

# **BIONUMERICS** Tutorial:

# Resistance detection genotyping: predicting acquired resistance for Gram-negative bacteria

# 1 Aim

In this tutorial we will screen whole genome sequences of any *Gram-negative* bacterial samples for phenotypic antibiotic resistance traits using the *Resistance detection plugin*.

The different steps are illustrated using the whole genome demonstration database of *Escherichia coli*. This database is available for download on our website (see 2) and contains 60 publicly available sequence read sets of *Escherichia coli* with already calculated de novo assemblies.

# 2 Preparing the database

#### 2.1 Introduction to the demonstration database

We provide a **WGS demo database** for *Escherichia coli* containing sequence read set data links for 60 samples, calculated de novo assemblies and wgMLST results (allele calls and quality information).



The wgMLST workflow and results will not be discussed in this tutorial.

The **WGS\_demo\_database\_for\_Escherichia\_coli** can be downloaded directly from the *BIONU-MERICS Startup* window (see 2.2), or restored from the back-up file available on our website (see 2.2).

Installation of the *Resistance detection plugin* is only possible when no spaces are present in the BIONUMERICS home directory and in the name of the database. Before downloading or restoring the **WGS demo database** for *Escherichia coli*, please check if your BIONUMERICS home directory does not contain any spaces:

- 1. Click on the Obstration, located in the toolbar in the BIONUMERICS Startup window and select **Change home directory...** to call the Home directory dialog box.
- 2. In case the currently specified home directory contains spaces, update the path to a path containing no spaces and close the *Home directory* dialog box.

#### 2.2 Option 1: Download demo database from the Startup Screen

3. Click the  $\checkmark$  button, located in the toolbar in the *BIONUMERICS Startup* window.

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

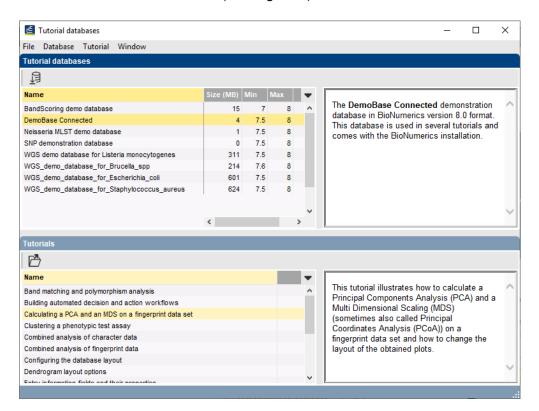


Figure 1: The *Tutorial databases* window, used to download the demonstration database.

- 4. Select WGS\_demo\_database\_for\_Escherichia\_coli from the list and select Database > Download ( ).
- 5. Confirm the installation of the database and press < *OK* > after successful installation of the database.
- 6. Close the *Tutorial databases* window with *File* > *Exit*.

The WGS\_demo\_database\_for\_Escherichia\_coli appears in the BIONUMERICS Startup window.

7. Double-click the **WGS\_demo\_database\_for\_Escherichia\_coli** 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 *Escherichia coli* is also available on our website. This backup can be restored to a functional database in BIONUMERICS.

8. Download the file WGS\_EC.bnbk file from https://www.bionumerics.com/download/ sample-data, under 'WGS\_demo\_database\_for\_Escherichia\_coli'.



In contrast to other browsers, some versions of Internet Explorer rename the WGS\_EC.bnbk database backup file into WGS\_EC.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.

- 9. In the *BIONUMERICS Startup* window, press the button. From the menu that appears, select **Restore database...**.
- 10. Browse for the downloaded file and select *Create copy*. Note that, if *Overwrite* is selected, an existing database will be overwritten.
- 11. Specify a new name for this demonstration database, e.g. "WGS\_Escherichia\_coli\_demobase".
- 12. Click < OK > to start restoring the database from the backup file.
- 13. Once the process is complete, click < Yes> to open the database.

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

	EIIII	h	0 ° °0	0,										
_	nent types				Data	base entries								Comparisons
-			<	<b>†</b>			86.16 2	<all entries=""></all>	1					
3	+ 🖒 🗞 🖡			Т -	£	1+0								
	≠ Name	Туре				Key	Organism	BioProject	BioSample	BioSampleModel	Isolation source	Center Name	🖵 <b>1</b> 2 3 4 5 6 7	Name Modified date
	1 wgs	Seque	ence read set typ			EC_0000001	Escherichia coli	PRJNA311907	SAMN04550946	Environment	Pizza flour mix	CFSAN	SRF • • • • • •	×
	2 wgMLST		scter types			EC_0000002	Escherichia coli	PRJNA311907	SAMN04497541	Environment	Food - Pizza Dough	CFSAN	SRF • • • • •	
	3 denovo		ence types			EC_0000003	Escherichia coli	PRJNA230969	SAMN04482456	Environment	dry powder	CFSAN	SRF • • • • • •	
	4 quality		acter types			EC_0000004	Escherichia coli	PRJNA311907	SAMN04550944	Environment	Pizza flour mix	CFSAN	SRF • • • • •	
	5 wgs_TrimmedState		ence read set typ			EC_0000005	Escherichia coli	PRJNA230969	SAMN04482457	Environment	dry powder	CFSAN	SRF • • • • •	
	6 wgMLST_CallType		scter types			EC_0000006	Escherichia coli	PRJNA230969	SAMN04429920	Environment	Flour	CFSAN	SRF • • • • •	
	7 wgsLong	Seque	ence read set typ			EC_0000007	Escherichia coli	PRJNA230969	SAMN04429921	Environment	Flour	CFSAN	SRF + + + + + +	
						EC_0000008 EC_0000009	Escherichia coli	PRJNA284275 PRJNA268206	SAMN05171462	Environment		CFSAN	SRF • • • • •	
						EC_0000009 EC_0000010	Escherichia coli Escherichia coli	PRJNA268206 PRJNA268206	SAMN04925767 SAMN04925766	Environment Environment	Product-Raw-Intact-Beef Product-Raw-Intact-Beef		SRF · · · · · ·	
						EC_0000010 EC_0000011	Escherichia coli	PRJNA268206 PRJNA268206	SAMN04925766 SAMN04908468	Environment	Product-Raw-Intact-Beef Product-Raw-Intact-Beef		SR0	
						EC_0000011 EC_0000012	Escherichia coli	PRJNA268206 PRJNA218110	SAMN04908468 SAMN04521095	Clinical	Product-Raw-Intact-Beat	EDLB-CDC	SRF • • • • •	
						EC_0000012 EC_0000013	Escherichia col	PRJNA210110 PRJNA230969	SAMN04521095 SAMN05289773	Environment	Unbleached White Flour	CFSAN	SRF • • • • •	<
						EC 0000014	Escherichia col	PRJNA230969	SAMN05294034	Environment	Enriched White Flour	CFSAN	SRF • • • • •	
1	ds Database design					EC_0000015	Escherichia coli	PRJNA218110	SAMN04521116	Cinical	Chinese Printer Paper	EDLB-CDC	586	Identification projects Decision networks
D	13 ⊗ 6.   6	2	† ↓			EC 0000016	Escherichia coli	PRJNA230969	SAMN05245392	Environment	All-Purpose Flour	CESAN	800	2日+円⊗風 品 ▽
		Field type				EC 0000017	Escherichia coli	PRJNA230969	SAMN05289777	Environment		CESAN	SR5 • • • • • •	Name Modified date
						EC 0000018	Escherichia coli	PRJNA230969	SAMN05215990	Environment		CESAN	585	Kame Modified date
	Organism	Fixed				EC 0000019	Escherichia col	PRJNA230969	SAMN04429924	Environment	Flour	CESAN	SRF • • • • • •	
		Fixed				EC_0000020	Escherichia coli	PRJNA230969	SAMN04429924	Environment	Flour	CFSAN	SRF • • • • • •	
	BioSample BioSampleModel	Fixed Fixed		_		EC_0000021	Escherichia coli	PRJEB11514	SAMEA3635259	Cattle	Cattle	UNIVERSITY OF ABE	ERR • • • • • •	
	Isolation source	Fixed		- 11		EC_0000022	Escherichia coli	PRJEB11514	SAMEA3835289	Clinical	Cinical	UNIVERSITY OF ABE	ERR • • • • • •	
	Center Name	Fixed				EC_0000023	Escherichia coli	PRJEB11514	SAMEA3635335	Clinical	Clinical	UNIVERSITY OF ABE	ERR • • • • • •	
	Run	Fixed				EC_0000024	Escherichia coli	PRJEB11514	SAMEA3635336	Clinical	Clinical	UNIVERSITY OF ABE	ERR • • • • • •	
	Experiment	Fixed				EC_0000025	Escherichia coli	PRJEB11514	SAMEA3635274	Clinical	Clinical	UNIVERSITY OF ABE	ERR • • • • • •	
	LoadDate	Fixed				EC_0000026	Escherichia coli	PRJEB11514	SAMEA3635334	Clinical	Clinical	UNIVERSITY OF ABE	ERR • • • • •	
	ReleaseDate	Fixed				EC_0000027	Escherichia coli	PRJEB11514	SAMEA3835339	Clinical	Clinical	UNIVERSITY OF ABE	ERR · · · · ·	
	SRA Sample	Fixed				EC_0000028	Escherichia coli	PRJEB11514	SAMEA3635209	Food/Environment	Food/Environment	UNIVERSITY OF ABE	ERR • • • • •	
	SRA Study	Fixed				EC_0000029	Escherichia coli	PRJEB11514	SAMEA3635210	Food/Environment	Food/Environment	UNIVERSITY OF ABE	ERR • • • • • •	
						EC_0000030	Escherichia coli	PRJEB11514	SAMEA3635256	Sheep	Sheep	UNIVERSITY OF ABE	ERR • • • • •	
į	nt files Power assembles	Anostation				EC_0000031	Escherichia coli	PRJNA277984	SAMN04510374	Environment	Bos taurus-raw mik	CFSAN	SR5 • • • • • •	Alignments BLAST projects Chromosome comparison
						EC_0000032	Escherichia coli	PRJNA277984	SAMN04510400	Environment	sol	CFSAN	SRF • • • • •	
	+ 🖒 ⊗ 🗟	6	Z. <ai< td=""><td></td><td></td><td>EC_0000033</td><td>Escherichia coli</td><td>PRJNA277984</td><td>SAMN04510408</td><td>Environment</td><td>soi</td><td>CFSAN</td><td>SRF • • • • •</td><td>+ 🖻 🛛 🗟 🗉 🔍 📣</td></ai<>			EC_0000033	Escherichia coli	PRJNA277984	SAMN04510408	Environment	soi	CFSAN	SRF • • • • •	+ 🖻 🛛 🗟 🗉 🔍 📣
	ile name t	xperiment t	ype Link	•		EC_0000034	Escherichia coli	PRJNA277984	SAMN04510394	Environment	water	CFSAN	SRF • • • • •	Name Modified date
						EC_0000035	Escherichia coli	PRJNA277984	SAMN04510376	Environment	Bos taurus-raw mik	CFSAN	SRF • • • • •	
						EC_0000036	Escherichia coli	PRJNA299490	SAMN04490312	Clinical	Bovine feces	CFSAN	SR5 • • • • •	
						EC_0000037	Escherichia coli	PRJNA299490	SAMN04490312	Clinical	Bovine feces	CFSAN	SRF • • • • •	
						EC_0000038	Escherichia coli	PRJNA230969	SAMN04490950	Environment	Bulk Flour	CFSAN	SRF • • • • •	
						EC_0000039 EC 0000040	Escherichia coli	PRJNA218110 PRJNA268206	SAMN04549544	Clinical	Stool	EDLB-CDC USDA-FSIS	SRF • • • • • •	
						EC_0000040 EC_0000041	Escherichia coli Escherichia coli	PRJNA268206 PRJNA268206	SAMN04942588 SAMN04621601	Environment Environment	Comminuted Beef Comminuted Beef	USDA-FSIS USDA-FSIS	SRF * * * * * *	
						EC_0000041 EC_0000042						USDA-FSIS UNIVERSITY OF ABE	SRF • • • • • •	
						EC_0000042 EC_0000043	Escherichia coli Escherichia coli	PRJEB11514 PRJEB11514	SAMEA3635202 SAMEA3635203	Food/Environment Food/Environment	Food/Environment Food/Environment	UNIVERSITY OF ABE	ENH EDD	
						EC_0000043 EC_0000044	Escherichia coli	PRJEB11514 PRJEB11514	SAMEA3635203 SAMEA3635247	Clinical	Cinical	UNIVERSITY OF ABE	FOD	
						EC_0000044 EC_0000045	Escherichia coli Escherichia coli	PRJEB11514 PRJEB11514	SAMEA3635247 SAMEA3635239	Food/Environment	Food/Environment	UNIVERSITY OF ABE	FRR · · · · · ·	
						EC_0000045 EC_0000046	Escherichia coli	PRJEB11514 PRJEB11514	SAMEA3635239 SAMEA3635316	FoodEnvironment	FoodEnvironment	UNIVERSITY OF ABE	ERR · · · · ·	
						<	Escretiona col	PRJED11514	SAMEA3635316	roowenvironment	roouchvironment	UNIVERSITY OF ABE	ERH .	

Figure 2: The Escherichia coli demonstration database: the Main window.

## 3 About the demonstration database

The **WGS\_demo\_database\_for\_Escherichia\_coli** contains data for a set of 60 samples. The sample information, stored in entry info fields (Isolation source, Center Name, etc.) was collected from the publications.

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

ଡ୍ୟ	#	+ 🚰 ⊗ 🔩   Name	🛍 🖳 <b>1</b> Туре	+	<all experiment="" types=""></all>	ט   ער די
⊒ ≩	1	wgs	Sequence read set types			
	2	wgMLST	Character types			
🗆 🛟	3	denovo	Sequence types			
	4	quality	Character types			
⊒≩	5	wgs_TrimmedStats	Sequence read set types			
	6	wgMLST_CallTypes	Character types			
_ ≹	7	wgsLong	Sequence read set 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 *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).

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

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 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 resistance detection plugin

1. Call the *Plugins* dialog box from the *Main* window by selecting *File* > *Install / remove plugins...* (, C=).

≝ EC_0000001	– 🗆 X								
File Preprocessing Analysis Window Help									
Sequence read set report									
Sequence read set information									
- Storage									
Storage by link: NCBI SRR3242188									
- Read set size	Help 88 5: 469412 5: 469412 5								
Number of sequences: 469412 Number of paired-end sequences: 469412									
Number of bases: 223023791, 1st end 1114	487506, 2nd end 111536283								
<ul> <li>Sequence length statistics</li> </ul>									
Average sequence length: 237.56, 1st end 2									
Standard deviation of the sequence length: 3 Minimum sequence length: 35, 1st end 35, 2									
Maximum sequence length: 251, 1st end 25									
- Quality statistics									
Average base quality: 35.28, 1st end 36.52, 2									
Minimum base quality: 2, 1st end 2, 2nd end									
Maximum base quality: 40, 1st end 40, 2nd e									
- Base statistics									
Number of bases A: 54471733 (24.42%)									
Number of bases C: 56917164 (25.52%) Number of bases G: 57060245 (25.58%)									
Number of bases T: 54560544 (24.46%)									
Number of other bases: 14105 (< 0.01%) Number of bases GC: 113977409 (51.11%)									
Analyses	Analysis info								
💾 i 🛇	Field name Field value 🗸								
Analysis name Analysis type 🗸 🗸									
^									
Ý	< >								

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

- 2. Select the *Resistance detection plugin* in the *Applications tab* and press the <*Activate*> button (see Figure 6).
- 3. Confirm the installation of the plugin.

During installation, the plugin downloads online knowledge bases from https://www.bionumerics.com, which requires a connection to the internet.

4. Click on < *Yes*> to review the settings.

The *Resistance detection settings* dialog box pops up, consisting of 2 tabs (see Figure 7). 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.

EC_0000001 (Sequence Viewer)		
File Sequence Header Annotation View Tools Window Help		
Sequence Editor		
TIGGATATEGT TAAAGAGGAT ACAGGCACCG TTGAAGCCCA CCTTACGCTC GGAAACCTGT TCCGTTCGCG TGGCGAAGTT GATCGCGCT TTCGCATCCA TCAGACCTA ATGGAAAGCG CCTCGCTGAC CTATGAACGG GGTCGTGGG GGATCAACA ACTAGGGCGT GATACATG CCGCCGGGTT ATACGACCGC GCGGAAGACA TGTTCAATCA GCTGGACGAT GGAACTGGCT TCCGCATTGG CGCTCTGCAA CAGTGCTAA AAATCTACCA GGCTACCAT GAGTGGCAGA AAGCAATTGA TGTTGCCGAA CGCCTGGTGA AGCTGGGTAA AGATAAACAG CGCGCCGAA TTGCCCATT CTACTGTGAG TTAGCCCTGC AGCATATGGC CAGCGACGT CTCGATCGTG CCATGACCTT GCTAAAAAA GGGCGGCG CAGATAAAAA CAGCCCCCGC GTATCCATCA TGATGGCAG CGTGTTTATG CGAAACGAG AATACCGCAA AGCCGTGAA AGTTGCCAA GCGTGATATC CCAGGACAG GACTGGTCA CGGAAGCGC GGAATGTG CGAAACGAG AATACCGCCA CGCCGCGAAAGAA CAGCGCGCACA GGTGGAAGAGA ACACCGGTGC CGATGCTGAA TTGAGCCTG GCAAAAGAA GGGCGGCG CAGAATTCC GCAACGCGCG GTGGAAGAGA ACACCGGTC CGAAGCGG ATGGTGGTC CTGATACAT CGAAGCGCG GACGGTGT CAGAATTCC GCAACGCGG GTGGAAGAGA ACACCGGTC CGAAGCGG CGCGGAAGGG AATACCGCT GGAAAGAA CGGGAGAAGG GCCTATATT ACGGCCACC TCAGCGTCA TCCGACCATG CGTGGTCC ATAAGTAAT GGATACCAC TTAAATGAA CGGAAAGAGG GCTGCCAAA GAGAGCCTGG GACGGTGGT GCCGACAGGG TCCGAATGGA CCACTGTAT CGAGCCCGC GACGGTGC AAGGCGCACA GGTCCCAAA GAGACCCTGG TGCGCG GCCGATGGG TGCCGAATAGG TACCTGCTAT CGAGCCCCA CGGAAGAGG CGCCGCCAA GGACCCTGAT GGCCTCGC GGCCTGGTCG TGCCGAAGGA CCATTAAACC GGAAGAGG GCCGCCAAA GAGCCCTGAT GGCATGCCC GCTTGTCGG GCCTGGTCA CATTAAACC GATTCGCGGT CTGATGGCC	G 8840 C 8730 A 8820 G 8910 C 9000 G 9090 G 9180 G 9270 A 9360	^
IGTAAttttt aaaaaaatcc tactttagtt acaacatact agttattgat gttccattgt gctccggcaa cgacggcgca gaaaaagcc		
gccaggggaa aaatcgcaac tgttaatttt ttatttccac gggtagaatg ctcgccgttt acctgtttcg cgccacttcc ggtgcccat	c 9630	~
Sequence Viewer		
	E E EC	12 EC_12399
Annotation		
Feature list Q   & O   ] & O		
		^
Feature key Start End Length %GC V 82869455		
→         M         4         →         CDS         4267         6234         1968         52.01         ^         /allele="10"           /locus_tag="EC_746" <td></td> <td></td>		
✓         4         ↓ ♦ CDS         4287         6234         1968         5201         //allele="10"           →         5         ↓ ♦ CDS         6298         6888         591         50.17         //outgag="EC_746"		
✓         ✓         CDS         4.267         6.234         1968         5.201         //allela="10"           ✓         5         ↓         CDS         6288         6888         591         50.17         //cute_tag="EC_746"           ✓         6         ↓         CDS         7058         7822         765         50.52         //note="fixed-1;start=4050;stop=5220;cid=denovo_1"           ✓         7         ↓         CDS         7971         8279         309         50.32         //raniation="fixed-1;start=4050;stop=520;cid=denovo_1"		
→         A         →         CDS         4267         6234         1968         5201         //allele="10"           →         5         ↔         CDS         6288         591         50.17         //ocus_tag="EC_746"           →         6         ↔         CDS         765         50.52         //note="twd=1; start=4050; stop=5220; cid=denovo_1"           →         7         ↔         CDS         8286         9455         1170         51.84	TYEORLLAIQOLGR	DYMAA
→         ↓	TYEQRLLAIQQLGRI DVAERLVKLGKDKQI RVFMAKGEYAKAVE:	DYMAA RVEIA SLQRV
→       ↓	TYEQRLLAIQQLGRI DVAERLVKLGKDKQI RVFMAKGEYAKAVE ADAELMLADIIEARI	DYMAA RVEIA SLQRV DGSEA
→       ↓       ↓       CDS       4267       6234       1968       52.01       //allele="10"         →       ↓       ↓       CDS       4267       6234       1968       52.01       //ocus_tag="EC_746"         →       ↓       ↓       ↓       CDS       628       6888       591       50.17       //ocus_tag="EC_746"         ↓       ↓       ↓       ↓       ↓       ↓       0.05       7058       7822       765       50.52       //note="fwd=1;start=4050;stop=5220;cid=denovo_1"         ↓       ↓       ↓       ↓       ↓       CDS       7971       8279       309       50.32       //note="fwd=1;start=4050;stop=5220;cid=denovo_1"         ↓ <td>TYEQRLLAIQQLGRI DVAERLVKLGKDKQI RVFMAKGEYAKAVE ADAELMLADIIEARI</td> <td>DYMAA RVEIA SLQRV DGSEA</td>	TYEQRLLAIQQLGRI DVAERLVKLGKDKQI RVFMAKGEYAKAVE ADAELMLADIIEARI	DYMAA RVEIA SLQRV DGSEA

Figure 5: The Sequence editor window.

Plugins				?	×
Applications	Utilities	Database Functiona	lity		
Activate func	tionality fo	or specific application	S.		
BandSco E. coli fu Custom d Listeria t MTBC fu Resistan Salmone	Inctional g Genotypin functional Inctional g Ince detect Illa function is function istance ITR Illine	enotyping genotyping enotyping ion nal genotyping nal genotyping	< <	Find resistance traits using one of the online knowledgebases.	
Show Manu	al				
				Activate	Exit



• **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.

	Resistance						
	resistance de	etection.					
Report	ing						
Included info fields			ganism Decised				
			Project Sample				
			SampleModel				
			lation source				
Exports directory		Ce	nter Name	*			
		C:\Use	Brov	vse			
input 3	equence expe	annon	denovo		~		
Enable	d features		Resistance				

Figure 7: The Resistance detection settings dialog box: General tab.

5. 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 *BioSample* to include in the report (see Figure 7).

The other tab groups the settings for the search of acquired resistance traits. It contains a *Knowl-edgebase*, *BLAST* and *Results* panel:

- Knowledgebase: in this panel the Version and Name of the knowledge base that is being used for this feature is shown. A different knowledge base version can be selected by pressing the < Change...> button. With Check for updates on startup checked, BIONUMERICS will check if a newer knowledge base version is available online for this feature each time the database is opened.
- BLAST: in this panel two settings for the BLAST algorithm can be specified; the Minimum percent identity (%) and the Minimum coverage (%) of your query sequence against the knowledge base's reference sequences. If the option Combine fragments is checked, genes that occur fragmented in the genome (i.e. split over two contigs) can still be detected.
- 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.
  - 6. In this tutorial, specify the experiment types and information fields in the *Resistance* tab by selecting the *<Create*> option in the drop-down lists and accepting the default names. Leave the other settings unaltered.

- 7. Click on *<OK>* in the *Resistance detection settings* dialog box.
- 8. When the *Resistance Detection plugin* is successfully installed, a confirmation message pops up. Press <*OK*>.
- 9. Press < *Exit*> to close the *Plugins* dialog box.
- 10. Close and reopen the database to activate the features of the Resistance Detection plugin.

The *Resistance Detection genotyping plugin* installs menu items in the main menu of the software under *Resistance detection* (see Figure 8).

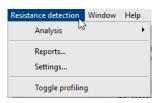


Figure 8: New menu-items after installation of the Resistance Detection plugin.



The settings specified during installation of the plugin can be called again at any time with *Resistance detection* > *Settings...*.

### 5 Screening of entries

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 for the acquired resistance can either be done for all tools checked in the *Resistance detection settings* dialog box (using *Resistance detection > Analysis > All Enabled*) or by *Resistance detection > Analysis > Resistance*.

4. Select *Resistance detection* > *Analysis* > *All Enabled* to screen the selected entries for all enabled traits.

A progress bar appears. The analysis time depends on the number of selected entries. When the analysis is finished, the progress bar disappears. The detected traits for the screened entries are stored in the database.

The character experiment types for **Resistance** are created and updated with the predicted acquired resistance 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 during installation.

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 resistance (see Figure 9):

- Resistance\_traits: contains the results for each antibiotic: 0 = not detected (sensitive), 1 = detected (resistant).
- **Resistance\_loci**: contains the results for each resistance gene: 0 = not detected (sensitive), when detected (resistant) the % identity of the best hit is shown.

Character	Value	Mapping	•	Character	Value	Mapping	
Unknown Beta-lactam	0	<->	^	vgbA	0	<->	-
linezolid	0	<->		vgb(B)	0	<->	
Netilmicin	0	<->		vgb(A)	0	<->	
Telithromycin	0	<->		vgaB	0	<->	
tetracycline	0	<->		vgaA	0	<->	
neomycin	0	<->		vga(E)	0	<->	
Lincomycin	0	<->		vga(D)	0	<->	
Streptomycin	0	<->		vga(C)	0	<->	
Unknown Fluoroquin	1	<+>		vga(B)	0	<->	
Butirosin	0	<->		vga(A)V	0	<->	
Ampicillin	0	<->		vga(A)LC	0	<->	
clindamycin	0	<->		vga(A)	0	<->	

Figure 9: Example output of the *Resistance\_traits* and the *Resistance\_loci* experiment types for sample EC\_0000012.

6. Close the character card(s).

### 6 Reports

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

The *Report* window contains a genotype report for each of the selected entries (see Figure 10).

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 *Resistance detection settings* dialog box are displayed in the *Genotype report* panel.

3. Select *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* > *Report styles* in the *Report* window (see Figure 11) 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* screening 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 12).

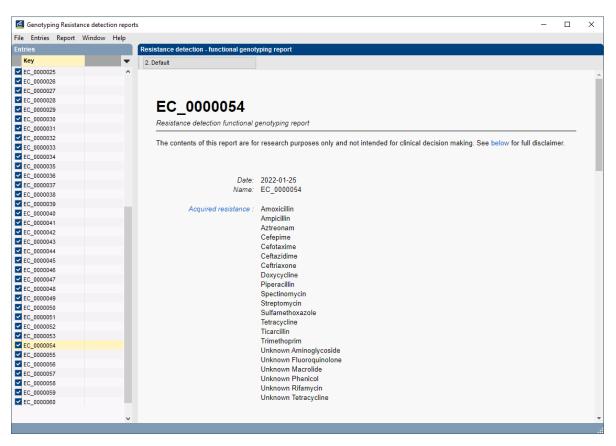


Figure 10: Example of a resistance detection report.

2. De	fault		
	1. Summary		
~	2. Default		
	3. Complete	13	

Figure 11: Report styles in the Report window.

6. Select *File* > *Exit* to close the *Report* window.

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

le Entries Report Window I	Help						
ntries		Resistance detection - function	al genotyping report				
Key	•	2. Default					
EC_0000025 EC_0000026	^	Resistance					
EC 0000027		Acquired resistance resu	ilte				
EC_0000028 EC_0000029		Trait	Locus	Coverage	Identity (%)	Position	Accession
EC_0000030				(%)			
EC_0000031		Unknown Macrolide	mdf(A)	100.00	97.65	23913022390070	Y08743
EC_0000033 EC_0000034		Unknown Aminoglycoside	mdf(A)	100.00	97.65	23913022390070	Y08743
EC_0000035 EC_0000036 EC_0000037		Unknown Tetracycline	mdf(A)	100.00	97.65	23913022390070	Y08743
EC_0000038		Unknown Fluoroquinolone	mdf(A)	100.00	97.65	23913022390070	Y08743
EC_0000040		Unknown Phenicol	mdf(A)	100.00	97.65	23913022390070	Y08743
EC_0000041 EC_0000042		Unknown Rifamycin	mdf(A)	100.00	97.65	23913022390070	Y08743
<pre> EC_0000043 EC 0000044 </pre>		Trimethoprim	dfrA17	100.00	100 00	255479255952	FJ460238
EC_0000045		Spectinomycin	aadA5	100 00	100.00	256083256871	AF137361
EC_0000046		Streptomycin	aadA5	100.00	100.00	256083256871	AF137361
EC_0000047							
EC_0000048		Streptomycin	aph(3")-lb	100.00	100.00	34147053415508	AF321551
EC_0000049 EC_0000050		Streptomycin	aph(6)-ld	100.00	100.00	34155143416344	CP000971
EC_0000050		Sulfamethoxazole	sul1	100.00	100.00	257418258257	U12338
EC_0000052 EC_0000053		Sulfamethoxazole	sul2	100.00	100.00	34138293414642, 34146813414718	FJ197818
EC_0000054		Amoxicillin	blaCTX-M- 27	100.00	100.00	36862473685372	AY156923
EC_0000056 EC_0000057 EC_0000058		Ampicillin	blaCTX-M- 27	100.00	100.00	36862473685372	AY156923
EC_0000059 EC_0000060		Aztreonam	blaCTX-M- 27	100.00	100.00	36862473685372	AY156923
-		Cefepime	blaCTX-M-	100.00	100.00	36862473685372	AY156923

Figure 12: Report details.