

BIONUMERICS Tutorial:

Custom genotyping plugin: predicting phenotypic traits from whole genome sequences

1 Aim

The *Custom genotyping plugin* allows you to detect and extract sequences from genome sequences using a BLAST or in silico PCR approach. Additionally, it allows the detection of mutations using a BLAST approach and the confirmation of species identity using sourmash.

In this tutorial we will use the custom knowledgebases which have been created in the "Creation of custom knowledgebases" tutorial to screen whole genome sequences of *Salmonella* with the *Custom genotyping* plugin for acquired and mutational resistance. We will also show how these custom knowledgebases can be used to perform an in-silico PCR analysis and to perform species confirmation.

2 Preparing the database

2.1 Introduction to the demonstration database

We provide a **WGS demo database** for *Salmonella* containing sequence read set data links for 62 samples, calculated de novo assemblies and wgMLST results (allele calls and quality information).

The **WGS demo database** for *Salmonella* can be downloaded directly from the *BIONUMERICS Startup* window (see 2.2), or restored from the back-up file available on our website (see 2.3).

2.2 Option 1: Download demo database from the Startup Screen

1. Click the **button**, located in the toolbar in the *BIONUMERICS Startup* window.

This calls the *Tutorial databases* window (see Figure 1).

- 2. Select WGS_demo_database_for_Salmonella_enterica from the list and select *Database* > *Download* (3).
- 3. Confirm the installation of the database and press < OK > after successful installation of the database.
- 4. Close the *Tutorial databases* window with *File* > *Exit*.

🖆 Tutorial databases					— 🗆	Х
File Database Tutorial Window						
Tutorial databases						
<u>1</u>						
Name	Size (MB)	Min	Max	-		
BandScoring demo database	15	7	8.1	^	The DemoBase Connected demonstration	^
DemoBase Connected	4	7.5	8.1		database. This database is used in several	
Neisseria MLST demo database	1	7.5	8.1		tutorials.	
SNP demonstration database	0	7.5	8.1			
WGS_demo_database_for_Brucella_spp	214	7.6	8.1			
WGS_demo_database_for_Burkholderia_cepacia	142	8	8.1			
WGS_demo_database_for_Escherichia_coli	601	7.5	8.1			
WGS_demo_database_for_Listeria_monocytogenes	288	7.5	8.1			
WGS_demo_database_for_MTBC	279	8	8.1			
WGS_demo_database_for_Salmonella_enterica	618	8	8.1			
WGS_demo_database_for_Staphylococcus_aureus	624	7.6	8.1			
						~
				*	1	-
Tutorials						
-7.						
¹						
Name				•		
Band matching and polymorphism analysis				~	This tutorial illustrates how to calculate a	^
Entry information fields and their properties					Principal Components Analysis (PCA) and a	
Configuring the database layout					Multi Dimensional Scaling (MDS) (sometimes also called Principal	
Selecting entries in a database					Coordinates Analysis (PCoA)) on a	
Combined analysis of fingerprint data					fingerprint data set and how to change the	
Calculating a PCA and an MDS on a fingerprint data set					layout of the obtained plots.	
calculating a FCA and an indo on a hingerprint data set						

Figure 1: The *Tutorial databases* window, used to download the demonstration database.

The **WGS_demo_database_for_Salmonella_enterica** appears in the *BIONUMERICS Startup* window.

5. Double-click the **WGS_demo_database_for_Salmonella_enterica** in the *BIONUMERICS Startup* window to open the database.

2.3 Option 2: Restore demo database from back-up file

A BIONUMERICS back-up file of the demo database for *Salmonella enterica* is also available on our website. This backup can be restored to a functional database in BIONUMERICS.

6. Download the file WGS_Salm.bnbk file from https://www.bionumerics.com/download/ sample-data, under 'WGS_demo_database_for_Salmonella_enterica'.



In contrast to other browsers, some versions of Internet Explorer rename the WGS_Salm.bnbk database backup file into WGS_Salm.zip. If this happens, you should manually remove the .zip file extension and replace with .bnbk. A warning will appear ("If you change a file name extension, the file might become unusable."), but you can safely confirm this action. Keep in mind that Windows might not display the .zip file extension if the option "Hide extensions for known file types" is checked in your Windows folder options.

- 7. In the *BIONUMERICS Startup* window, press the button. From the menu that appears, select **Restore database...**.
- 8. Browse for the downloaded file and select *Create copy*. Note that, if *Overwrite* is selected, an existing database will be overwritten.

- 9. Specify a new name for this demonstration database, e.g. "WGS_Salmonella_demobase".
- 10. Click < OK > to start restoring the database from the backup file.
- 11. Once the process is complete, click < *Yes*> to open the database.

The *Main* window is displayed (see Figure 2).

Example 2	Type Baquetor read set by Datactor types Charactor types		RR4119785 RR1752879 RR1643091 RR1427104 RR3402078 RR2585567 RR1653383 RR4842191	Modified date 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Corpanism All Entries> Organism Samonela enterica Samonela enterica Samonela enterica Samonela enterica subsp. enterica Samonela enterica Samonela enterica Samonela enterica	Serovar_NCBI 28:e,h:- Aberdeen Agona Ajobo Amsterdam Anatum Bangkok	Resistance genes_NDARO aph(3°)-b,fosA7,su2,tet(A) blaCMY-2,tet(C)	Run_accession SRR3724530 SRR4119765 SRR1643091 SRR1643091 SRR16427104 SRR3402078	1	2 3 4 5		Name	⊗ 6, 6	lified date
Ausset	Type Bequence read set by TO Character types Dependence types Character types Character types Character types	>es : : : : : : : : : : : : : : : : : : :	ey RR3724530 RR4119785 RR1752879 RR1643091 RR1643091 RR1643091 RR2585567 RR1653383 RR4842191 RR4842191 RR3049577 RR2421532	Modified date 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Organism Satnonela enterica Satnonela enterica Satnonela enterica subsp. enterica Satnonela enterica subsp. enterica Satnonela enterica Satnonela enterica Satnonela enterica	Serovar_NCBI 28:e,h:- Aberdeen Agona Ajobo Amsterdam Anatum Bangkok	Resistance genes_NDARO sph(3")-b,fosA7,sul2,tet(A)	SRR3724530 SRR4119785 SRR1752879 SRR1643091 SRR1427104	1		•••	Name	Modi	lified date
Ausset	Type Bequence read set by TO Character types Dependence types Character types Character types Character types	>es : : : : : : : : : : : : : : : : : : :	ey RR3724530 RR4119785 RR1752879 RR1643091 RR1643091 RR1643091 RR2585567 RR1653383 RR4842191 RR4842191 RR3049577 RR2421532	Modified date 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Organism Satnonela enterica Satnonela enterica Satnonela enterica subsp. enterica Satnonela enterica subsp. enterica Satnonela enterica Satnonela enterica Satnonela enterica	Serovar_NCBI 28:e,h:- Aberdeen Agona Ajobo Amsterdam Anatum Bangkok	Resistance genes_NDARO sph(3")-b,fosA7,sul2,tet(A)	SRR3724530 SRR4119785 SRR1752879 SRR1643091 SRR1427104	1		•••	Name	Modi	lified date
2 2	Sequence read set by Bopartice read set by Character types Character types Internet types Character types Internet types Sequence read set by Internet types Sequence types	bes 1 bes 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1	RR3724530 RR4119785 RR1752879 RR1643091 RR1427104 RR3402078 RR2585567 RR1653383 RR4842191 RR3049577 RR3049577 RR2421532	2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Samonela enterica Samonela enterica Salmonela enterica Salmonela enterica subsp. enterica Salmonela enterica subsp. enterica Salmonela enterica Salmonela enterica subsp. enterica	28:e,h:- Aberdeen Agona Ajobo Amsterdam Anatum Bangkok	aph(3")-lb,fosA7,sul2,tet(A)	SRR3724530 SRR4119785 SRR1752879 SRR1643091 SRR1427104	•		•••	Salmone		
 vegating vegating<	ng Sequence read set by TC Character types Sequence types Character types trimmedStats Sequence read set by TC_CalTypes Character types second	bes 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1	RR4119785 RR1752879 IRR1643091 IRR16427104 IRR3402078 IRR2585567 IRR1653383 IRR4842191 IRR3049577 IRR2421532	2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Salmonella enterica Salmonella enterica subsp. enterica Salmonella enterica subsp. enterica Salmonella enterica Salmonella enterica Salmonella enterica	Aberdeen Agona Ajobo Amsterdam Anatum Bangkok		SRR4119785 SRR1752879 SRR1643091 SRR1427104	:		••		la outoreax 2020	-07-15 1
3 wydu 27 4 4 derova 5 wydu 27 4 4 derova 5 wydu 27 6 wydu 27 7 wydu 27 7 wydu 27 7 wydu 27 7 wydu 27 9 20 <t< td=""><td>T Character types Sequence types Character types rimmedStats Sequence read set typ T_CalTypes Character types</td><td>> C S</td><td>RR1752879 RR1643091 RR1427104 RR3402078 RR2585567 RR1653383 RR4842191 RR3049577 RR3049577</td><td>2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11</td><td>Salmonella enterica Salmonella enterica subsp. enterica Salmonella enterica subsp. enterica Salmonella enterica Salmonella enterica Salmonella enterica subsp. enterica</td><td>Agona Ajobo Amsterdam Anatum Bangkok</td><td></td><td>SRR1752879 SRR1643091 SRR1427104</td><td>÷</td><td></td><td></td><td></td><td></td><td></td></t<>	T Character types Sequence types Character types rimmedStats Sequence read set typ T_CalTypes Character types	> C S	RR1752879 RR1643091 RR1427104 RR3402078 RR2585567 RR1653383 RR4842191 RR3049577 RR3049577	2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Salmonella enterica Salmonella enterica subsp. enterica Salmonella enterica subsp. enterica Salmonella enterica Salmonella enterica Salmonella enterica subsp. enterica	Agona Ajobo Amsterdam Anatum Bangkok		SRR1752879 SRR1643091 SRR1427104	÷					
 denova Squarky Squarky Squarky Squarky Squarky Squarky Toparan Strange <	Sequence types Character types rimmedSlats. Sequence read set by T_CallTypes. Character types	> C S	RR1643091 RR1427104 RR3402078 RR2585567 RR1653383 RR4842191 RR3049577 RR2421532	2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Salmonella enterica subsp. enterica Salmonella enterica subsp. enterica Salmonella enterica Salmonella enterica Salmonella enterica subsp. enterica	Ajobo Amsterdam Anatum Bangkok		SRR1643091 SRR1427104	÷					
C gualty C gual	Character types rimmedStats. Sequence read set by T_CalTypes. Character types	> = = = = = = = = = = = = = = = = = = =	RR1427104 RR3402078 RR2585567 RR1653383 RR4842191 RR3049577 RR2421532	2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Salmonella enterica subsp. enterica Salmonella enterica Salmonella enterica Salmonella enterica subsp. enterica	Amsterdam Anatum Bangkok	blaCMY-2,tet(C)	SRR1427104	1					
Constants	rimmedStats Sequence read set typ T_CallTypes Character types	> = = = = = = = = = = = = = = = = = = =	RR3402078 RR2585567 RR1653383 RR4842191 RR3049577 RR2421532	2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Salmonella enterica Salmonella enterica Salmonella enterica subsp. enterica	Anatum Bangkok	blaCMY-2,tet(C)							
feds Database deal feds Database deal feds Database deal Crganism Grows_NCBI Serves_NCBI BioPhysel BioPhysel BioPhysel BioPhysel Serves Surve	ST_CalTypes Character types	,	RR2585567 RR1653383 RR4842191 RR3049577 RR2421532	2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Salmonella enterica Salmonella enterica subsp. enterica	Bangkok								
Constant of the second se	tesign ⊾ E⊡ र. ↑	<pre></pre>	RR1653383 RR4842191 RR3049577 RR2421532	2020-07-15 20:50:11 2020-07-15 20:50:11 2020-07-15 20:50:11	Salmonella enterica subsp. enterica			SRR2585567	1.					
felds Database desil Name Companism Servers rNCBi Resistance genes Servers rNCBi Resistance genes Servers rNCBi BioProject BioProject Source Source MLST PubMLST AI C	, 6 ~ ↑	,	RR4842191 RR3049577 RR2421532	2020-07-15 20:50:11 2020-07-15 20:50:11		Bardo		SRR1653383						
felds Database desil Name Companism Servers rNCBi Resistance genes Servers rNCBi Resistance genes Servers rNCBi BioProject BioProject Source Source MLST PubMLST AI C	, 6 ~ ↑	,	RR3049577 RR2421532	2020-07-15 20:50:11		Braenderup		SRR4842191						
felds Database desu Name Companism Serovar NCBI Resistance genes Serovar NCBI Resistance genes Serovar NCBI Serovar NCBI	, 6 ~ ↑	> 	RR2421532		Salmonella enterica	Bron		SRR3049577						
felds Database desu Name Companism Serovar NCBI Resistance genes Serovar NCBI Resistance genes Serovar NCBI Serovar NCBI	, 6 ~ ↑			2020-07-15 20:50:11	Salmonella enterica	Cerro		SRR2421532						
Name Name Organism Servery NCBI Resistance genes Run accession BioProject BioProject BioProject Source Source MLST PubMLST A C Promet Resi	, 6 ~ ↑			2020-07-15 20:50:11	Salmonella enterica Salmonella enterica subso, enterica	Chicago		SRR1220728				<		
Name Name Organism Servery NCBI Resistance genes Run accession BioProject BioProject BioProject Source Source MLST PubMLST A C Promet Resi	, 6 ~ ↑		RR3194565	2020-07-15 20:50:11	Salmonella enterica	derby	aac(3)-Vla,aadA1,fosA7,g	SRR1220720 SRR3194565	1.			Identification p	piects Decision no	otwork
Name Organism Servors_IVCBI Resistances Run_accession BioProject Isolato Source Source MLST PubMLST A Power assi			RR1030845	2020-07-15 20:50:11	Salmonella enterica subsp. enterica		adu(3)-Via,aduk (,105k//,q	SRR1030845	1.					
Serovar_NCBI Resistance genes, Run_accession BioProject Isolate Isolate Source MLST PubMLST Av C			RR1183899	2020-07-15 20:50:11	Salmonella enterica subsp. enterica	Enteritidis		SRR1030645	1.			_2∃ +		
Serovar_NCBI Resistance genes, Run_accession BioProject Isolation_source Source MLST PubMLST Av < pprint files Power ass			RR1646564	2020-07-15 20:50:11		Hadar	aph(3")-lb,aph(6)-ld,tet(A)	SRR1646564	1.1			Name	Modifie	ed date
Serovar_NCBI Resistance genes, Run_accession BioProject Isolation_source Source MLST PubMLST Av < pprint files Power ass	Fixed		RR1046564	2020-07-15 20:50:11	Salmonella enterica Salmonella enterica subsp. enterica	Heidelberg	aph(3")-ib,aph(6)-id,tet(A)	SRR1040504 SRR1105667	1.					-
Resistance genes Run_accession BioProject Isolate Isolation_source Source MLST PubMLST Av < rprint files Power ass			RR3289809	2020-07-15 20:50:11		Heidelberg		SRR1105667 SRR3289809	1.1					
Run_accession BioProject Isolate Isolate Source MLST PubMLST A C print files Power ass	Fixed Fixed		RR1574295	2020-07-15 20:50:11	Salmonella enterica Salmonella enterica		aac(3)-Vla,aadA1,blaHER-3	SRR1574295	1.					
BioProject Isolate Isolation_source Source MLST PubMLST Av < print files Power ass			RR1574295	2020-07-15 20:50:11		lla 18:z4,z23:-	aac(3)-Vla,aadA1,qacEdelt	SRR1574295 SRR1534841	1.1					
solate solation_source Source MLST PubMLST Av C			RR3476365	2020-07-15 20:50:11	Salmonella enterica	Kentucky	aph(3")-lb,aph(6)-ld,tet(B)	SRR1534841 SRR3476365	1.					
Isolation_source Source MLST PubMLST Av < rprint files Power ass	Fixed			2020-07-15 20:50:11	Salmonella enterica	Manhattan	(00. 0		1.1					
Source MLST PubMLST A < rprint files Power ass	Fixed		RR1783160		Salmonella enterica	Montevideo	aac(6')-lb,aadA1,aph(3")-lb	SRR1783160	1.					
MLST PubMLST A				2020-07-15 20:50:11	Salmonella enterica	Newport	aph(3")-lb,aph(6)-ld,blaCM	SRR1203019						
<pre>c print files Power as:</pre>	Fixed		RR3098669	2020-07-15 20:50:11	Salmonella enterica	Saintpaul	blaTEM-1	SRR3098669	•		•••			
print files Power as:	T Achtman ST Flexible		RR1574259	2020-07-15 20:50:11	Salmonella enterica	Schwarzengrund	aph(3")-lb,aph(6)-ld,blaTEM	SRR1574259	•					
print files Power as:	>		RR1107480	2020-07-15 20:50:11	Salmonella enterica subsp. enterica	Thompson		SRR1107480	•		•••			-
	,		RR340757	2020-07-15 20:50:11	Salmonella enterica subsp. enterica	Typhi		ERR340757	•		•••			_
	assemblies Annotations			2020-07-15 20:50:11	Salmonella enterica subsp. enterica	Typhimurium		SRR3476793				Alignments	LAST projects Ch	hrom, C
		1000	RR3330219	2020-07-15 20:50:11	Salmonella enterica	Typhimurium var. 0:5-	blaCMY-2,sul2,tet(A)	SRR3330219						
100	⊗ 6., 6. ⊽.		RR3205698	2020-07-15 20:50:11	Salmonella enterica	Waycross		SRR3205698				+ 🖻	⊗ 6, 6	Ô.
File name	Experiment type		RR6128335	2020-08-10 10:11:01	Salmonella enterica subsp. enterica	Typhimurium		SRR6128335				Name	Modifi	ied dat
			RR6128336	2020-08-10 10:11:06	Salmonella enterica subsp. enterica	Typhimurium		SRR6128336	•		•••			-
			RR6128337	2020-08-10 10:11:10	Salmonella enterica subsp. enterica	Typhimurium		SRR6128337	•	• • •	•••			
			RR6128338	2020-07-15 20:50:11	Salmonella enterica subsp. enterica	14\[5]12:I:-		SRR6128338						
			RR6128339	2020-08-10 10:11:14	Salmonella enterica subsp. enterica	Typhimurium		SRR6128339	•	•••	•••			
			RR6128340	2020-08-10 10:11:18	Salmonella enterica subsp. enterica	Typhimurium		SRR6128340	•		•••			
			RR6128341	2020-07-15 20:50:11	Salmonella enterica subsp. enterica	14\[5]12:i:-		SRR6128341	•	•••	•••			
				2020-07-15 20:50:11	Salmonella enterica subsp. enterica	14\[5]12:I:-		SRR6128342	•		•••			
			RR6128343	2020-07-15 20:50:11	Salmonella enterica subsp. enterica	14\[5]12:i:-		SRR6128343	•		•••			
			RR6128344	2020-07-15 20:50:11	Salmonella enterica subsp. enterica	14\[5]12:i:-		SRR6128344	•		•••			
			RR6128345	2020-07-15 20:50:11	Salmonella enterica subsp. enterica	14\[5]12:i:-		SRR6128345	•		•••			
<			RR6128346	2020-07-15 20:50:11	Salmonella enterica subsp. enterica	14\[5]12:i:-		SRR6128346	1			~ <		

Figure 2: The Salmonella demonstration database: the Main window.

3 About the demonstration database

The WGS demo database contains links to sequence read set data on NCBI's sequence read archive (SRA) for 62 publicly available sequencing runs. Additional information (in entry info fields Organism, Serovar etc.) was collected from the corresponding publications and added to the demonstration database.

Seven experiments are present in the demo database and are listed in the *Experiment types* panel (see Figure 3).

Experim	nent ty	pes	
କ୍ଷ	-	+ 🗗 🛛 🔍	All Experiment types>
	#	Name	Туре 🗸 🗸 🗸
≣ ≩	1	wgs	Sequence read set types
≣ [2	wgsLong	Sequence read set types
	3	wgMLST	Character types
🗆 🛟	4	denovo	Sequence types
	5	quality	Character types
≣ ≣	6	wgs_TrimmedStats	Sequence read set types
	7	wgMLST_CallTypes	Character types

Figure 3: The Experiment types panel in the Main window.

1. Click on the green colored dot for one of the entries in the first column in the *Experiment presence* panel. Column 1 corresponds to the first experiment type listed in the *Experiment types* panel, which is **wgs** in the default configuration.

In the *Sequence read set experiment* window, the link to the sequence read set data on NCBI (SRA) with a summary of the characteristics of the sequence read set is displayed: *Read set size*, *Sequence length statistics*, *Quality statistics*, *Base statistics* (see Figure 4).

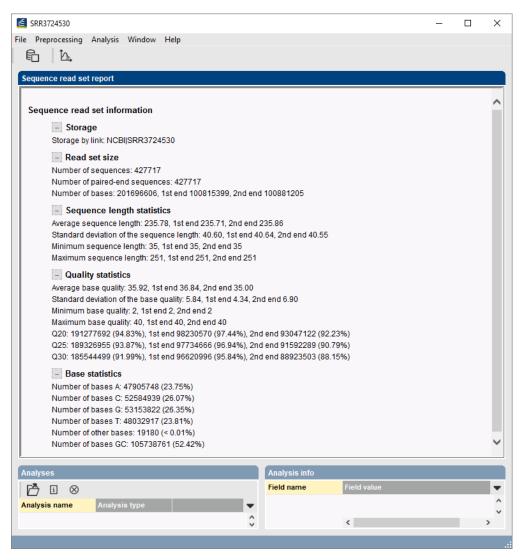


Figure 4: The sequence read set experiment card for an entry.

- 2. Close the Sequence read set experiment window.
- 3. Click on the green colored dot for one of the entries in the fourth column in the *Experiment presence* panel. Column 4 corresponds to the fourth experiment type listed in the *Experiment types* panel, which is **denovo** in the default configuration.

The *Sequence editor* window opens, containing the results from the de novo assembly algorithm, i.e. concatenated de novo contig sequences (see Figure 5).

4. Close the Sequence editor window.

The sequence read set experiment type **wgs_TrimmedStats** contains some data statistics about the reads retained after trimming, used for the de novo assembly.

The sequence read set experiment type **wgsLong** contains the links to long read sequence read

	530 (Sec	juence Vie	ewer)													_			×
e Sequer	nce He	eader Ar	notation	View	Tools V	Vindow	Help												
		<u>}</u>	<u>n r</u>	χ β] [3	\otimes	<u>Q</u>]	Þ	=	⊽1 '	▽2 ▽	∿ ⊗							
Sequence	Editor																		
TGCGTA GGCGTC GGAAAC GCTTAC GCTTAC ATCGGT TACACA GCagtt gtccga ccgtat	CATCI CCTIG CCGGA CCGGA CCGGC ACCAI GCACA CCCAI	GCCG TACA CATC GGGT TACC ACCA ACCA ggga tggt	GGCTG ACAAA TGGCC CGGTC CCATC TCGCG CTCAG aacac aatgg	A AATT GCTC G AAGC I ATAC I GGAA A ACCC I AAtc g tatt g sgtt	TGAA. GATG GACAA. CCATG GTGGG CCGAC CCGAC CCGAC	AA TO CC GO AG CO AC CO CA CI CG GI GG GI GG GI AC CO CA CI CG GI CG GI CG GI CG GI CG GI CG GI	CAGG CTTC. CCTG GGCC TACC TACC CCG CCG CCG CCG CCG CCG CCG CCG CCG	GCGT AGCC GATA TCAG TAAC TGGA aaca ggca ttcc	C AG G CC T TC C GC G TG G CT g Cg g gt c Cg	TCGC GATA CGCC CACC ACAG CCAA acat agtg gttg	AGCC TTAA GTTG ACGA GCGA CTGT Caat IGCGG IGCGG IGCGG	TG1 CA1 TA1 CA1 CA1 ACC GC2 TCA CA1 ACC CA1 TTA ACC CA1 CA1 CA1 CA1 CA1 CA1 CA1 CA1 CA	CGAA CGAC ATTT TCTG CCAG ATGA cagg rggkg	TCC CTG CCC TTA CAA act ggr rtt	ACAC GGGG CAGA CGAT ACCG GGCT ttat gktt tata	TTTTC ACGCG TCACG CTTTT GAGTG TAACG gctgc wttka acctt	G G A A A A C A A C A C A C A C A C A C	15750 15820 15890 15960 16030 16100 16170 16240 16310	~
] @,		S4	'5000	X≻ Ç SA S	10000 SAL		4970	SALM_5	i050 \$	⇒—	0000 ' 	SALM_112	1 '250 (08	00		'3000 ⊐K	- 	-
¥ Ч		<	S4	'5000' ' '	X, Ç	10000 SAL	C)(i050 \$	⇒—	_	SALM_112				'3000 ⊐K	- 	-
Annotation		<	S4		ĴÇ, Ç SA Ş	- '10000 SAL	EXI <mark>SALI</mark>	490		5050 \$	⇒—	_	SALM_112				'3000 ⊐K	- 	۰ ۸_1197 ک
- 1		<	SA SA	'5000 L	tr) p sa s)]	EX) Bally	4920			⇒—	_	SALM_112				'3000 (1) (3)	- 	-
Annotation	1 st	Feature	SA SA	Q Start		SAL	CA) SALM	4920	SALUI_S	⊗ ∕alle	SALM_708	9 S	16122				'3000	- 	-
Annotation Feature list			SA key S S S			SAL .	EM Sal I	49°0 2' ^ 8 61	SALIN_S	× /alle ocus_t /evide	SALM_708	9 S 5253 27" ALM_50 00.0	16122	08	Sque_104	(\		ALSALM	-
Annotation Feature list	10 11 12	Feature	s/ key S S S S	Q Start 12076 15253		SAL End 14247 16122		492°0		/alle ocus_t /evide /no	SALM_708	5253 27" ALM_50 00.0 vd=1;s	16122 50" tart=1	5252;	344 104 stop=10	15 5122;ci	d=d	enovo	-
Annotation Feature lis -→ ☆	10 11 12 13 14	Feature	sø sø s s s s s s	Start 12076 15253 18389		End 14247 16122 18991		2 [·] ^ 8 [·] 61		/alle ocus_t /evide /no	BALM_708 BALM_708 15 1e="12 agg="S2 ince=10 vte="fv te="fv	5253 27" ALM_50 0.0 wd=1;s 5YVQRI YYGGTA	16122 50" tart=1 ITVEFT GTQITL	5252; LSDGR YIWGL	stop=1; TFDNGK	(\	d=d	enovo CFATV AQSTA	-
Annotation Feature lis → ☆ ∜ -→ ◇	10 11 12 13 14	Feature	key S S S S S S S	Start 12076 15253 18389 21151		End 14247 16122 18991 22005		4900 2 ^ ^ 8 61 81 11 v	SALUL 5	Alle bcus_t (evide /no hslati	BALM_708 BALM_708 15 1e="12 agg="S2 ince=10 vte="fv te="fv	5253 27" ALM_50 0.0 wd=1;s 5YVQRI YYGGTA	16122 50" tart=1 ITVEFT GTQITL	5252; LSDGR YIWGL	stop=1; TFDNGK	5122;ci	d=d	enovo CFATV AQSTA	-

Figure 5: The Sequence editor window.

data (typically PacBio or MinION datasets). In this demo database, no links are defined for this experiment.

The other three experiments contain data related to the wgMLST analysis performed on the samples:

- Character experiment type wgMLST contains the allele calls for detected loci in each sample, where the consensus from assembly-based and assembly-free calling resulted in a single allele ID.
- Character experiment type **quality** contains quality statistics for the raw data, the de novo assembly and the different allele identification algorithms.
- · Character experiment type wgMLST_CallTypes: contains details on the call types.

4 Installing the custom genotyping plugin

Proceed as follows to install the *Custom genotyping plugin*:

- 1. Call the *Plugins and Scripts* dialog box from the *Main* window with *File* > *Install / remove plugins...* (,B).
- 2. Select the *Custom genotyping plugin* from the list and press the *<Install*> button.

5

3. Confirm the installation of the plugin.

A message appears, confirming the installation of the plugin and prompting you to restart BIONU-MERICS.

- 4. Press < OK > in the confirmation message.
- 5. Press < *Close* > to close the *Plugins and Scripts* dialog box.
- 6. Close and re-open the database to complete the installation of the plugin.

The *Custom genotyping plugin* installs menu items in the main menu of the software under *Geno-typing* (see Figure 6).

Genotyping Window Help	
Manage knowledgebases Manage models	
Reports Settings Run active model	8 Ex Entries>
Tools •	Export Minhash signatures Export knowledgebase fasta

Figure 6: New menu items, available after installation of the *Custom genotyping plugin*.

5 Managing knowledgebases

Before a genotyping model can be created, at least one knowledgebase should be available. The following custom knowledgebases have been created in the "Creation of custom knowledgebases" tutorial:

- Disinfectant_resistance: A BLAST-based knowledgebase for the detection and extraction of disinfectant resistance genes.
- In-silico PCR: An in-silico PCR-based knowledgebase for the detection and extraction of markers for the identification of *Salmonella enterica* serovars.
- Mutational_resistance_Salmonella: A BLAST-based knowledgebase for the detection of mutational antibiotic resistance.
- Species_confirmation_Salmonella: A minhash-based knowledgebase for *Salmonella* species confirmation.

These four custom knowledgebases can be downloaded from the BIONUMERICS website (https://www.bionumerics.com/download/sample-data, click on "Custom knowledgebases").

1. In the *Main* window of the BIONUMERICS database select *Genotyping* > *Manage knowl-edgebases...* to open the *Manage knowledge bases* dialog box (see Figure 7).

The *Manage knowledge bases* dialog box shows all currently available genotyping knowledgebases in the BIONUMERICS database. Initially, this dialog is empty.

2. Press < *Add local...*> to open the *Register local knowledge base* dialog box.

inage knov	vledgebases				?	×
Гуре	Name	Version	Checksum	Path		
Add local	Delete Create example					
						Close

Figure 7: The Manage knowledge bases dialog box.

3. Press < *Browse...*> and browse for the Disinfectant_resistance knowledgebase folder in the downloaded Custom knowledgebases folder (see Figure 8).

Register local	knowledgebase	?	×
	e folder which contains the local knowledgebase. The files must have the co the feature in which they will be used. Please consult the plugin manual for r		me
\Example kno	wledgebases\Custom knowledgebases\Disinfectant_resistance Brows	se	
Validate for:	Sequence detection \checkmark Check		
	ОК	Cano	cel

Figure 8: The Register local knowledge base dialog box.

Optionally, the knowledgebase can be validated before it is added.

4. Select the type of features for which the knowledgebase is intended for (i.e. Sequence detection, Sequence extraction and Acquired traits detection) and press < *Check* >.

Any detected issues will be reported in the *Knowledge base validation issues* dialog box. An error means that the knowledge base is not usable for the selected feature type: a different knowledge base should be specified or the knowledgebase's files should be corrected according to the error description. It will not be possible to add a knowledgebase with a validation error.

A knowledge base for which only warning messages are raised might be usable for the selected feature, but not all options of the feature are applicable.

- 5. Press <*OK*> in the *Register local knowledge base* dialog box to add the validated knowledgebase to the list in the *Manage knowledge bases* dialog box.
- 6. Repeat the previous steps to add the other three downloaded knowledgebases as well (i.e. "Insilico PCR", "Mutational_resistance_Salmonella" and "Species_confirmation_Salmonella"). The "In-silico PCR" knowledgebase can be validated for the in-silico PCR detection and extraction features, the "Mutational_resistance_Salmonella" knowledgebase for the mutational traits detection and mutation scanning features and the "Species_confirmation_Salmonella" knowledgebase for the Species Confirmation feature.

The *Manage knowledge bases* dialog box should now look like Figure 9.

7. Press <*Close*> to close the *Manage knowledge bases* dialog box.

ype	Name	Version	Checksum	Path	
cal	Disinfectant_resistance	1.0	5e379f9ec90c117f72812693b8a0d428	C:\Users\10029961\Desktop\Example knowledgebases\Cus	
ocal	In-silico PCR	1.0	9fd5ec8d5030dd79223e2bd3d5fa8d8c	C:\Users\10029961\Desktop\Example knowledgebases\Cus	
ocal	Mutational_resistance_Salmonella	1.0	9eda0cc7b0e4e2aff6cc57e8a460d4c2	C:\Users\10029961\Desktop\Example knowledgebases\Cus	
ocal	Species_confirmation_Salmonella	1.0	2f7e8ca5013f579965210c798b7a3d20	C:\Users\10029961\Desktop\Example knowledgebases\Cus	
					_

Figure 9: The Manage knowledge bases dialog box.

6 Creating a genotyping model

A genotyping model defines which genotyping analyses (i.e. features) will be executed as well as which knowledgebases and other feature settings will be used during the execution of the model. A user can create several models and switch between these models to change the active model. Only one model can be active, and hence be executed, at the same time.

 Select Genotyping > Manage models... to open the Manage genotyping models dialog box (see Figure 10).

Manage genotyping models								
Active Name Version Featur	es							
Create new Create update Import Export Delete								
Set as active Extract knowledgebases	CI	ose						

Figure 10: The Manage genotyping models dialog box.

The *Manage genotyping models* dialog box lists all genotyping models available in the BIONU-MERICS database, with their 'Name', 'Version' and the number of features present in the model ('Features'). Only one of the available models is the *active model*, i.e. the model that can be executed. Initially, this list shows up empty.

- 2. Press < *Create new...*> to start the *Create model* wizard (see Figure 11).
- Specify a *Model name* e.g. "MySalmonellaModel" and *Version* e.g. ""1.0". This combination should be unique for each model. Optionally, enter a *Description* for the model. Keep the default *Translation table* i.e.'Bacterial' (see Figure 12).



Do not use spaces in the model name or in the feature names as this interferes with the functionality of the BLAST-based features!

Create model		?	×
Model info Set the overal	l info and settings of the model		
Model name	[
Version			
Description			
	×		
Translation table	Bacterial V		
	< Back Next >	Can	cel

Figure 11: The Create model wizard.

Create model		?	×
Model info Set the overa	Il info and settings of the model		
Model name	MySalmonellaModel		
Version	1.0		
Description	^		
	~		
Translation table	Bacterial		
	< Back Next >	Can	cel

Figure 12: The Create model wizard.

4. Press < *Next* > to proceed to the second page of the *Create model* wizard.

In the second page of the *Create model* wizard, one or more features can be added to the genotyping model. Note that a model should contain at least one feature. We will add all features to our model except for the "sequence detection" feature as this functionality (i.e. the detection of sequences) is also included in the "Acquired traits detection" feature.

5. In the tree control, highlight the "Sequence extraction" feature and press the *Add...>* button.

The *New genotyping feature name* dialog box prompts you to enter a name for the new genotyping feature.

- 6. Enter a name (do not use spaces!) and press the *<OK>* button.
- 7. Repeat the previous actions to add the "Acquired traits detection", "Mutational traits detection", "Mutation scanning", "In-Silico PCR detection", "In-Silico PCR extraction" and "Species confirmation" features to the model.

The new features will be listed in the bottom part of the *Create model* wizard (see Figure 13).

Create model		?	\times
Model features Add features to the model			
BLAST Sequence detection Acquired traits detection Mutational traits detection Mutation scanning PCR In-Silico PCR detection In-Silico PCR detection In-Silico PCR extraction In-Silico PCR e	Minhash-based taxonomic confirmation genus, species and subspecies	n of	
Name	Feature		
Disinfectants_extraction	Sequence extraction		
Disinfectants	Acquired traits detection		
Antibiotics	Mutational traits detection		
Antibiotics_mutations	Mutation scanning		
Antigen_detection	In-Silico PCR detection		
Antigen_extraction	In-Silico PCR extraction		
Species_confirmation	Species Confirmation		
	< Back Next >	Cance	el

Figure 13: The Create model wizard.

8. Press the <*Next...*> button to start the *Create feature* wizard for each feature in the model (see Figure 14).

The *Knowledgebases* section is available for all genotyping features.

The **BLAST** section is available for all features using BLAST, i.e. sequence detection, sequence extraction, acquired traits detection, mutational traits detection and mutation scanning features. In the **BLAST** panel, two settings for the BLAST algorithm can be specified:

- *Minimum identity (%)* is the minimum sequence identity (as percentage) of the query sequence against the knowledge base's reference sequences.
- *Minimum length for coverage* specifies the minimum overlap (as percentage) between the subsequence found in the target assembly sequence and the reference sequence from the knowledge base.

Create model			?	×
Disinfectants Use BLAS above cer	T to detec	t and extract sequer	nces	
Knowledgeba	se			
Name	Disinfect	ant_resistance		
Version	1.0			
Change				
BLAST				
Minimum ident	ity (%)	95.0		
Minimum cove	rage (%)	95.0		
Combine fr	agments			
Extraction				
Sequence cor	rection	No correction		
		O CDS (conserva	ative)	
		O Trimming patter	rns	
	< Back	Next	Can	cel
	< back	wext	Can	Cer

Figure 14: The Create feature wizard.

If the option *Combine fragments* is checked, genes that occur fragmented in the genome (i.e. split over two contigs) can still be detected.

The *Extraction* section is only available for BLAST-based sequence extraction features. When mismatches occur at the edges of a query sequence, BLAST may return a truncated sequence to optimize the similarity score. Therefore, when the knowledgebase contains only a single allele or a very limited set of alleles for a certain gene, a *Sequence correction* might be needed. For the latter, one out of three options should be selected:

- *No correction*: The BLAST hit is taken as-is. This is the default option since no correction is needed when the knowledgebase covers sufficient diversity.
- **CDS** (conservative): The BLAST hit is extended to retrieve a full protein coding sequence (CDS), i.e. starting from the first encountered start codon upstream and ending at the first encountered stop codon downstream.
- *Trimming patterns*: The BLAST hit is extended and trimmed to length using the trimming patterns, present in the knowledgebase.
 - 9. Press <*Change...*> to display the *Manage knowledge bases* dialog box, from which a knowledgebase can be specified.
 - 10. For the "sequence extraction" feature select the "Disinfectant_resistance" knowledgebase from the list of knowledgebases and press the $\langle OK \rangle$ button.

When a knowledgebase is selected, its *Name* and *Version* are indicated (see Figure 14).

11. Press the *<Next...>* button to open the *Create feature* wizard of the next feature.

- 12. Repeat the previous steps to add the appropriate knowledgebase to each feature i.e. the "Disinfectant_resistance" knowledgebase for the "Acquired traits detection" feature, the "Mutational_resistance_Salmonella" knowledgebase for the "Mutational traits detection" and "Mutation scanning" features, the "In-Silico PCR" knowledgebase for the "In-Silico PCR detection" and "In-Silico PCR extraction" features and the "Species_Confirmation_Salmonella" knowledgebase for the "Species confirmation" feature.
- 13. Press < *Finish*> to complete the creation of the features in the model.

The *Create feature* wizard runs for each feature in the model. When the model creation is complete, the question "Do you want to set the new model as active?" pops up. In the *Custom geno-typing plugin*, only the active model can be executed, so typically you will want to answer < **Yes**> to this question.

14. Press < *Yes*>.

Next, the software asks "Do you want to modify the new model settings now?".

15. Press < **Yes**> to open the *Settings* dialog box.

The *Settings* dialog box pops up and consists of a general tab and a tab for each feature that was added to the model.



The settings for the active genotyping model in the *Custom genotyping plugin* can be accessed at any time via *Genotyping* > *Settings...* in the *Main* window.

7 Managing genotyping model settings

In the General tab the following general settings need to be specified:

- *Included info fields*: In this list the entry information fields that will be displayed in the genotyping report can be specified.
- *Exports directory*: With < *Browse...*> you can specify an export directory to store all exports from the genotyping reports.
- *Input Sequence experiment*: From the drop-down list you can specify the sequence experiment that holds the (whole) genome sequences that will be screened.
- **Enabled features**: This list contains all offered features of the genotyping plugin. Features which are not required can be disabled in this list to save on processing time and omit the corresponding sections from the report. By default, all features are enabled.
 - In our demonstration database, the assembled sequences are stored in the *denovo* sequence experiment. Make sure this experiment is selected from the drop-down list and optionally check the *Isolate* to include in the report (see Figure 15).

The other tabs group the settings for each possible search: Resistance (Acquired and mutational resistance), in-silico PCR detection and extraction and Species confirmation.

All tabs contain a *Knowledgebase* and *Results* panel and the BLAST-based features also include a *BLAST* panel:

1. *Knowledgebase*: in this panel the *Version* and *Name* of the knowledge base that is being used for this feature is shown.

Antibiotics_mutation						
		tigen_detection	Antigen_ext		Species_confirm	
General	Disinfe	ctants_extraction	Disi	nfectants	Antibioti	s
Reporting						
ncluded info fields	Or Se Re Ru	ST PubMLST Achtn ganism rovar_NCBI sistance genes_NE n_accession pProject		< >		
xports directory	C:\Use	ers\10029961\Docu	ments\BIONUM	ERICS8.1\W	Browse	
rocessing						
nput sequence exp	eriment	denovo			\sim	
nabled features		 Disinfectants. Disinfectants Antibiotics Antibiotics_m Antigen_dete Antigen_extra 	utations ction		~	

Figure 15: The Settings dialog box: General tab.

- BLAST: in this panel the two settings for the BLAST algorithm are specified; the *Minimum percent identity (%)* and the *Minimum coverage (%)* of your query sequence against the knowledge base's reference sequences.
- 3. Results: in this panel the output database information fields and experiments to which the screening results will be written can be dictated. Use the drop-down list to choose an existing experiment type or field, or the <*Create*> option to create new experiments and fields. A default name for the experiment or information field is suggested, but you can adjust this if you want to. Check *Annotate sequence experiment* to annotate the input sequence with the detected genotyping features.
 - In this tutorial, specify the experiment types and information fields in all tabs by selecting the <*Create*> option in the drop-down lists and accepting the default names. Leave the other settings unaltered.
 - 3. In the *Results* panel of the *Antigen_extraction* tab click on <*Change...*> next to *Sequence extraction* to open the *Change sequence experiment* dialog box. Click on <*Auto configure...*> to set the sequence experiments automatically and click on <*OK*>.
 - 4. In the *Results* panel of the *Disinfectants_extraction* tab click on <*Change...*> next to *Sequence extraction* to open the *Change sequence experiment* dialog box. In the drop-down list of the formA PCR target select the <*Create*> option and accept the default name. Click on <*OK*> to close the *Change sequence experiment* dialog box.
 - 5. Click on < OK > in the *Settings* dialog box to close the dialog box.
 - 6. Press <*Close*> to close the *Manage genotyping models* dialog box.

8 Running the active genotyping model

The screening can be done on any selection of entries in the database.

1. Select a single entry in the *Database entries* panel by holding the **Ctrl**-key and left-clicking on the entry. Alternatively, use the **space bar** to select a highlighted entry or click the ballot box next to the entry.

Selected entries are marked by a checked ballot box () and can be unselected in the same way.

2. In order to select a group of entries, hold the Shift-key and click on another entry.

A group of entries can be unselected the same way.

3. Make sure a few entries are selected in the *Database entries* panel of the demonstration database.

Screening selected entries with the active genotyping model can be done using *Genotyping* > *Run active*.

The analysis time increases proportionally with the number of selected entries and the number of genotyping features in the model. A complete analysis may take up to several minutes or even hours. The progress bar disappears when the analysis is finished.

4. Select *Genotyping* > *Run active* to start the screening of the selected entries.

The species confirmation results (*Species confirmation*) are written to the information field in the *Database entries* panel (see Figure 16). Please note that the shown name of the information field is the one that was created per default, but can be different in your case depending on whether you have chosen an alternative name.

Data	base entries				
ц _‡ 1] + 🖻	⊗ & ₽	<all entries=""></all>	U ا	
	Key	Modified date	Organism	Serovar_NCBI	Species confirmation
~	SRR3194565	2022-04-13 17:22:57	Salmonella enterica	derby	Salmonella enterica enterica
~	SRR1030845	2022-04-13 17:23:08	Salmonella enterica subsp. enterica	Dublin	Salmonella enterica enterica
~	SRR1183899	2022-04-13 17:23:18	Salmonella enterica subsp. enterica	Enteritidis	Salmonella enterica enterica
~	SRR1646564	2022-04-13 17:23:29	Salmonella enterica	Hadar	Salmonella enterica enterica
~	SRR1105667	2022-04-13 17:23:40	Salmonella enterica subsp. enterica	Heidelberg	Salmonella enterica enterica
~	SRR3289809	2022-04-13 17:23:53	Salmonella enterica	Heidelberg	Salmonella enterica enterica
~	SRR1574295	2022-04-13 17:24:04	Salmonella enterica	Illa 18:z4,z23:-	Salmonella enterica arizonae
~	SRR1534841	2022-04-13 17:24:15	Salmonella enterica	Kentucky	Salmonella enterica enterica
~	SRR3476365	2022-04-13 17:24:25	Salmonella enterica	Manhattan	Salmonella enterica enterica



The character experiment types for Resistance (Acquired and mutational resistance) and in-silico PCR screening are created and updated with the predicted traits. Please note that the shown names of the experiment types are those created per default, but can be different in your case depending on whether you have chosen an alternative name.

5. Open a character card for one of the analyzed entries by clicking on the corresponding green colored dot in the *Experiment presence* panel.



The characters in the characters experiments are displayed in the same order they are listed in their knowledge base. However, it might be more convenient for interpretation to have them displayed alphabetically. This can be done in the *Character type* window with the option *Characters* > *Arrange characters by field...* (\downarrow).

Below, the interpretation of the results gathered in the character experiment types is given.

Acquired traits detection (see Figure 17):

- **Disinfectants_traits**: contains the results for each disinfectant: 0 = not detected (sensitive), 1 = detected (resistant).
- **Disinfectants_detections**: contains the results for each resistance gene: 0 = not detected (sensitive), when detected (resistant) the % identity of the best hit is shown.

Character		Mapping	•	Character	Value	Mapping	-
Ciprofloxacin	0	<->		ClpL	0	<->	/
Nalidixic acid	0	<->		OqxB	0	<->	
Formaldehyde	1	<+>		OqxA	0	<->	
Chlorhexidine	0	<->		sitABCD	0	<->	
Chloramphenicol	0	<->		qacZ	0	<->	
Cetylpyridinium Chloride	0	<->		qacJ	0	<->	
Temperature	0	<->		qacA4	0	<->	
Ethidium Bromide	0	<->		qacH	0	<->	
Benzylkonium Chloride	0	<->		qacG	0	<->	
Hydrogen peroxide	0	<->		qacF	0	<->	
				qacE	0	<->	
				qacD	0	<->	
				qacC	0	<->	
				qacB	0	<->	
				qacA	0	<->	
				formA	100	<+>	

Figure 17: Example output of the *Disinfectants_traits* and the *Disinfectants_detections* experiment types for sample SRR1574259.

Sequence extraction (see Figure 18):

If sequence experiments have been created in the settings tab of the "sequence extraction" feature, the detected sequences are stored in the corresponding sequence type experiments.

- Click on the green colored dot of the **formA** sequence experiment for the entry with Key SRR1574259. The *Sequence editor* window opens and displays the extracted sequence (see Figure 18).
- 7. Close the Sequence editor window.

Mutational traits detection (see Figure 19):

- **Antibiotics**_**traits**: contains the results for each antibiotic: 0 = not detected (sensitive), 1 = detected (resistant).
- **Antibiotics_mutations**: contains the results for each known resistance mutation: -2 = partially indecisive, -1 = fully indecisive, 0 = not detected (sensitive), 1 = detected (resistant).

Mutation scanning (see Figure 20):

• **Antibiotics_mutations_mutations_all**: contains the results for each mutation: 1 = detected.

In-silico PCR detection (see Figure 21):

• **Antigen_detection_amplicons**: contains the result for the in-silico PCR: 0 = no amplicon, 1 = amplicon generated.

In-silico PCR extraction (see Figure 22):

SRR1574259 (Sequence Viewer)	-		×
File Sequence Header Annotation View Tools Window Help			
$\blacksquare \boxtimes \blacksquare \square \square \land \% \ \textcircled{a} \ \textcircled{b} \otimes \bigcirc \square \land \square \neg 2 \lor \otimes$			
Sequence Editor			
atgaaatcac gtgcagctgt agcatttgct cctggtaagc ccctcgagat cgttgaaatt gatgtggagc cgcctcgtaa	80		^
gggtgaagta ctggtaaaaa tcacccatac cggcgtctgc cacactgatg catttacctt gtccggtgat gatccggaag	160		
gtgtgttccc ggcagtactg ggtcatgaag gtgcgggtgt tgttgtggaa gtcggcgaag gggtcaccag tgtgaaacct			
ggcgatcatg ttattccgct ttacacggca gaatgcggcg agtgtctgtt ctgtaaatcc ggaaaaacta acctgtgtgt			
ctctgttcgc gccacccagg gtaaaggact tatgcctgat ggcacgaccc gtttctctta caagggccag cctcttttcc			
actatatggg ctgctctaca ttcagtgagt ataccgtcgt cgcagaagtg tctttagcca aaattaatcc acaggcgaat			
catgaacatg totgtotgot gggttgtggg gtaacgacag gtatcggtge ggttcacaac actgogaaag tacaaccagg			
tgacacggtt gctatttttg gcctgggtgg tattggcctt gcagcggtac agggcgcgcg tcaggcaaaa gcgggtcgta	040		×
Sequence Viewer			
			Ξ
<pre> '500' '1000 <</pre>		>	*
Contigs			
Start End Length			•
1 0 1110 1110			
Annotation Header Custom Fields Sequence Search Contigs Frame Analysis Restriction Analysis			
		_	
Sequence: SRR1574259 Experiment: formA 1 1110 bp			

Figure 18: Example output of the formA experiment type for the entry with Key SRR1574259.

SRR1574259				SRR1574259			
Character	Value	Mapping	-	Character	Value	Mapping	
Ciprofloxacin	0	<->		gyrB_pL447E	0	<->	
Nalidixic acid	0	<->		gyrB_pS464Y	0	<->	
Spectinomycin	0	<->		gyrB_pS464F	0	<->	
Azithromycin	0	<->		gyrB_pS464T	0	<->	
Colistin	0	<->		gyrB_pE466D	0	<->	
Nalidixic acid,Ciprofl	-2	<->		parC_pT57S	1	<+>	
				parC_pT66I	0	<->	
				parC_pG78D	0	<->	
				parC_pS80R	0	<->	
				parC_pS80I	0	<->	
				parC_pE84K	0	<->	
				parC_pE84G	0	<->	
				parE_pM438I	0	<->	
				parE_pE454G	0	<->	
				parE_pS458P	0	<->	
				parE_pH462Y	0	<->	
ress Insert to add chara	acter			Press Insert to add o	character		

Figure 19: Example output of the *Antibiotics_traits* and the *Antibiotics_mutations* experiment type for the entry with Key SRR1574259.

If sequence experiments have been created in the settings tab of the "in-silico PCR extraction" feature, the detected sequences are stored in the corresponding sequence type experiments.

- 8. Click on the green colored dot of the **fliCcom-fliCd** sequence experiment for the entry with Key SRR1574259. The *Sequence editor* window opens and displays the extracted sequence (see Figure 22).
- 9. Close the character and sequence card(s).

Character	/alue	Mapping	•
parC_dC170G	1	<+>	1
parC_dC369T	1	<+>	
parC_dC450G	1	<+>	
parC_dT672A	1	<+>	
parC_dT702C	1	<+>	
parC_dT708G	1	<+>	
parC_dT769C	1	<+>	
parC_dA783G	1	<+>	
parC_dA792G	1	<+>	
parC_dC825T	1	<+>	
parC_dT1161C	1	<+>	
parC_dC1170T	1	<+>	
parC_dC1305T	1	<+>	
parC_dC1479T	1	<+>	
parC_dT1617C	1	<+>	
parC_dC1806T	1	<+>	~

Figure 20: Example output of the *Antibiotics_mutations_mutations_all* experiment type for the entry with Key SRR1574259.

Character	Value	Mapping	
fliCcom-fliCa	0	<->	
fliCcom-fliCd	1	<+>	
viaB	0	<->	
prt	0	<->	
tyv	0	<->	

Figure 21: Example output of the *Antigen_detection_amplicons* experiment type for the entry with Key SRR1574259.

9 Reports

1. Open the genotype report for the selected entries with *Genotyping* > *Reports...*.

The *Report* window contains a genotype report for each of the selected entries (see Figure 23)).

2. Select another entry in the *Entries* panel to update the results in the *Genotype report* panel.

The creation date of the report (*Date*), the Key (*Name*), and information fields checked in the *Settings* dialog box are displayed in the *Genotype report* panel.

3. Select *Report > Report styles* in the *Report* window and make sure the option *Summary* is selected.

A summary of the results of all analyzed traits is displayed in the *Report* window.

4. Select *Report styles* in the *Report* window (see Figure 24) and select the option *Complete*.

In the *Complete* view, the summarized results as well as all available details are shown. All hits that passed the settings for resistance (Acquired and mutational resistance), in-silico PCR

SRR1574259 (Sequence Viewer) – – – – – – – – – – – – – – – – – – –
A CARACTER STATE S
atcaacac aacctgcagc gtgtgcgtga actggcggt cagtctgcta acggbacyaa ctcccagtct gaccttgact 80 tatccagge tgaaatcace cagegyetga acgaaatega eegggtga aactattgat attgattaa aagaaattag 240 tggegeagg acaacaeet gaccatecag gttggtgea acgaeggtga aactattgat attgattaa aagaaattag 240 tetaaaaca etgggaettg ataagettaa tgtecaagat geetaeaeee egaagaaae tgetgtaaee gttgataaaa 320 taeetaaa aaatggtaee gataetgtta eageeeagg caataetgat ategaaaet geetgetgetgetgetgetgetgetgetgetgetgetget
atcaacaac aacctgcagc gtgtgcgtga actggcggtt cagtctgcta acggbacyaa ctcccagtct gaccttgact 80 tatccaggc tgaaatcacc cagcgyctga acgaaatcga ccgtgtatcc ggycagactc agttcaacgg cgtgaaagtc 160 tggcgcagg acaacaccct gaccatccag gttggtgcca acgacggtga aactattgat attgattaa aagaaattag 240 tctaaaaca ctgggacttg ataagcttaa tgtccaagat gcctacacc cgaaagaac tgctgtaacc gttgataaaa 320 tacctataa aaatggtaca gatactgtta cagcccagag caatactgat atcgaaactg cattggcg tggtgcaacg 400 gggttactg gggctgatat caaatttaa gatggtcaat actatttaga tgttaaaggc ggtgcttctg ctggtgtta 480 aaagccact tatgatgaaa ctacaaagaa agttaatatt gatacgactg ataaaaaccc gttagcaact gcggaagca 560 agctattcg gggaacggcc actataaccc acaaccaaat tgctgaagta acaaaagagg gtgttgatac gaccacagtt 640
tatccaggc tgaaatcacc cagcgyctga acgaaatcga ccgtgttcc ggycagactc agttcaacgg cgtgaaagtc 160 tggcgcagg acaacaccct gaccatccag gttggtgcca acgacggtga aactattgat attgattaa agaaattag 240 tctaaaaca ctgggacttg ataagcttaa tgtccaagat goctacacce cgaaagaaac tgctgtaacc gttgataaaa 320 tacctataa aaatggtaca gatactgtta cagcccagag caatactgat atcgaaactg caattggcg tggtgcaacg 400 gggttactg gggctgatat caaatttaaa gatggtcaat actatttaga tgttaaagge ggtgcttctg ctggtgtta 480 aaagccact tatgatgaaa ctacaaagaa agttaatat gatacgactg ataaaactce gttagcaact gcggaagcta 560 agctattcg gggaacggce actataacce acaaccaaat tgctgaagta acaaaagagg gtgttgatac gaccacagtt 640
tggcgcagg acaacacct gaccatccag gttggtgcca acgacggtga aactattgat attgatttaa aagaaattag 240 tctaaaaca ctgggacttg ataagcttaa tgtccaagat gcctacaccc cgaaagaaac tgctgataac gttgataaaa 320 tacctataa aaatggtaca gatactgtta cagcccagag caatactgat atcgaaactg caattggcg tggtgcaacg 400 gggttactg gggctgatat caaatttaaa gatggtcaat actatttaga tgttaaaggc ggtgctctg ctggtgtta 480 aagccact tatgatgaaa ctacaaagaa agttaatatt gatacgactg ataaaactcg gtagcaact gcggaagcta 560 agctattcg gggaacggcc actataaccc acaaccaaat tgctgaagta acaaaagagg gtgttgatac gaccacagtt 640 quence Viewer
totaaaaca otgggaottg ataagottaa tgtocaagat gootacacoo ogaaagaaac tgotgtaaco gttgataaaa 320 taootataa aaatggtaca gatactgtta cagoocagag caatactgat atogaaactg caattggogg tggtgcaacg 400 gggttactg gggotgatat caaatttaaa gatggtoaat actatttaga tgttaaagoo ggtgottotg otggtgtta 480 aaagooact tatgatgaaa otacaaagaa agttaatatt gatacgaotg ataaaactoo gttagcaact goggaagota 560 agotattog gggaacggoo actataacoo acaaccaaat tgotgaagta acaaaagagg gtgttgatac gaccacagtt 640 equence Viewer
tacctataa aaatggtaca gatactgtta cagcccagag caatactgat atcgaaactg caattggcgg tggtgcaacg 400 gggttactg gggctgatat caaatttaaa gatggtcaat actatttaga tgttaaaggc ggtgcttctg ctggtgtta 480 aaagccact tatgatgaaa ctacaaagaa agttaatatt gatacgactg ataaaactcc gttagcaact gcggaagcta 560 agctattcg gggaacggcc actataaccc acaaccaaat tgctgaagta acaaaagagg gtgttgatac gaccacagtt 640 quence Viewer Image: Stage in the
gggttactg gggctgatat caaatttaaa gatggtcaat actatttaga tgttaaaggc ggtgcttctg ctggtgttta 480 aaagccact tatgatgaaa ctacaaagaa agttaatatt gatacgactg ataaaactcc gttagcaact gcggaagcta 560 agctattcg gggaacggcc actataaccc acaaccaaat tgctgaagta acaaaagagg gtgttgatac gaccacagtt 640 quence Viewer Image: See Stage
aaagccact tatgatgaaa ctacaaagaa agttaatatt gatacgactg ataaaactcc gttagcaact gcggaagcta 560 agctattcg gggaacggcc actataaccc acaaccaaat tgctgaagta acaaaagagg gtgttgatac gaccacagtt 640 quence Viewer Image: The second se
agctattcg gggaacggcc actataaccc acaaccaaat tgctgaagta acaaaagagg gtgttgatac gaccacagtt 640 quence Viewer Image: Table Image: Tab
aquence Viewer
200 400 600
٢
٢
ntigs
0 750 750
antelia Uradan Casta Sinta Casara Casara Casara Casara Antelia Destinia Antelia
notation Header Custom Fields Sequence Search Contigs Frame Analysis Restriction Analysis
ence: SRR1574259 Experiment: fliCcom-fliCd 1 750 bp

Figure 22: Example output of the fliCcom-fliCd experiment type for the entry with Key SRR1574259.

screening and species confirmation are listed and described.

- 5. Click on a hyperlink of one of the predicted traits to display the detailed results in the *Genotype report* panel (see Figure 25).
- 6. Select *File* > *Exit* to close the *Report* window.

For more detailed information on the genotyping analyses and interpretation of the reported results, please check the genotyping plugin manual.

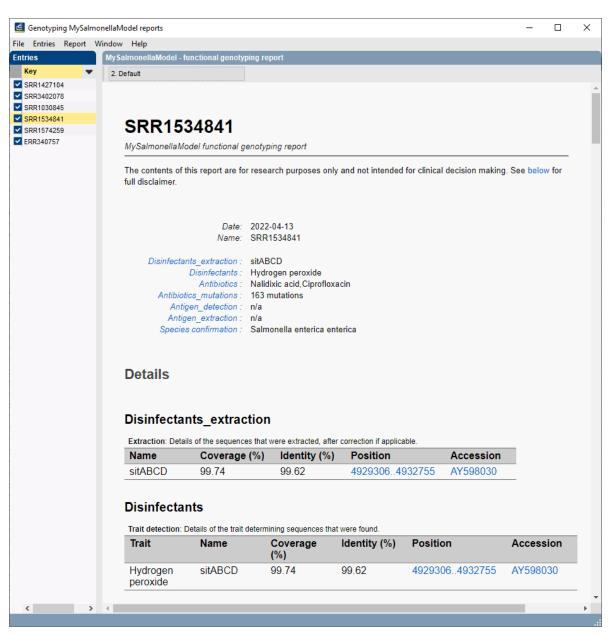


Figure 23: Example of a functional genotyping report.

E	. coli	- functional genotyping rep	ort
	1. Su	mmary	
	~	1. Summary	nce :
		2. Default	
		3. Complete	

Figure 24: Report styles in the Report window.

intries Report	Window	w Heln								
	Window	MySalmonellaModel - fu	actional genotyping re	nort						
y y	-	2. Default	redonal genotyping re	pore						
R1427104		2. Deiduk								
R3402078		Details								
1030845 1534841		Dotano								
R1534841 R1574259										
340757		Disinfectan	ts extractior							
		Disinicular		•						
				were extracted, after c		ible.		_		
		Name	Coverage (%)	Identity (%)	Position		Accession			
		formA	100.00	99.64	479397847	795087	X73835			
		Disinfectan	ts							
		Trait detection: Dr	tails of the trait determ	ining sequences that w	ioro found					
		Trait	Name	Coverage (%)		%) Pos	sition	Accessio	n	
						'	3978479508			
		Formaldehyde	s that are present, and	100.00 may express their trait	99.64 Related mutation				t may not be fully	
		Antibiotics All known mutation determined, check	s that are present, and the 'Requirements' coli	may express their trait umn.	. Related mutation	is that are a p	prerequisite to the	expression of a trait		ate
		Antibiotics All known mutation determined, check Trait	s that are present, and the 'Requirements' cole Locus	may express their trait umn. Level	Related mutation	is that are a p	prerequisite to the eference	expression of a trait	Requiremer	nts
		Antibiotics All known mutation determined, check	s that are present, and the 'Requirements' colu Locus parC	may express their trait umn.	. Related mutation	is that are a p	prerequisite to the eference	expression of a trait		nts
		Antibiotics All known mutation determined, check Trait Nalidixic acid, Ciproflox Antibiotics	s that are present, and the 'Requirements' coli Locus parC acin _mutations	may express their trait umn. Level	Related mutation Positio 57	is that are a p	prerequisite to the eference	expression of a trait	Requiremer	nts
		Antibiotics All known mutation determined, check Trait Nalidixic acid, Ciproflox Antibiotics	s that are present, and the 'Requirements' coli Locus parC acin _mutations	may express their trait urm. AA D the reference sequen	Related mutation Positio 57	is that are a p	prerequisite to the	expression of a trait	Requiremer	nts
		Antibiotics All known mutation determined, check Trait Nalidixic acid, Ciproflox Antibiotics All mutations that w	s that are present, and the 'Requirements' coli parC acin _mutations vere detected relative to	may express their trait urm. AA D the reference sequen	Related mutation Position 57 ce. Reference	ns that are a p n R T	prerequisite to the	expression of a trait	Requiremer	nts
		Antibiotics All known mutation determined, check Trait Nalidixic acid, Ciproflox Antibiotics All mutations that w Locus	s that are present, and the 'Requirements' colu parC acin _mutations vere detected relative to Level	may express their trait urm. Level AA b the reference sequen Position F	Related mutation Position 57 ce. Reference	is that are a p n R T T Mutatio	prerequisite to the	expression of a trait	Requiremer	nts
		Antibiotics All known mutation determined, check Trait Nalidixic acid, Ciproflox Antibiotics All mutations that w Locus acrB	s that are present, and the 'Requirements' colu parC acin _mutations vere detected relative to Level DNA	may express their trait imm. AA b the reference sequen Position F 82 7	Related mutation Position 57 ce. Reference	n R R T Mutatio C	prerequisite to the	expression of a trait	Requiremer	nts
		Antibiotics All known mutation determined, check Trait Nalidixic acid, Ciproflox Antibiotics All mutations that w Locus acrB acrB	s that are present, and the 'Requirements' colu parC acin _mutations vere detected relative to Level DNA DNA	may express their trait Imm. Level AA to the reference sequen Position F 82 119 204	Related mutation Position 57 ce. Reference	n R R T Mutatio C C	prerequisite to the	expression of a trait	Requiremer	nts
		Antibiotics All known mutation determined, check Trait Nalidixic acid, Ciproflox Antibiotics All mutations that v Locus acrB acrB acrB acrB	s that are present, and the 'Requirements' colu- parC acin _mutations vere detected relative to Level DNA DNA DNA DNA	may express their trait imm. AA the reference sequen Position F 82 1 119 1 204 1 291 0	Related mutation Position 57 ce. Reference T	n R R T Mutatio C C C C	prerequisite to the	expression of a trait	Requiremer	nts
		Antibiotics All known mutation determined, check Trait Nalidixic acid, Ciproflox Antibiotics All mutations that w Locus acrB acrB acrB acrB acrB	s that are present, and the 'Requirements' cold parC acin _mutations vere detected relative to Level DNA DNA DNA DNA	may express their trait Level AA Define reference sequen Position 82 119 204 291 381	Related mutation Position 57 ce. Reference T C	n R T T Mutatio C C C T	prerequisite to the	expression of a trait	Requiremer	nts
		Antibiotics All known mutation determined, check Trait Nalidixic acid, Ciproflox Antibiotics All mutations that w Locus acrB acrB acrB acrB acrB acrB	s that are present, and the 'Requirements' colu- parC acin _mutations vere detected relative to Level DNA DNA DNA DNA DNA DNA	may express their trait Level AA Define reference sequen Position 82 119 204 291 381	Related mutation Position 57 ce. Reference T C 3 3 3	n R T T Mutatio C C C C T A	prerequisite to the	expression of a trait	Requiremer	nts
		Antibiotics All known mutation determined, check Trait Nalidixic acid, Ciproflox Antibiotics All mutations that w Locus acrB acrB acrB acrB acrB acrB acrB acrB	s that are present, and the 'Requirements' colu- parC acin _mutations rere detected relative to Level DNA DNA DNA DNA DNA DNA DNA DNA	may express their trait Level AA b the reference sequen Position 82 119 204 381 60 819 01452	Related mutation Position 57 ce. Reference T C 3 3 3	n R T T Mutatio C C C T A A	prerequisite to the	expression of a trait	Requiremer	nts

Figure 25: Report details.