Gen2Epi GUI Quick Tutorial

Locations and Name of the Test DATA Used in this tutorial:

Main Data directory: "/home/gen2epi/Desktop/Test_DATA"

Fastq files: "/home/gen2epi/Desktop/Test_DATA/WHO_Data"

Genome files:

"/home/gen2epi/Desktop/Test_DATA/WHO_Full_Reference_genome/Chromosome"

Annotation files:

"/home/gen2epi/Desktop/Test_DATA/WHO_Genome_Annotation/Chromosome"

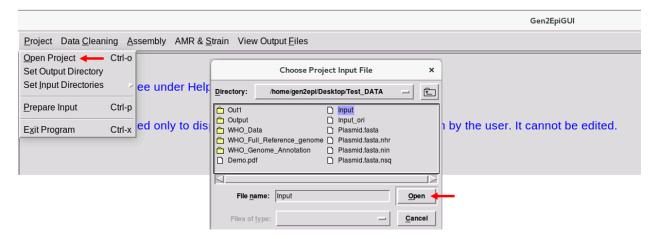
- Centrol State Mail State Mail
- 1) Click on the "Gen2EpiGUI" icon on the Desktop to start the program.

- To use the program please follow the instruction as given in "Introductory Demo" under Help menu on the top right-hand side. For detailed usage please see the description below:
 - a. <u>Introductory demo</u>: A quick tutorial on "how to use Gen2Epi GUI on test dataset".
 - b. User Manual: A step by step guide on "how can users analyze their own data using Gen2Epi GUI".
 - c. Reference Guide: This document describes the complete command line usage of Gen2Epi v0.1.

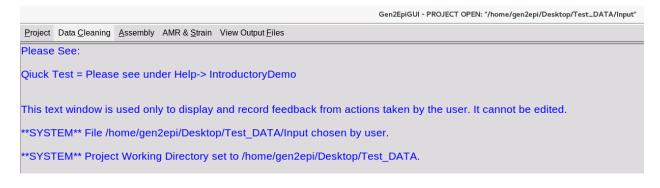
<u>Please Note</u>: For beginners "Introductory Demo" is highly recommended.

Gen2EpiGUI	- *
Project Data Gleaning Assembly AMR & Strain View Output Elles	В
Please See: Qiuck Test = Please see under Help-> IntroductoryDemo	FAQ Introductory Dem User Manual Reference Guide VM Help
This text window is used only to display and record feedback from actions taken by the user. It cannot be edited.	About
VIDO-InterVac Vaccee and Infections Diesase Organization	

3) Load the input text (tab separated text file describes the name of all fastq files) file by clicking on the "Open Project" and browse (and open) the file as shown in the picture below. "Input" file is present in the Test_DATA folder at Desktop.



You will see following ******SYSTEM****** message in the window which shows that the "Input" file is successfully loaded.



4) Set the output directory by clicking on "Set Output Directory" under "Project" menu. Browse the path (where you want to save the output files generated by Gen2Epi e.g. in this case path is "/home/gen2epi/Desktop/Test DATA/") and click open.

	Gen2EpiGUI -	PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly AMR & Strain	View Output Eiles	
Open Project Ctrl-o		
Set Output Directory 🔶	Choose a Directory ×	
Set Input Directories ee under Help-> I		
	Directory: /home/gen2epl/Desktop/Test_DATA	
Prepare Input Ctrl-p	C Out1	
Exit Program Ctrl-x ed only to display	C Output	the user. It cannot be edited.
	WHO_Data	
SYSTEM File /home/gen2epi/Deskto	WHO_Full_Reference_genome WHO Genome Annotation	
SYSTEM Project Working Directory s		
	File <u>n</u> ame: Open •	-
	Files of type: Cancel	

5) Write the name of your output directory (e.g. Output1) in the search box and click confirm.

Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"		- •	×
Project Data Cleaning Assembly AMR & Strain View Output Elles			<u>H</u> elp
Type the name of your new output subdirectory, or leave the field blank and click confirm if the listed directory will be your output directory:			
home/genZep/Desktop/Test_DATA/ Output!	Confirm	Cance	1
Please See:	4		
Qiuck Test = Please see under Help-> IntroductoryDemo	1		
This text window is used only to display and record feedback from actions taken by the user. It cannot be edited.			
SYSTEM File /home/gen2epi/Desktop/Test_DATA/Input chosen by user.			
SYSTEM Project Working Directory set to /home/gen2epi/Desktop/Test_DATA.			
SYSTEM Directory /home/gen2epi/Desktop/Test_DATA chosen by user.			

6) Set Input directories for "Fastq" files by clicking on "Set Input Directories->Fastq". Browse the folder "WHO Data" and click "Open" as shown in the figure below:

	Gen2EpiGUI - PROJECT OPEN	: "/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly AMR & Strain View Output	iles	
Open Project Ctrl-o Set Output Directories F Prepare Input Ctrl-p Genome File Ctrl-g	Demo	
Annotation Ctrl-a Exit Program Ctrl-x ed only to dispray and record **SYSTEM** File /home/gen2epi/Desktop/Test DA	Choose a Directory X	cannot be edited.
SYSTEM Project Working Directory set to /home	Count Count Count	
SYSTEM Directory /home/gen2epi/Desktop/Tes	Cutput1 Cutpu	
SYSTEM The output directory has been set to /I	HHO_Genome_Annotation	
SYSTEM The System Logs directory has been s	File name: Open	utput1/SystemLogs
	Files of type:	

 Set Input directories for "genome" files by clicking on "Set Input Directories"-> Genome File". Select the folder

"/home/gen2epi/Desktop/Test_	DATA/WHO_Full_	_Reference_	_genome/Chromos	ome"
and click "Open" as shown in t	the figure below:			

Gen2	2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly AMR & Strain View Output Files	
Open Project Ctrl-o Set Output Directory Fastq Ctrl-f Input Directories Fastq Ctrl-f	
Prepare Input Ctrl-p Genome File Ctrl-g Exit Program Ctrl-x Ctrl-x Ctrl-x	ken by the user. It cannot be edited.
SYSTEM File /home/gen2 Choose a Directory	×
SYSTEM Project Working	ome/Chromosome 💼
SYSTEM Directory /home/	
SYSTEM The output direc	
SYSTEM The System Log	put1/SystemLogs
File name:	<u>Open</u>
Files of type:	<u> </u>

<u>Please Note</u>: You will not be able to see anything when you choose a directory as shown above but that's all right.

8) Set Input directories for "annotation" files by clicking on
 "Set Input Directories ->Annotation". Select the folder
 "/home/gen2epi/Desktop/Test_DATA/WHO_Genome_Annotation/Chromosome" and

click on "Open" as shown in the figure below:

	Gen2EpiGUI - PROJECT OPEN: "/hom	e/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly AMR & Strain	View Output Eiles	
Open Project Ctrl-o Set Output Directory Eastq Ctrl-f Eastq	ntroductoryDemo	
Prepare Input Ctrl-p Genome File Ctrl-g Annotation Ctrl-a		
Exit Program Ctrl-x	Choose a Directory	× dited.
SYSTEM File /home/gen2epi/Deskto	Directory: /home/gen2epl/Desktop/Test_DATA/WHO_Genome_Annotation/Chromosome	- 1
SYSTEM Project Working Directory		
SYSTEM Directory /home/gen2epi/D		
SYSTEM The output directory has b		
SYSTEM The System Logs directory	File name:	Open MLogs
	Files of type:	<u>C</u> ancel

Again, you will not be able to see anything under your chosen directory as shown above but that's all right. Once all paths are set properly you will see following **SYSTEM** MESSAGES in Gen2Epi window

		GenzepiGUI - PROJECT OPEN: "/nome/genzepi/Desktop/Test_DATA/Input"
Project Data Cleaning	AMR & Strain	View Output <u>F</u> iles
Please See:		
Qiuck Test = Pleas	e see under Help-> Ir	ntroductoryDemo
This text window is	used only to display	and record feedback from actions taken by the user. It cannot be edited.
SYSTEM File /h	ome/gen2epi/Deskto	p/Test_DATA/Input chosen by user.
SYSTEM Projec	t Working Directory	set to /home/gen2epi/Desktop/Test_DATA.
SYSTEM Direct	ory /home/gen2epi/D	esktop/Test_DATA chosen by user.
SYSTEM The o	utput directory has be	een set to /home/gen2epi/Desktop/Test_DATA/Output1
SYSTEM The S	ystem Logs directory	has been set to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs
SYSTEM Direct	ory /home/gen2epi/D	esktop/Test_DATA/WHO_Data chosen by user.
SYSTEM Direct	ory /home/gen2epi/D	esktop/Test_DATA/WHO_Full_Reference_genome/Chromosome chosen by user.
SYSTEM Direct	ory /home/gen2epi/D	esktop/Test_DATA/WHO_Genome_Annotation/Chromosome chosen by user.

10) Now to clean your data: Under "Data Cleaning" menu – click on "Quality Control" submenu to check the quality of your raw fastq files.

				Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
/	Data <u>C</u> leaning		AMR & <u>s</u>	train View Output <u>F</u> iles
Please	<u>R</u> ead Mapping		Ctrl-r	
Qiuck [·]	Quality Control	←	Ctrl-q	-> IntroductoryDemo
	Tr <u>i</u> mming <u>B</u> oth		Ctrl-i Ctrl-b	
This te	Reset Trimming	g Parameters	5	play and record feedback from actions taken by the user. It cannot be edited.
SYST	FEM File /h	nome/gen2	2epi/De	sktop/Test_DATA/Input chosen by user.
SYS1	FEM Projec	ct Working	Direct	ory set to /home/gen2epi/Desktop/Test_DATA.
SYS1	FEM Direct	tory /home	/gen2e	pi/Desktop/Test_DATA chosen by user.
SYST	FEM The o	utput dire	ctory h	as been set to /home/gen2epi/Desktop/Test_DATA/Output1
SYS1	FEM The S	System Log	gs dire	ctory has been set to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs
SYS1	FEM Direct	tory /home	/gen2e	pi/Desktop/Test_DATA/WHO_Data chosen by user.
SYS1	FEM Direct	tory /home	/gen2e	pi/Desktop/Test_DATA/WHO_Full_Reference_genome/Chromosome chosen by user.
SYS1	FEM Direct	tory /home	/gen2e	pi/Desktop/Test_DATA/WHO_Genome_Annotation/Chromosome chosen by user.

Start of the analysis

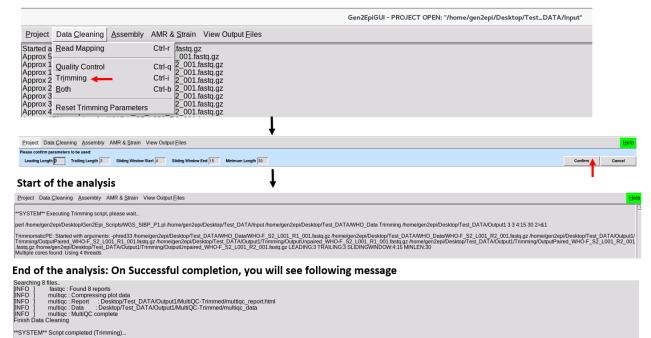
Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly AMR & Strain View Output Eiles
SYSTEM Executing Quality_Control script, please wait
perl /home/gen2epi/Desktop/Gen2Epi_Scripts/WGS_SIBP_P1.pl /home/gen2epi/Desktop/Test_DATA/Input /home/gen2epi/Desktop/Test_DATA/WHO_Data qualitycheck /home/gen2epi/Desktop/Test_DATA/Output1 2>&1
Started analysis of WHO-F_S2_L001_R1_001.fastq.gz Approx 5% complete for WHO-F_S2_L001_R1_001.fastq.gz Approx 15% complete for WHO-F_S2_L001_R1_001.fastq.gz Approx 20% complete for WHO-F_S2_L001_R1_001.fastq.gz
End of the analysis: On Successful completion, you will see following message

End of the analysis: On successful completion, you will see following mes	sat
Analysis complete for WHO L SE LOO1 P2 001 facto az	

	ipiete ioi who-c_55_coot_rz_oot.iasiq.gz
[WARNING]	multiqc : MultiQC Version v1.7 now available!
[INFO]	multige : This is MultiQC v1.5
(INFO)	multiqc:Template :default
[INFO]	multiqc : Searching /home/gen2epi/Desktop/Test_DATA/Output1/QualityControl/WHO-F_S2_L001_R1_001_fastqc.zip
[INFO]	multiqc : Searching '/home/gen2epi/Desktop/Test_DATA/Output1/QualityControl/WHO-F_S2_L001_R2_001_fastqc.zip'
[INFO]	multiqc : Searching '/home/gen2epi/Desktop/Test_DATA/Output1/QualityControl/WHO-G_S3_L001_R1_001_fastqc.zip'
[INFO]	multiqc : Searching /home/gen2epi/Desktop/Test_DATA/Output1/QualityControl/WHO-G_S3_L001_R2_001_fastqc.zip
(INFO)	multiqc : Searching '/home/gen2epi/Desktop/Test_DATA/Output1/QualityControl/WHO-K_S4_L001_R1_001_fastqc.zip'
[INFO]	multiqc : Searching '/home/gen2epi/Desktop/Test_DATA/Output1/QualityControl/WHO-K_S4_L001_R2_001_fastqc.zip'
[INFO]	multiqc : Searching '/home/gen2epi/Desktop/Test_DATA/Output1/QualityControl/WHO-L_S5_L001_R1_001_fastqc.zip'
[INFO]	multiqc : Searching '/home/gen2epi/Desktop/Test_DATA/Output1/QualityControl/WHO-L_S5_L001_R2_001_fastqc.zip'
Searching 8	
[INFO]	fastqc : Found 8 reports
[INFO]	multiqc : Compressing plot data
[INFO]	multiqc : Report : Desktop/Test_DATA/Output1/MultiQC-Raw/multiqc_report.html
[INFO]	multiqc : Data : Desktop/Test_DATA/Output1/MultiQC-Raw/multiqc_data
(INFO)	multiqc : MultiQC complete
SYSTEM	Script completed (Quality_Control)
**CVCTEM*	Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/Ouality_Control.log
SISIEM	Logged script output to momergenzepirzesktopri esc. DA rA/Output/SystemLogs/Quality_Control.log

<u>Note</u>: Quality control results are present under "/home/gen2epi/Desktop/Test_DATA/Output1". There will be two folders "MultiQC-Raw" and "QualityControl". Please open the .html file under these folders to visualize the quality of each sample.

11) In order to trim the raw reads, users have to click on the "trimming" submenu that will further ask for the parameters confirmation. Users have the option to use either the default parameters or change them according to their requirements. After confirmation, the program will start trimming the reads.



SYSTEM Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/Trimming.log

<u>Note</u>: Trimming results are present under "/home/gen2epi/Desktop/Test_DATA/Output1". There will be three folders "MultiQC-Trimmed", "Trimming" and "Trimmed_QC". Please open the .html file under these folders to visualize the quality of each sample.

12) It is also possible to run step 10 and 11 together as a single command. First open "/home/gen2epi/Desktop/Test_DATA/Output1" folder and delete "MultiQC-Raw", "quality control", "MultiQC-Trimmed", "Trimming" and "Trimmed_QC" folders. Now go to Gen2Epi GUI and click on the "both" tab under "Data Cleaning" menu.

		Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/In	put"
Project Data Cleaning Assembly AMR & Strain	View Output <u>F</u> iles		
Started a Approx 5 Approx 5 Approx 1 Approx 1 Approx 2 Approx 2 Approx 2 Ctrl-r Quality Control Ctrl-q Ctrl-q 2 001 fas 2 001 fas	stq.gz stq.gz stq.gz stq.gz stq.gz stq.gz stq.gz		
Project Data Cleaning Assembly AMR & Strain View Output Files			Help
Please confirm parameters to be used:	Indow End 15 Minimum Length 30		Confirm Cancel
	Gen2EpiGUI - PROJECT OPEN: */	home/gen2epi/Desktop/Test_DATA/Input"	
Project Data Cleaning Assembly AMR & Strain View Output Files			
SYSTEM Executing Quality_Control_AND_Trimming script, please wai	ait		
perl /home/gen2epi/Desktop/Gen2Epi_Scripts/WGS_SIBP_P1.pl /home/ge	en2epi/Desktop/Test_DATA/Input /home/gen2epi/Deskto	p/Test_DATA/WHO_Data both /home/gen2epi/Desktop/Test_DATA/Output1 3 3 4:15 30 2>	&1
Started analysis of WHOF-52 L001, R1_001.fastq.gz Approx 5% complete for WHOF-52 L001, R1_001.fastq.gz Approx 15% complete for WHOF-52 L001, R1_001.fastq.gz Approx 15% complete for WHOF-52 L001, R1_001.fastq.gz Approx 25% complete for WHOF-52 L001, R1_001.fastq.gz			
End of the analysis: On Successful of	completion, you will see fo	llowing message	
Finish Data Cleaning			
SYSTEM Script completed (Quality_Control_	_AND_Trimming)		
SYSTEM Logged script output to /home/gen2	2epi/Desktop/Test_DATA/Output1/Sy	stemLogs/Quality_Control_AND_Trimming.log	

<u>Note</u>: - You will find all the results under "/home/gen2epi/Desktop/Test_DATA/Output1" in five different folders i.e. "MultiQC-Raw", "quality control", "MultiQC-Trimmed", "Trimming" and "Trimmed_QC"

13) Now, perform the *de novo* assembly of trimmed reads by clicking on "*De Novo*" under Assembly tab. You will see the output as shown in the picture below.

<u>Please Note</u>: - This step may take a while depending on individual computer configurations. It could be possible that nothing happens when you try to press enter in the VM image, no need to worry VM will resume itself after completing the *de novo* assembly part. You can minimize the VM image window and come back later.

Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"	- •
Project Data Geaning Assembly AMR & Strain View Output Elles	H
inght reads: [homesjen2epiDeskspfTest_DATA/OutputI/Trimming/OutputPaired_WHO-F_S2_L001_R2_001.fastq.gz] interlaced reads: not specified single reads: [homesjen2epiDesktop/Test_DATA/OutputI/Trimming/OutputUnpaired_WHO-F_S2_L001_R1_001.fastq.gz', 'homesjen2epiDesktop/Test_DATA/OutputI/Trimming/OutputUnpaired_WHO-F_S2_L001_R2_001.fastq.gz', 'homesjen2epiDesktop/Test_DATA/OutputI/Trimming/OutputUnpaired_WHO-F_S2_L001_R2_001.fastq	
Other pränneters: Dir for temp files: home/gen2epi/Desktop/Test_DATA/OutputJ/Chrom_AssemblyTrimmedReads/WHO-F/mp Threads: 2 Memory limit (in Gb): 2	
===== Read error correction started.	
Enuming read error correction tool: home/gen2epi/DownloadS/SPAdes-312.0-Linuxbhirbspades-hammer / home/gen2epi/DesktopT des_DATA/OutputL/Chrom_AssemblyTrimmedReads/WHO-F/corrected/configs/config.info correction tool: home/gen2epi/DownloadS/SPAdes-312.0-Linuxbhirbspades-hammer / home/gen2epi/DesktopT des_DATA/OutputL/Chrom_AssemblyTrimmedReads/WHO-F/corrected/configs/config.info correction tool: home/gen2epi/DesktopT dest	

14) Once the *de novo* assembly (step 13) complete. You will see the following message

===== Mismatch correction finished.
* Corrected reads are in /home/gen2epi/Desktop/Test_DATA/Output1/Plasmid_AssemblyTrimmedReads/WHO-L/corrected/ * Assembled contigs are in /home/gen2epi/Desktop/Test_DATA/Output1/Plasmid_AssemblyTrimmedReads/WHO-L/contigs.fasta * Assembled scatfolds are in /home/gen2epi/Desktop/Test_DATA/Output1/Plasmid_AssemblyTrimmedReads/WHO-L/assembly_graph.fasta * Assembley graph is in /home/gen2epi/Desktop/Test_DATA/Output1/Plasmid_AssemblyTrimmedReads/WHO-L/assembly_graph.fastg * Assembly graph in GFA format is in /home/gen2epi/Desktop/Test_DATA/Output1/Plasmid_AssemblyTrimmedReads/WHO-L/assembly_graph.fastg * Paths in the assembly graph corresponding to the contigs are in /home/gen2epi/Desktop/Test_DATA/Output1/Plasmid_AssemblyTrimmedReads/WHO-L/assembly_graph_with_scatfolds.gfa * Paths in the assembly graph corresponding to the scatfolds are in /home/gen2epi/Desktop/Test_DATA/Output1/Plasmid_AssemblyTrimmedReads/WHO-L/assemblyTrimmedReads/WHO-L/assemblyTrimmedReads/WHO-L/assembly_graph.vith_scatfolds.gta
======= SPAdes pipeline finished.
l SPAdes log can be found here: /home/gen2epi/Desktop/Test_DATA/Output1/Plasmid_AssemblyTrimmedReads/WHO-L/spades.log
Thank you for using SPAdes!
SYSTEM Script completed (De_Novo)
SYSTEM Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/De_Novo.log

<u>Note</u>: Users will find the results under "/home/gen2epi/Desktop/Test_DATA/Output1" in four folder i.e. Chrom_AssemblyTrimmedReads (Assembled chromosome contigs in FASTA format generated from trimmed reads), ChromContigAssemblyTrimmedStat (Chromosome assembly statistics), Plasmid_AssemblyTrimmedReads (Assembled plasmid contigs in FASTA format generated from trimmed reads), and PlasmidContigAssemblytrimmedStat (Plasmid assembly statistics)

15) Perform the scaffolding of *de novo* assembly by clicking on "Assembly->Scaffolding->Chromosome->Multiple Genomes" as shown below:

	Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly AMR & Strain View Output Eiles	
[bwa_index] Construct BW De Novo Ctrl-d	
[bwa_index] 0.01 seconds Scaffolding Chromosome	
[bwa_index] Opdate BW1	上 Multiple Genomes 🗲 🗕
[bwa_index] Pack forward Preassembled Genome Plasmid Ct	1-I Single Genomes
[bwa_index] Construct SA from BWT and Occ 0.01 sec	
	a index -a is /home/gen2epi/Desktop/Test_DATA/Output1/Plasmid_AssemblyTrimmedReads/WHO-L/misc/a
[main] Real time: 0.054 sec; CPU: 0.029 sec	a mack -a is momengenzep//beskup/rest_DATA/output/reashing_Assembly minineukeads/wiro-Dimiscr
	pp : 173) Running bwa mem:/home/gen2epi/Downloads/SPAdes-3.12.0-Linux/bin/spades-bwa mem -
	DATA/Output1/Trimming/OutputUnpaired WHO-L S5 L001 R2 001.fastg.gz >/home/gen2epi/Desktop/Te
sT6/lib2_VXTd9s/tmp.sam	
[main] Version: 0.7.12-r1039	
	a mem -v 1 -t 2 /home/gen2epi/Desktop/Test_DATA/Output1/Plasmid_AssemblyTrimmedReads/WHO-L/mis
ng/OutputUnpaired_WHO-L_S5_L001_R2_001.fastq.gz	
[main] Real time: 0.025 sec; CPU: 0.034 sec	nn . 224) Adding comfle (home/gon2eni/Declar (fort DATA/Output) (Decmid Accombly TrimmedDecd

Start of the Scaffolding process

The following objects are masked from anopackage:basean:

intersect, setdiff, setequal, union Joining, by = "samples"

Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input" _				
Project Data Gleaning Assembly AMR & Strain View Output Elies	Help			
SYSTEM Executing Multiple_Genomes script, please wait perf AnomelgenZeepiDesktop/Test_DistributOS_SIBP_P3-Chr-C1.pl AnomelgenZepiDesktop/Test_DATA/hput AnomelgenZepiDesktop/Test_DATA/WHO_Full_Reference_genome/Chromosome AnomelgenZepiDesktop/Test_DATA/OutputI/Chrom_Assembly medReads AnomelgenZepiDesktop/Test_DATA/WHO_Genome_Annotation/Chromosome 2 TXT AnomelgenZepi/Desktop/Test_DATA/OutputI 2>&1 093 5331 (NFC: Starting Regult v2.0 199 35334 (NFC: Running Sibelia with block size 5000	yTrim			
Scaffolding process completion				
Finished: 2019-06-03 09:41:13 Elapsed time: 0:00:07.934315 Total NOTICEs: 5; WARNINGs: 5; non-fatal ERRORs: 0				
Thank you for using QUAST!				
Attaching package: 800dpty1800				
The following objects are masked from @Dpackage:stats@D:				
filter, lag				

SYSTEM Script completed (Multiple_Genomes)...

SYSTEM Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/Multiple_Genomes.log

Note: Output files generated from this step will be stored in folder "Chr_Scaffolds" under "/home/gen2epi/Desktop/Test_DATA/Output1".

16) To identify the different types of *Neisseria gonorrhoeae* plasmids using assembled plasmid contigs (in step 13), select "Assembly->Scaffolding->Plasmid".

			Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
Project	Data <u>C</u> leaning	Assembly AMR & Strain View Output Files	
		De Novo Ctrl-d Scaffolding Preassembled Genome Plasmid	

Once you click on "Plasmid". You will see the following output



<u>Note</u>: Output generated from this step will be stored in folder "Plasmid_Identification" under "/home/gen2epi/Desktop/Test_DATA/Output1".

17) NgMAST Typing

a. Update the underlying MAST database by clicking MAST->MAST Update tab

		Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly	AMR & Strain View Output Files	
	MASI MAST Update MLST MAST Typing STAR TETRES	

MAST database update

Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"

 $\underline{P}roject \quad Data \, \underline{C}leaning \quad \underline{A}ssembly \quad AMR \& \, \underline{S}train \quad View \, Output \, \underline{F}iles$

SYSTEM Executing MAST_Update script, please wait...

perl /home/gen2epi/Desktop/Gen2Epi_Scripts/MASTdbUpdate.pl 2>&1

SYSTEM Script completed (MAST_Update)...

SYSTEM Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/MAST_Update.log

b. Perform the MAST typing by selecting MAST->MAST Typing

Gen2EpiGUI - PROJECT	OPEN: "/home/gen2epi/Desktop	/Test_DATA/Input"
----------------------	------------------------------	-------------------

Project	Data <u>C</u> leaning	Assembly	AMR & Strain	View Output <u>F</u> iles
5 1				MAST Update MAST Typing
Strain	typing			

Gen2EpiGUI - PROJECT OPEN: "home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly AMR & Strain View Output Eiles
SYSTEM Executing MAST_Typing script, please wait
perl /home/gen2epi/Desktop/Gen2Epi_Scripts/WGS_SIBP_P4_Epi.pl /home/gen2epi/Desktop/Test_DATA/Input /home/gen2epi/Desktop/Test_DATA/Output1/Chr_Scaffolds NGMAST /home/gen2epi/Desktop/Test_DATA/Output1 2>&
SYSTEM Script completed (MAST_Typing)
SYSTEM Logged script output to /home/gen/gen/Desktop/Test_DATA/Output1/SystemLogs/MAST_Typing.log

<u>Note</u>: MAST strain typing for each sample is present in file "NgMAST.txt" under "/home/gen2epi/Desktop/Test_DATA/Output1" as shown below.

	A second se	-					
MultiQC-Trimmed	NgMAST.txt	Asse	Plasmid_ emblyTrimmedRe ads	PlasmidContigAssen blytrimmedStat	n Plasmid_Identification		
Open 👻 🎛			NgMA: ~/Desktop/Test.		Save = -		×
Samples NgMAST POR WHO-F_scaffolds.fas WHO-G_scaffolds.fas WHO-K_scaffolds.fas WHO-L_scaffolds.fas	ta 621 ta 1424	2002 90 917 915	165 18 10 21				
				Plain Text 🔻 Tab V	Nidth: 8 🗸 🛛 Ln 1, Col 1	•	INS

18) NgMLST Typing

a. Update the underlying MLST database by clicking MLST->MLST Update tab



	Senzepioor Thoseer of EN. /home/genzepi/Desktop/Test_DATA/input			
Project Data Cleaning Assembly AMR & Strain View Output Files				
SYSTEM Executing MLST_Update script, please wait				
perl /home/gen2epi/Desktop/Gen2Epi_Scripts/MLSTdbUpdate.pl 2>&1				
Read and write (return) sequences				
SYSTEM Script completed (MLST_Update)				
SYSTEM Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/ML	.ST_Update.log			

b. Perform the MLST typing by selecting MLST->MLST Typing

Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"					
oject Data Cleaning Assembly AMR & Strain View Output Ejles	Project Data Cleaning Assemb				
MAST MLST MLST Update STAB TETRES					
start of the strain typing analysis					

SYSTEM Executing MLST_Typing script. please wait.. pert.Mone/gen2epi/Desktop/Gen2Epi_Scripts/MLST_eileles.tasta.home/gen2epi/Desktop/Gen2Epi_Scripts/MLST_profile_bt 2>&1 Building a new DB, current time: 06/03/2019 10:38:30 New DB name: /home/gen2epi/Desktop/Gen2Epi_Scripts/MLST_alleles.fasta New DB name: /home/gen2epi/Desktop/Gen2Epi_Scripts/MLST_alleles.fasta Sequence type: Nucleotide BLAST database named /home/gen2epi/Desktop/Gen2Epi_Scripts/MLST_alleles.fasta Keep MBits: T Maximum file size: 100000000B Adding sequences from FASTA; added 7219 sequences in 0.144738 seconds. **SYSTEM** Script completed (MLST_Typing)... *SYSTEM** Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/MLST_Typing.log

<u>Note</u>: MLST strain typing for each sample is present in file "NgMLST.txt" under "/home/gen2epi/Desktop/Test DATA/Output1" as shown below.

	nContigAssembl TrimmedStat	Chr_Scaffolds	GenomeSta	inteAll.txt Mul	tiQC-Raw		
MultiQC-Trimmed	JgMAST.txt	And the second s	Plasmi AssemblyTri ads	mmedRe blytr	IContigAssem immedStat		
Open 🗸 🖻			LST.txt st_DATA/Output1		Save = -		×
Sample ST abcZ WHO-F_scaffolds_pubMLS WHO-G_scaffolds_pubMLS WHO-K_scaffolds_pubMLS WHO-L_scaffolds_pubMLS	T.txt 1903 T.txt 7363	fumC go	dh pdhC 9 67 9 67 9 67	pgm clona 157 148 157 148 78 148 78 148 78 149	al_complex 153 65 153 65 153 65 153 65 153 65		
			Plain Text 🔻	Tab Width: 8 🗸	Ln 1, Col 1	•	INS

19) NgSTAR

a. Update the underlying STAR database by clicking STAR->STAR Update->STARdbUpdate tab

	Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"				
Project Data Cleaning Assembly AMR & Strain View Output Files					
MASI MLST STAR TETRES STAR Update ► STARdbUpdate ← STAR Typing STARdbMetaDataUpdate	Ite				
STAR database update					
	Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"				
Project Data Cleaning Assembly AMR & Strain View Output Files					
SYSTEM Executing STAR_Database_Update script, please wait					
perl /home/gen2epi/Desktop/Gen2Epi_Scripts/ngSTARdb.pl 2>&1					
Read and write (return) sequences					
SYSTEM Script completed (STAR_Database_Update)					
SYSTEM Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/S	TAR_Database_Update.log				

b. Update the underlying STAR database metadata by clicking STAR->STAR Update->STARdbMetaDataUpdate tab

 Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"

 Project
 Data <u>C</u>leaning <u>A</u>ssembly
 AMR & <u>S</u>train
 View Output <u>E</u>iles

 MASI
 M_ST
 STAR Update
 STARdbUpdate

 STAR database metadata update
 STAR Typing
 STARdbMetaDataUpdate

 Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"

 Project
 Data <u>C</u>leaning <u>A</u>ssembly <u>AMR & Strain</u> View Output <u>E</u>iles

 SYSTEM Executing STAR_Database_META_Update script, please wait...

 perl /home/gen2epi/Desktop/Gen2Epi_Scripts/NgSTARmeta.pl 2>&1

 SYSTEM Script completed (STAR_Database_META_Update)...

SYSTEM Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/STAR_Database_META_Update.log

c. Perform the STAR typing by selecting STAR->STAR Typing

		Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly	AMR & Strain View Output Eiles	
, , , , , , , , , , , , , , , , , , , ,	MASI MLST STAR TETRES STAR Update STAR Typing	

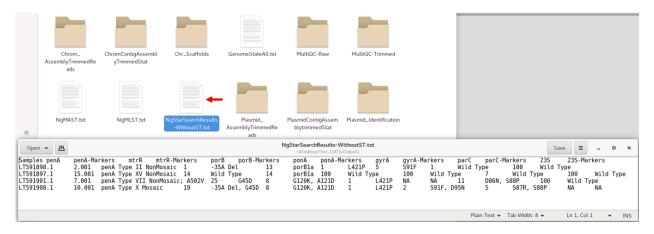
Start of the STAR typing process

Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"	_ •
Project Data Cleaning Assembly AMR & Strain View Output Elles	<u>H</u> e
SYSTEM Executing STAR_Typing script. please wait perf /home/gen2epi/Desktop/Gen2Epi_Scripts/WGS_SIBP_P4_Epi.pl /home/gen2epi/Desktop/Test_DATA/Input /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scatfolds ngstar /home/gen2epi/Desktop/Test_DATA/Out AMR-Genes-NgStar.tasta /home/gen2epi/Desktop/Gen2Epi_Scripts/AMR-Genes-NgStar-allelies.tasta 2>&1	put1 /home/gen2epi/Desktop/Gen2Epi_Scripts/
Completion of the STAR typing process	
The following objects are masked from â@package:statsâ@:	
filter, lag	
The following objects are masked from âmpackage:baseâm:	

intersect, setdiff, setequal, union
SYSTEM Script completed (STAR_Typing)...

SYSTEM Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/STAR_Typing.log

<u>Note</u>: NgSTAR output for each sample is present in file "NgStarSearchResults-WithoutST.txt" under "/home/gen2epi/Desktop/Test_DATA/Output1" as shown below.



20) Tetracycline Resistance

1) Click on "Nucleotide" under TETRES Tab to extract the rpsJ nucleotide sequences. These fasta sequences can be visualized under any multiple sequence alignment program.

	Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly AMR & Strain View Output Files	
MASI MLST STAR TETRES P Nucleotide	
Starting	
Gen2EpiGUI - PROJECT OPEN	"/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly AMR & Strain View Output Eiles	
SYSTEM Executing Nucleotide script, please wait perl /home/gen2epi/Desktop/Gen2Epi_Scripts/TetRes.pl /home/gen2epi/Desktop/Gen2Epi_Scripts/rpsJ.fasta /home/gen2epi/De	sktop/Test_DATA/Output1/Chr_Scaffolds/All_Sequences /home/gen2epi/Desktop/Test_DATA/Output1 2>&1
Completion	
Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-F.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-F.nsq Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-G.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-G.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-G.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-G.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-G.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-K.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-K.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-K.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-K.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-K.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-K.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-L.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-L.nin Can't open /home/gen2epi/Desktop/Test_DATA/OutputI/Chr_Scaffolds/All_Sequences/WHO-L.nin	Scaffolds_nucl.fasta: No such file or directory. scaffolds_nucl.fasta: No such file or directory.
SYSTEM Script completed (Nucleotide)	
SYSTEM Logged script output to /home/gen2epi/Desktop/Test_DATA/Output1/SystemLogs/Nucl	eotide.log

<u>Note</u>: Users can see the rpsJ nucleotide sequence for each sample in file "Nucl_rpsJ.fasta" under "/home/gen2epi/Desktop/Test_DATA/Output1" as shown below.

Chrom_ AssemblyTrimmedRe ads	ChromContigAssembl yTrimmedStat	Chr_Scaffolds	GenomeStateAll.txt	MultiQC-Raw	MultiQC-Trimmed				
NgMAST.txt	NgMLST.txt	NgStarSearchResults	Nucl_rpsJ.fasta	Plasmid_	PlasmidContigAssem				
		-WithoutST.txt		AssemblyTrimmedRe	blytrimmedStat				
Open 👻 🖭				Nucl_rpsJ.fasta			Save =	_ 0	ı x
ATGGCAAACCCAAAAATCCGTATCCGCTGCAAAGCTTATGATTACGCCCTGATTGACCGTTCTGCACAAG AAATCGTTGAAACGTACCGTGCTGGTGTGTGAAAAGGCCGCGATTGCAAGCGGCCCAATGGA GCGTTTCAACATTTTGCGTTCTCCGCACGTGAACAAAACTTCCCGTGACGAACGGCGCCGCCAT TTGCGCCTGATGGACTACAAACCCGAATAA >chr ENALLTS91898LT591898.1 1867 # 1966250 # 1966561 # 1 # ID=1 1867;partial=00;start_type=ATG;rbs_motif=AGGA;rbs_spacer=5-10bp;gc_cont=0.465 ATGGCAAACCCAAAAATCGCGATCGGCTGTTGTAAAAGGCCCGATCGCTTGTGCACAAG AAATCGTTGAAACCGAACAGCCGATAAACTTGCGACGGCTGTGTGGACGACTTGGCGACCCAA TTGCGCAACCCAAAAAGCGACCGGATCGATGACGCGGCGGCGGTGTGGACGACGGGCGCGGCGGCGGCGGGCG									

2) Click on "Protein" under TETRES Tab to extract the rpsJ protein alignment. These fasta sequences can be visualized under any multiple sequence alignment program.

		Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly	AMR & Strain View Output Files	
	MASI MLST STAR TETRES Nucleotide Protein	
		Gen2EpiGUI - PROJECT OPEN: "/home/gen2epi/Desktop/Test_DATA/Input"
Project Data Cleaning Assembly	AMR & <u>S</u> train View Output <u>F</u> iles	
SYSTEM Executing Protein script	t, please wait	
perl /home/gen2epi/Desktop/Gen2Ep	i_Scripts/SeqProt.pl /home/gen2epi/Desktop/Test_DA	TA/Output1/TetResOut /home/gen2epi/Desktop/Test_DATA/Output1 2>&1
Read and write (return) sequences		
SYSTEM Script completed (Prote	in)	
SYSTEM Logged script output to	/home/gen2epi/Desktop/Test_DATA/Output1/SystemLo	ogs/Protein.log

<u>Note</u>: Users can see the rpsJ nucleotide sequence for each sample in file "Prot_rpsJ.fasta" under "/home/gen2epi/Desktop/Test_DATA/Output1" as shown below.

Plasmid_Identification	Prot_rpsJ.fasta	QualityControl	SystemLogs	TetResOut	Trimmed_QC
Trimming	•				"Prot_rpsJ.fasta" selected (468 bytes)
Open 🔻 🖭		~/De	Prot_rpsJ.fasta esktop/Test_DATA/Output1		Save = - • ×
>WH0-F MANQKIRIRLKAYDYALID WH0-FSREQLEIRTHLRLM >WH0-G MANQKIRIRLKAYDYALID WH0-GSREQLEIRTHLRLM >WH0-K MANQKIRIRLKAYDYALID WH0-LSREQLEIRTHLRLM >WH0-LSREQLEIRTHLRLM	DIVDWTDKTTDALMKLE RSAQEIVETAKRTGAV\ DIVDWTDKTTDALMKLE RSAQEIVETAKRTGAV\ DIVDWTDKTTDALMKLE RSAQEIVETAKRTGAV\	DLPAGVDVEIKVQ /KGPIPLPTKIERFNILR: DLPAGVDVEIKVQ /KGPIPLPTKIERFNILR: DLPAGVDVEIKVQ /KGPIPLPTKIERFNILR:	SPHMNKT SPHMNKT		

21) View Output Files: - Finally, the text and HTML outputs generated from the above steps are accessible under "View Output File" tab.

		Gen2EpiGUI
Project Data Cleaning Assembly AMR &	<u>& S</u> train View Output <u>F</u> iles	
	View Text File in Main Window - Click of View Text in New Window View HTML in New Window	n
Project Data Cleaning Assembly AMR & Strain View Output	Gen2EpiGUI	
H.	Open Text File x tory: /homeigen2epiDesktopTesLDATA/Output1 Image: Constraint of the second s	Select the output file
		Gen2EpiGUI
Project Data Cleaning Assembly AMR & Strain	Niew Output <u>F</u> iles	
Samples NgMAST POR TBPB WHO-F_scaffolds.fasta 3303 2002 16 WHO-G_scaffolds.fasta 621 90 18 WHO-K_scaffolds.fasta 1424 917 10 WHO-L_scaffolds.fasta 1422 915 23	Ō	 Visualize the results