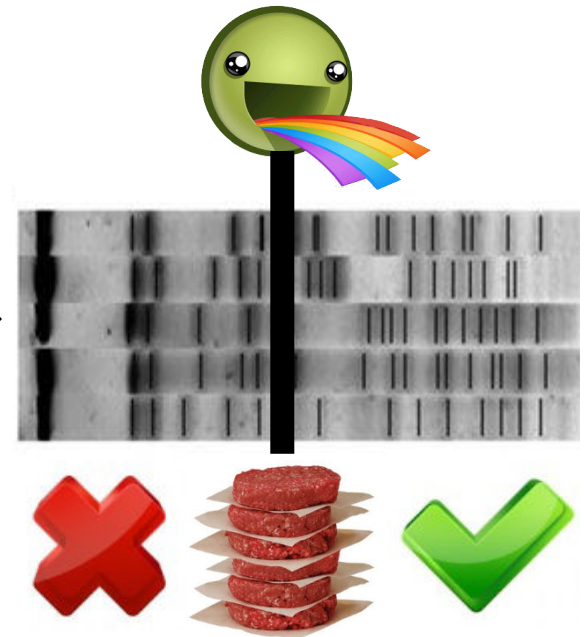
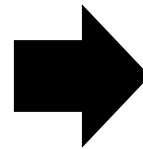
- The Public Health Agency of Canada is involved in foodborne disease outbreak surveillance and response
- Uses PFGE fingerprinting



Sample Source



Culture

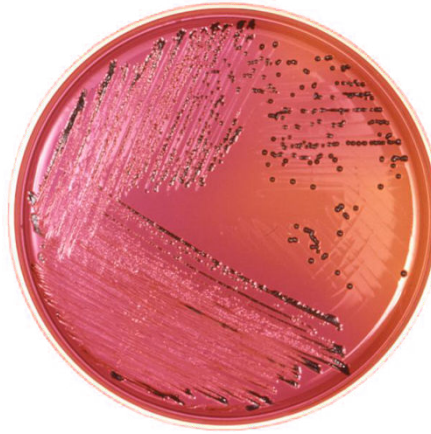
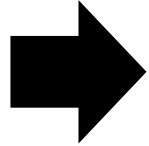


PFGE Fingerprint

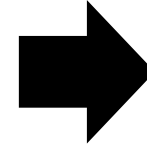
Whole Genome Sequencing



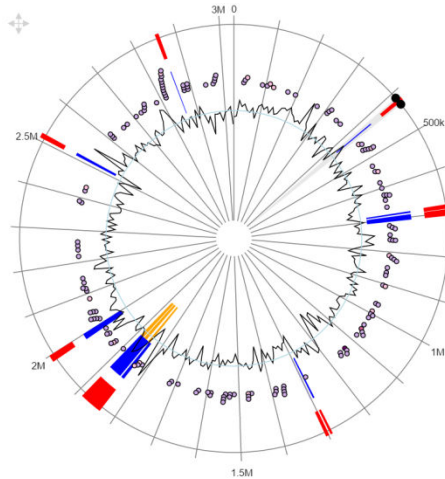
Sample Source



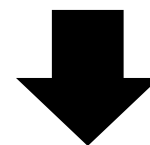
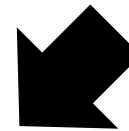
Culture



Sequence



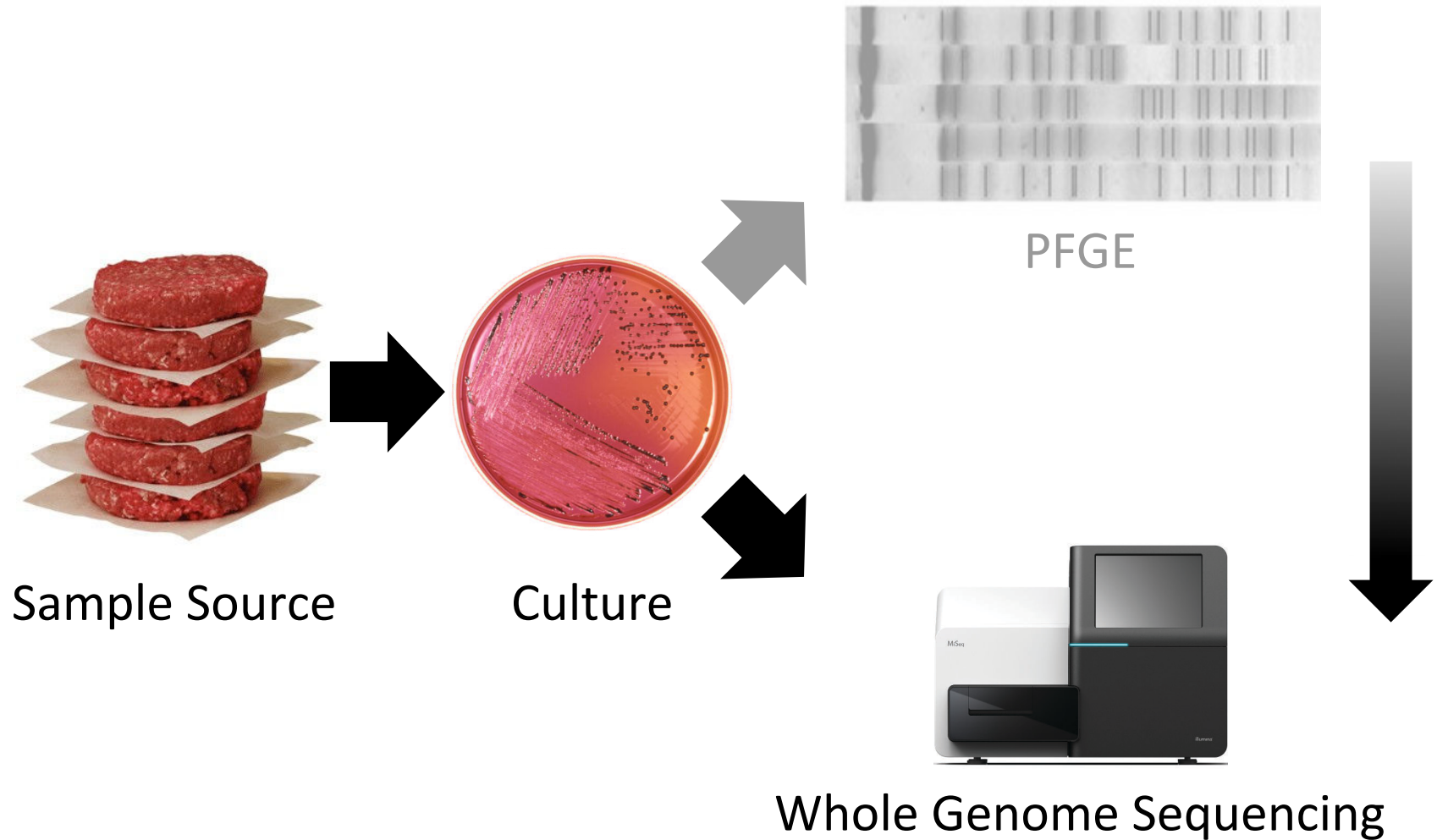
Additional Information



ATCATTTC
CGGAATG
GCCGAAT

SNV Fingerprint

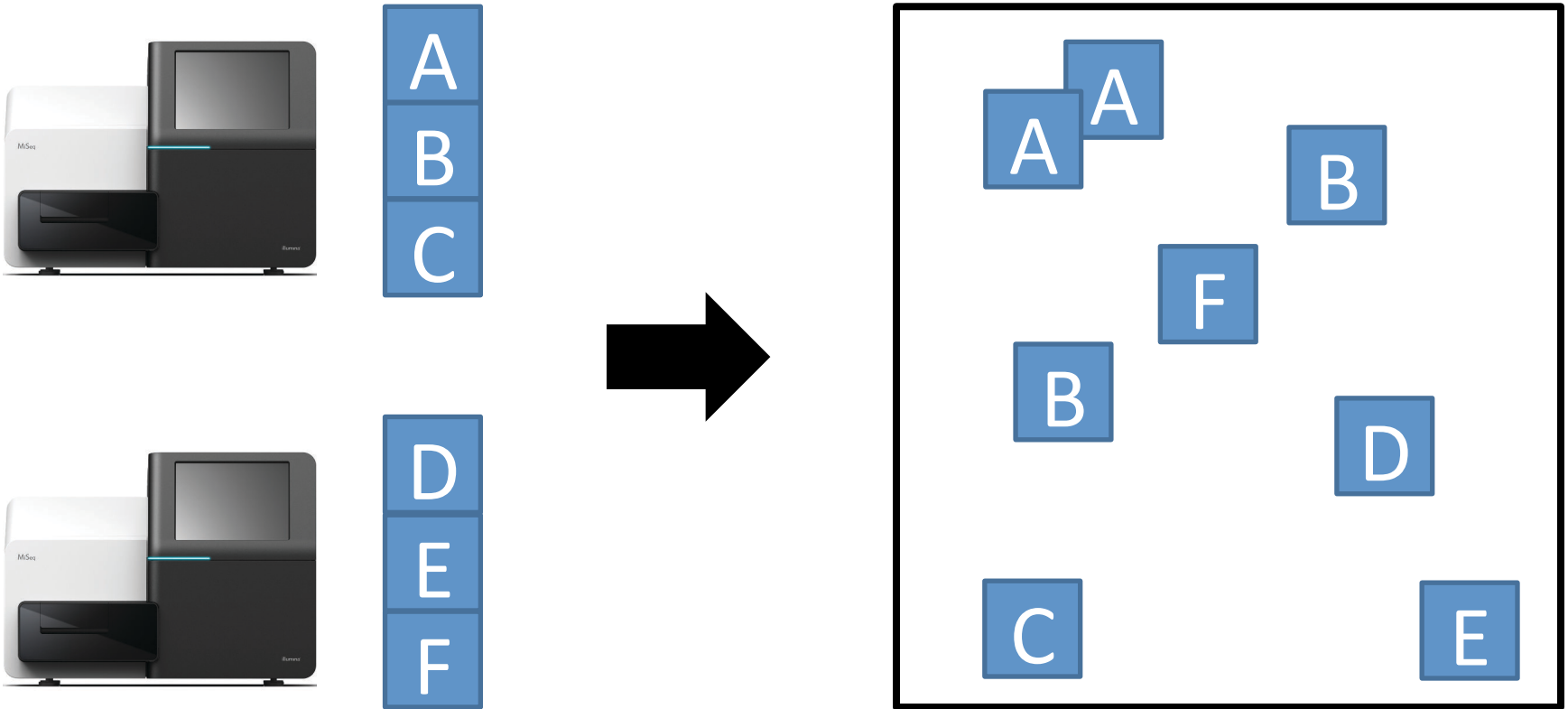
Future



Issues

(Computational and Bioinformatics)

Data Management

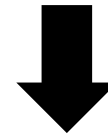


Data Analysis

- Lack of standards
- Constantly changing software
- Quality control and assessment
- Requires heavy compute resources



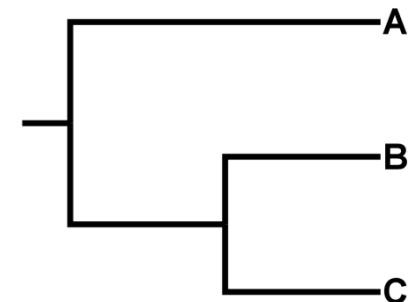
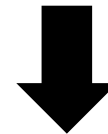
A B C



A: AATC

B: TTGA

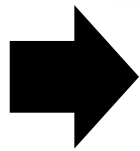
C: TTCT



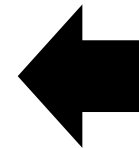
Data Sharing and Security



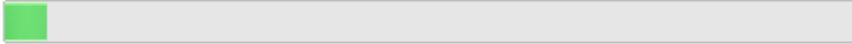
ERR632096.sra 174 MB [Downloads](#)
ftp-trace.ncbi.nlm.nih.gov 2% at 519 KB/sec
5 min 34 sec remaining



Download, Import,
Export, Upload



ERR632096.sra



2 minutes, 34 seconds remaining — 10.3 of 174 MB



Public

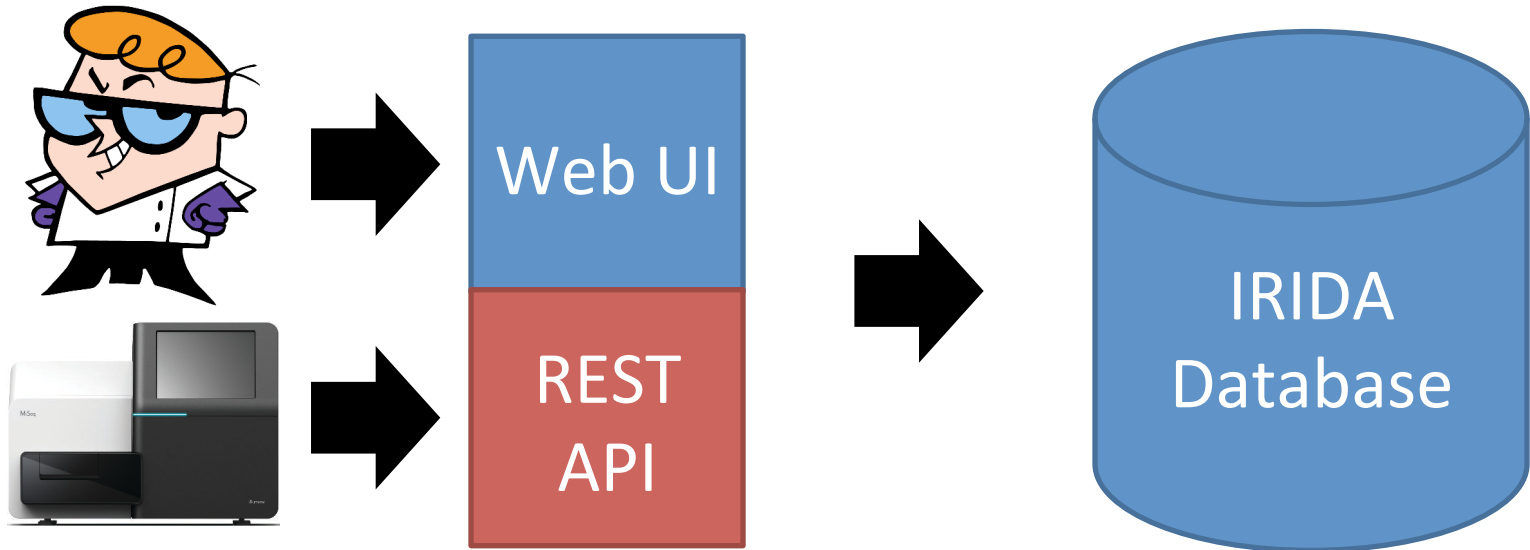
Private



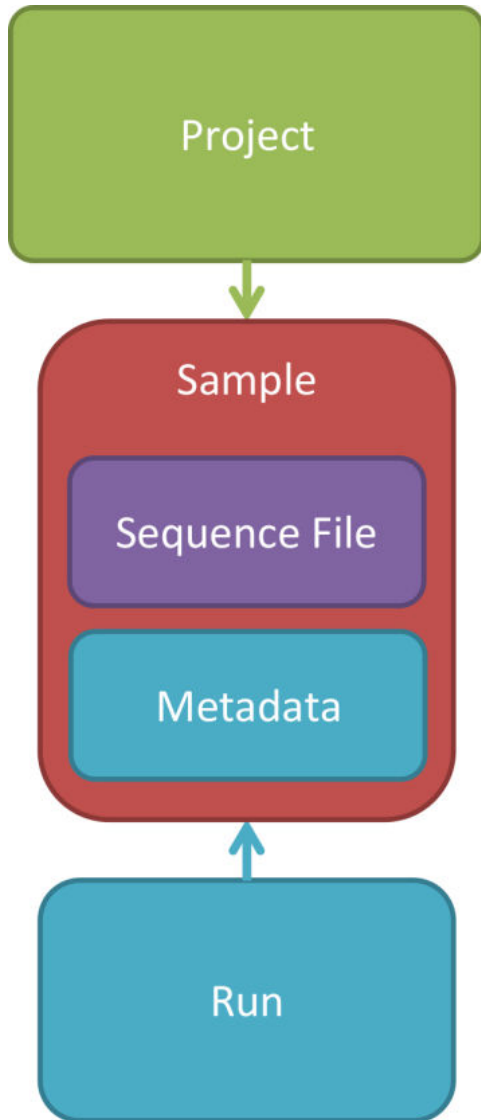
IRIDA

Integrated Rapid Infectious Disease
Analysis

Data Management



Data Management



Listeria Example Samples

1

Select ▾

Display ▾

Samples ▾

Export ▾

 Add to Cart

Name

◆

Organism

◆

Collected By

◆

Project

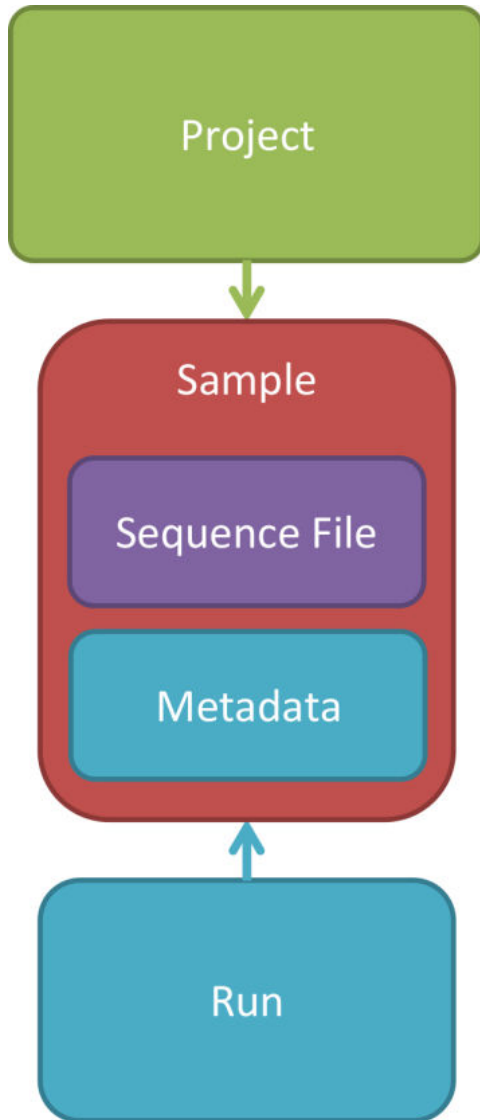


08-5578



08-5923

Data Management



IRIDA Platform Projects Analysis Users

Projects all

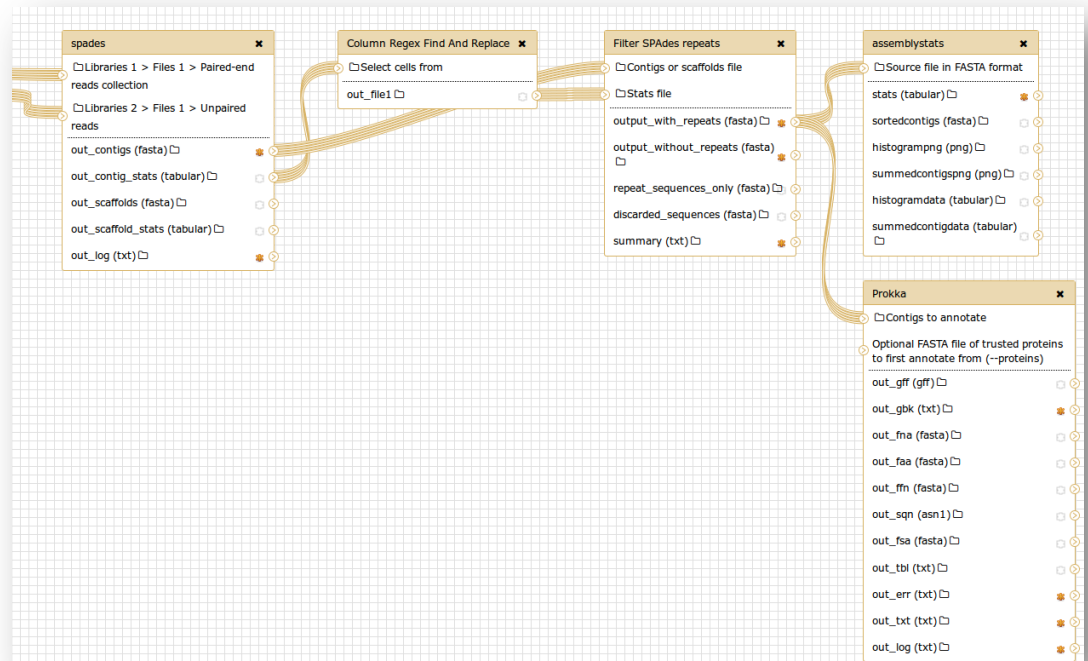
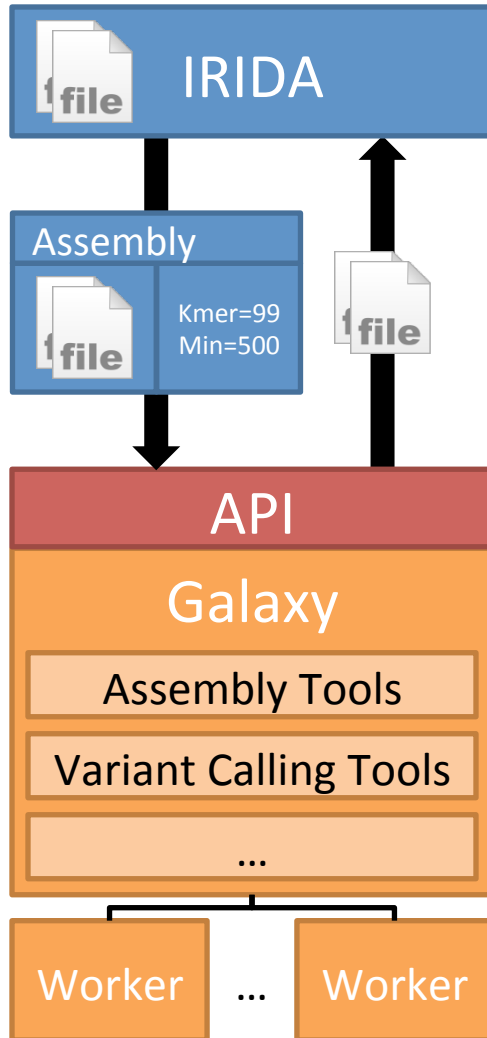
Projects

Show 10 entries

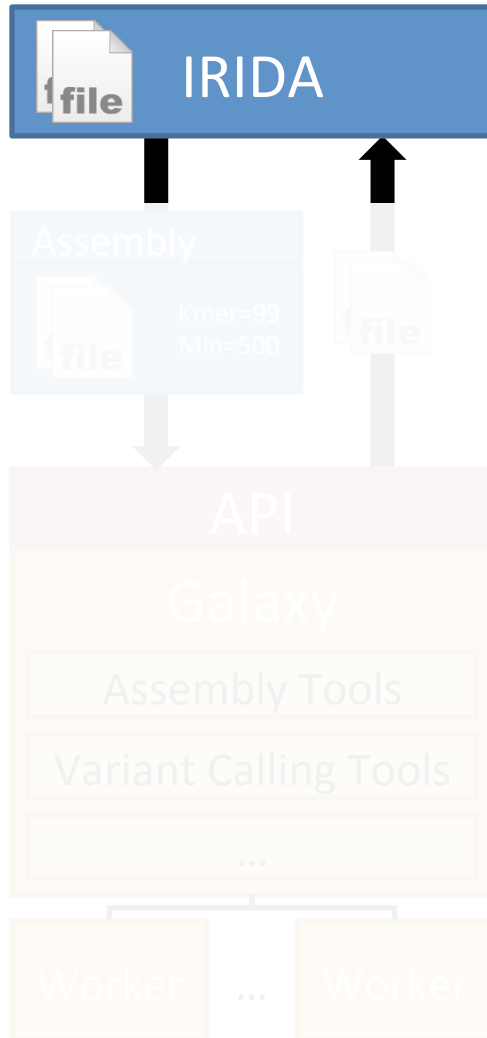
☐	ID	Name	Organism	Samples
☐	1	Listeria Example	<i>Listeria monocytogenes</i>	3
☐	2	E Coli Example	<i>Escherichia coli</i>	1
☐	3	Other Example		0

Showing 1 to 3 of 3 entries

Data Analysis



Sample Selection



Listeria Example Samples

1 Select ▾ Display ▾ Samples ▾ Export ▾ [Add to Cart](#)

	Name	Organism	Collected By	Project
☒	08-5578			
☐	08-5923			

Search Cart

- ▼ Listeria Example 2 / 2 ×
 - 08-5923 ×
 - 08-5578 ×
- ▼ E Coli Example 1 / 1 ×
 - 11-1111 ×

[Select a Pipeline](#)

Pipeline Selection



Select a Pipeline

Your cart contains 1 project with 3 samples.

ASSEMBLY AND ANNOTATION PIPELINE

Generate an assembled and annotated genome from the reads within a sample using FLASH, SPAdes, and Prokka. Outputs analyzed and produced separately for each sample include: log files, assembly statistics, the contigs (all contigs, filtered contigs with repeats, filtered contigs without repeats), and annotations from Prokka.

Select

ASSEMBLY AND ANNOTATION COLLECTION PIPELINE

The assembly and annotation collection pipeline provides the same results as the assembly and annotation pipeline, but all samples are analyzed together which allows you to download a single package for all samples submitted.

Select

PHYLOGENOMICS PIPELINE

Generate a Whole Genome Phylogeny from a set of samples and a reference genome based on Single Nucleotide Polymorphisms (SNVs). This will provide a dendrogram as well as a table of all SNVs used and a SNV distance matrix between each sample.

Select

Assembly and Annotation

Ready to Launch?

Pipeline Name AssemblyAnnotation_20150618

Parameters Default Parameters

Customize

Files Listeria Example / 08-5578

Remove

08-5578_S1_L001_R1_001.fastq

65.9 MB
File Size

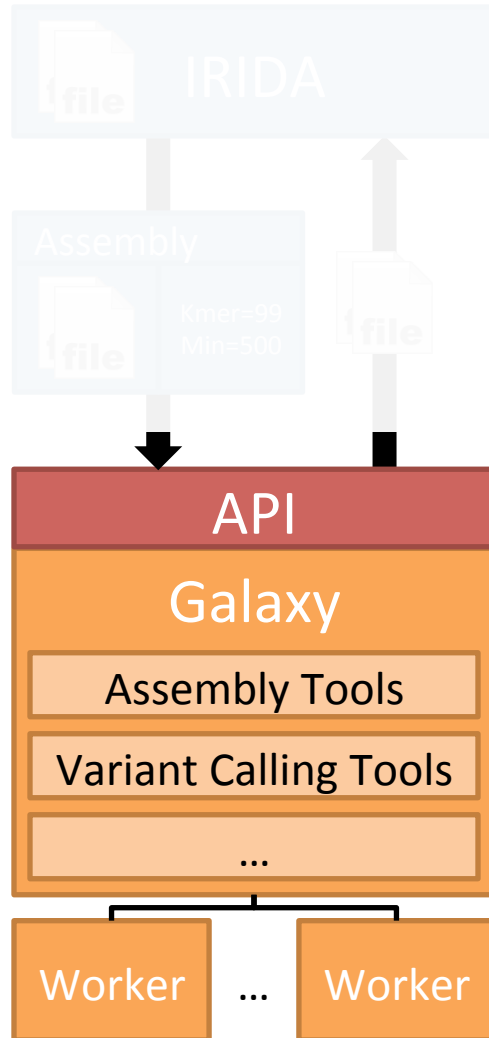
15 Jun 2015
Date Created

08-5578_S1_L001_R2_001.fastq

65.9 MB
File Size

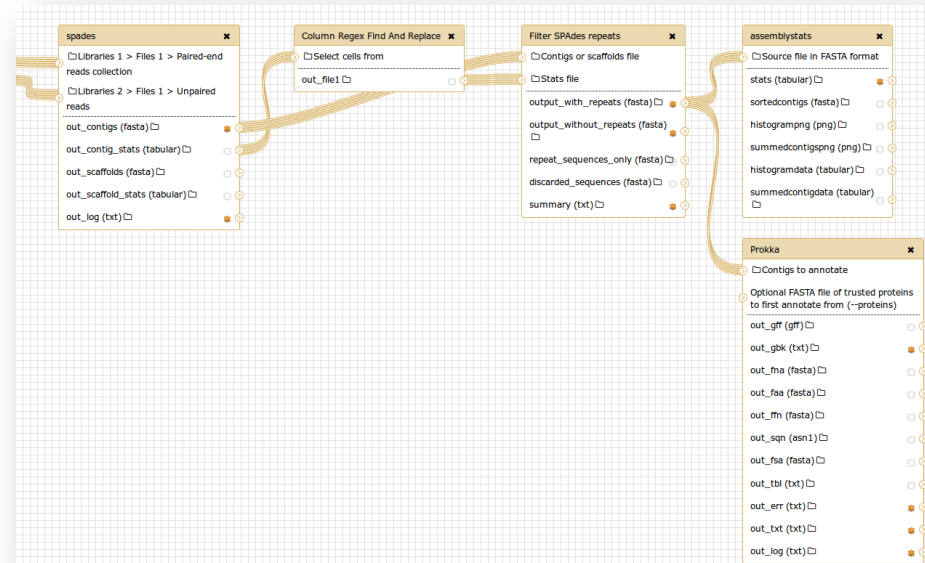
15 Jun 2015
Date Created

Analysis Execution

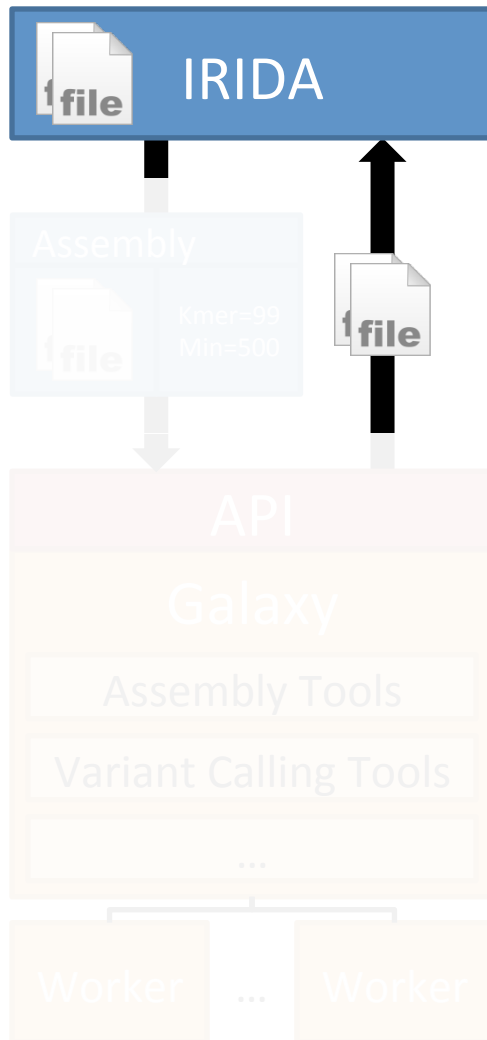


Your Analyses

State	Name	Type	Date Created
Running 25%	AssemblyAnnotation_20150630_08-55 78	Assembly and Annotation Pipeline	Jun 30, 2015



Analysis Results



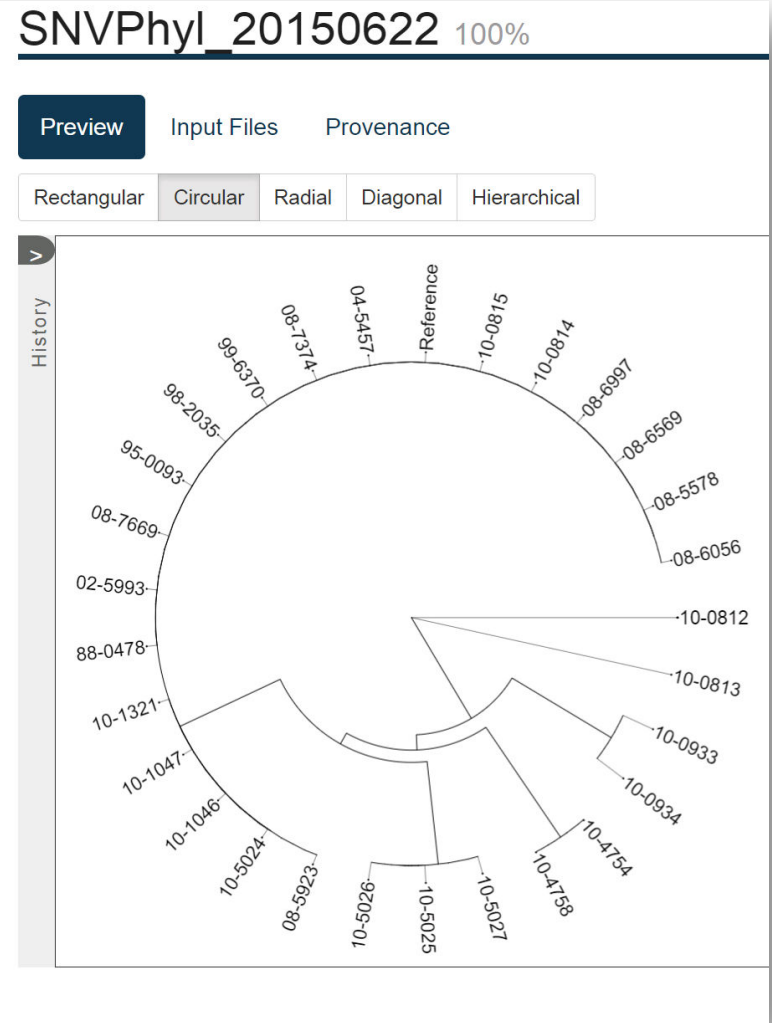
Analysis Details

ID	76
Type	Phylogenomics Pipeline
State	Completed
Date Created	22 Jun 2015

Output Files

- phylogeneticTree.tre
- snpMatrix.tsv
- snpTable.tsv

[Download Files](#)



Data Sharing

- IRIDA Instances



- External tools



Galaxy Exporter

The screenshot displays the Galaxy NML interface for the IRIDA Platform. The main content area shows a table of '2010Listeria_resequence Samples' with columns for Name, Project, and Created On. A dropdown menu is open for the 'Export' button, with 'Send to Galaxy' highlighted by a red box. The 'Tools' sidebar on the left has 'Get Data' highlighted with a red box. The 'History' sidebar on the right has '1: IRIDA Export' highlighted with a red box.

Galaxy / NML | Analyze Data | Workflow | Shared Data | Visualization | Admin | Help | User

Tools

- search tools
- Get Data**
 - IRIDA server
 - Upload File from your computer
- Text Manipulation
- Convert Formats
- Filter and Sort
- Join, Subtract and Group
- Statistics
- Experimental
- NGS: QC and manipulation
- NGS: Assembler
- NGS: VCFTools
- NGS: SAMTools
- NGS: In silico MLST
- NGS: RNA-seq
- NGS: Mapping

IRIDA Platform

Projects 7

2010Listeria_resequence Samples

3 Select Export Add to Cart

Name	Project	Created On
☒ 08-5578		Nov 25, 2013
☒ 08-5923		Nov 25, 2013
☒ 08-7669		Nov 25, 2013
☐ 08-7374		Nov 25, 2013

Download
Command-line Linker
Send to Galaxy

History

search datasets

Unnamed history
1 shown, 1 hidden
5.1 KB

1: IRIDA Export

Release Schedule

- Current
 - IRIDA instance deployed and in use locally
 - Deploying instances to external collaborators
- Future
 - Public release, available as free and open source software
 - Support and addition of new pipelines

More Information

- <http://irida.ca>
- IRIDA-mail@sfu.ca
- Poster P10 -
[IRIDA: A Genomic Epidemiology Platform Built on top of Galaxy](#)
- Poster P12 -
[Bioinformatics Evolving at Canada's National Microbiology Laboratory](#)

Acknowledgements



Public Health
Agency of Canada

Agence de la santé
publique du Canada

PHSA Laboratories

Public Health Microbiology & Reference Laboratory



McMaster University

Andrew McArthur
Daim Sardar

European Nucleotide Archive

Guy Cochrane
Petra ten Hoopen
Clara Amid

European Food Safety Agency

Leibana Criado Ernesto
Vernazza Francesco
Rizzi Valentina

Project Leaders

Fiona Brinkman – SFU
Will Hsiao – PHMRL
Gary Van Domselaar – NML

University of Lisbon

João Carriço

National Microbiology Laboratory (NML)

Franklin Bristow
Aaron Petkau
Thomas Matthews
Josh Adams
Adam Olsen
Tara Lynch
Shaun Tyler
Philip Mabon
Philip Au
Celine Nadon
Matthew Stuart-Edwards
Morag Graham
Chrystal Berry
Lorelee Tschetter

Laboratory for Foodborne Zoonoses (LFZ)

Eduardo Toboada
Peter Kruckiewicz
Chad Laing
Vic Gannon
Matthew Whiteside
Ross Duncan
Steven Mutschall

Simon Fraser University (SFU)

Emma Griffiths
Geoff Winsor
Julie Shay
Matthew Laird
Bhav Dhillon
Raymond Lo

BC Public Health Microbiology & Reference Laboratory (PHMRL) and BC Centre for Disease Control (BCCDC)

Judy Isaac-Renton
Patrick Tang
Natalie Prystajeky
Jennifer Gardy
Linda Hoang
Kim MacDonald
Yin Chang
Eleni Galanis
Marsha Taylor
Cletus D'Souza
Ana Paccagnella

University of Maryland

Lynn Schriml

Canadian Food Inspection Agency (CFIA)

Burton Blais
Catherine Carrillo
Dominic Lambert

Dalhousie University

Rob Beiko
Alex Keddy