

Genotypic Characterization of a New Reference Set of Gram-Negative Bacteria Utilizing a Novel Annotation Pipeline

Juan Lopera¹, Ph.D., Xiaochuan Liu², Ph.D., Quan Zhang², Ph.D., Jung-Woo Son², Ph.D. Wujun Zhan², M.Sc., Joyce Sutcliffe², Ph.D., Cynthia Long¹, B.A., and Raul Cano², Ph.D.

¹ATCC, Manassas, VA

²ATCC-CTM, Union, NJ

Background

According to the World Health Organization, antimicrobial resistance (AMR) among Gram-negative bacteria continues to increase on a global scale¹. It is estimated that more than 23,000 people in the U.S. alone die each year from infections with multidrug-resistant (MDR) bacteria². New therapeutic agents are critical to stem this trend, but new technologies are required for shortening the time from discovery to production. To support this effort, ATCC has developed a collection of 33 fully characterized Gram-negative isolates representing current MDR disease strains from around the globe. Strains were evaluated using whole genome sequencing (WGS) and a novel annotation program to identify AMR genes and protein targets. Using public databases, ATCC developed and validated an accurate and efficient bioinformatics pipeline for the automated assembly and annotation of microbial genomes. Next-generation sequencing (NGS) data from all 33 individual isolates in combination with the novel bioinformatics pipeline were used to identify AMR genes and predictive targets that could be associated with the observed phenotype. Using our proprietary bioinformatics pipeline, we created a searchable database of AMR determinants containing the WGS information as well as a list of known AMR genes with their corresponding nucleotide sequences.

Methods

ATCC® Clinical Isolates Collection:

Table 1. Strains available in the ATCC® Clinical Isolates Collection

ATCC® No.	Species	Strain	Isolation Source
BAA-2774™	<i>Escherichia coli</i>	1093925	50-year-old male, abscess, general surgery, Argentina, Latin America, 2014
BAA-2775™	<i>Escherichia coli</i>	1099675	75-year-old male, decubitus, Czech Republic, 2014
BAA-2776™	<i>Escherichia coli</i>	1099705	78-year-old female, urine, general medical clinic, Czech Republic, 2014
BAA-2777™	<i>Escherichia coli</i>	1106769	40-year-old female, urine, general surgery, Nigeria, 2014
BAA-2778™	<i>Escherichia coli</i>	1123005	46-year-old female, urine, emergency room, Nigeria, 2014
BAA-2779™	<i>Escherichia coli</i>	1136922	83-year-old male, urine, emergency room, Israel, 2014
BAA-2780™	<i>Escherichia coli</i>	1138276	70-year-old female, abscess, general medicine, Mexico, 2014
BAA-2781™	<i>Escherichia coli</i>	1144993	52-year-old female, blood, medicine ICU, Turkey, 2014
BAA-2782™	<i>Klebsiella pneumoniae</i>	931476	55-year-old female, peritoneal fluid, general surgery, Brazil, 2013
BAA-2783™	<i>Klebsiella pneumoniae</i>	972863	35-year-old male, large colon, surgery ICU, Serbia, 2013
BAA-2784™	<i>Klebsiella pneumoniae</i>	1147843	53-year-old female, urine, medicine ICU, United States, 2014
BAA-2785™	<i>Klebsiella pneumoniae</i>	BK33650	United States, 2010
BAA-2786™	<i>Klebsiella pneumoniae</i>	BK34774	Deep wound, United States, 2010
BAA-2787™	<i>Klebsiella pneumoniae</i>	BK34786	Urine, United States, 2010
BAA-2788™	<i>Klebsiella pneumoniae</i>	BK34907	Urine, United States, 2011
BAA-2789™	<i>Klebsiella pneumoniae</i>	BK35106	Blood, United States, 2012
BAA-2790™	<i>Klebsiella pneumoniae</i>	BK35639	Wound, United States, 2012
BAA-2791™	<i>Proteus mirabilis</i>	927889	79-year-old male, urine, general medical clinic, Greece, 2013
BAA-2792™	<i>Proteus mirabilis</i>	1104843	50-year-old male, urine, general medicine, Greece, 2014
BAA-2793™	<i>Pseudomonas aeruginosa</i>	1077994	26-year-old female, urine, general medicine, Chile, 2014
BAA-2794™	<i>Pseudomonas aeruginosa</i>	1079232	28-year-old female, abscess, general medicine, Thailand, 2014
BAA-2795™	<i>Pseudomonas aeruginosa</i>	1106432	70-year-old male, urine, general medicine, Kenya, 2014
BAA-2796™	<i>Pseudomonas aeruginosa</i>	1106434	32-year-old male, blood, general medicine, Kenya, 2014
BAA-2797™	<i>Pseudomonas aeruginosa</i>	1109196	51-year-old male, endotracheal aspirate, surgery ICU, Colombia, 2014
BAA-2798™	<i>Pseudomonas aeruginosa</i>	1124989	34-year-old male, sputum, general medicine, Spain, 2014
BAA-2799™	<i>Pseudomonas aeruginosa</i>	1145009	55-year-old male, blood, medicine ICU, Turkey, 2014
BAA-2800™	<i>Acinetobacter baumannii</i>	1074318	57-year-old male, abscess, general surgery, Spain, 2014
BAA-2801™	<i>Acinetobacter baumannii</i>	1103521	36-year-old female, endotracheal aspirate, medicine ICU, Venezuela, 2014
BAA-2802™	<i>Acinetobacter baumannii</i>	1109222	37-year-old male, blood, Colombia, 2014
BAA-2803™	<i>Acinetobacter baumannii</i>	1125209	73-year-old male, sputum, general medicine, Italy, 2014
BAA-2806™	<i>Enterobacter cloacae</i>	976689	33-year-old male, urine, general medicine, Russia, 2013
BAA-2807™	<i>Citrobacter freundii</i>	990954	54-year-old male, skin ulcer, general surgery, Brazil, 2013
BAA-2808™	<i>Serratia marcescens</i>	938677	72-year-old male, urine, Romania, 2013

Development of a Bioinformatics Platform:

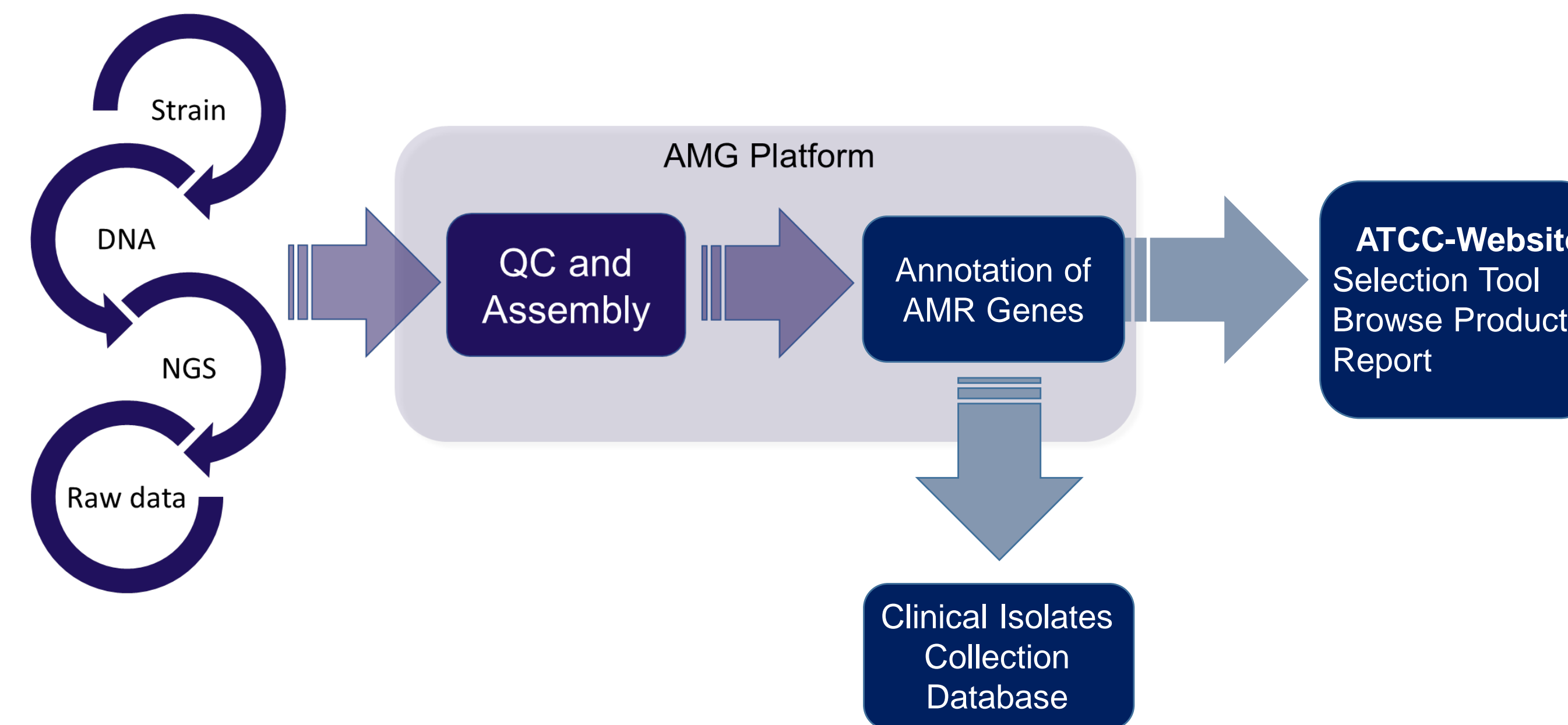


Figure 1. Genome sequencing and data analysis workflow. DNA was extracted from isolates and sequenced. NGS data from all isolates were quality controlled, assembled, and annotated using a proprietary novel bioinformatics platform to identify predicted AMR genes and antibiotics targets.

Results

Automated Assembly and Annotation:

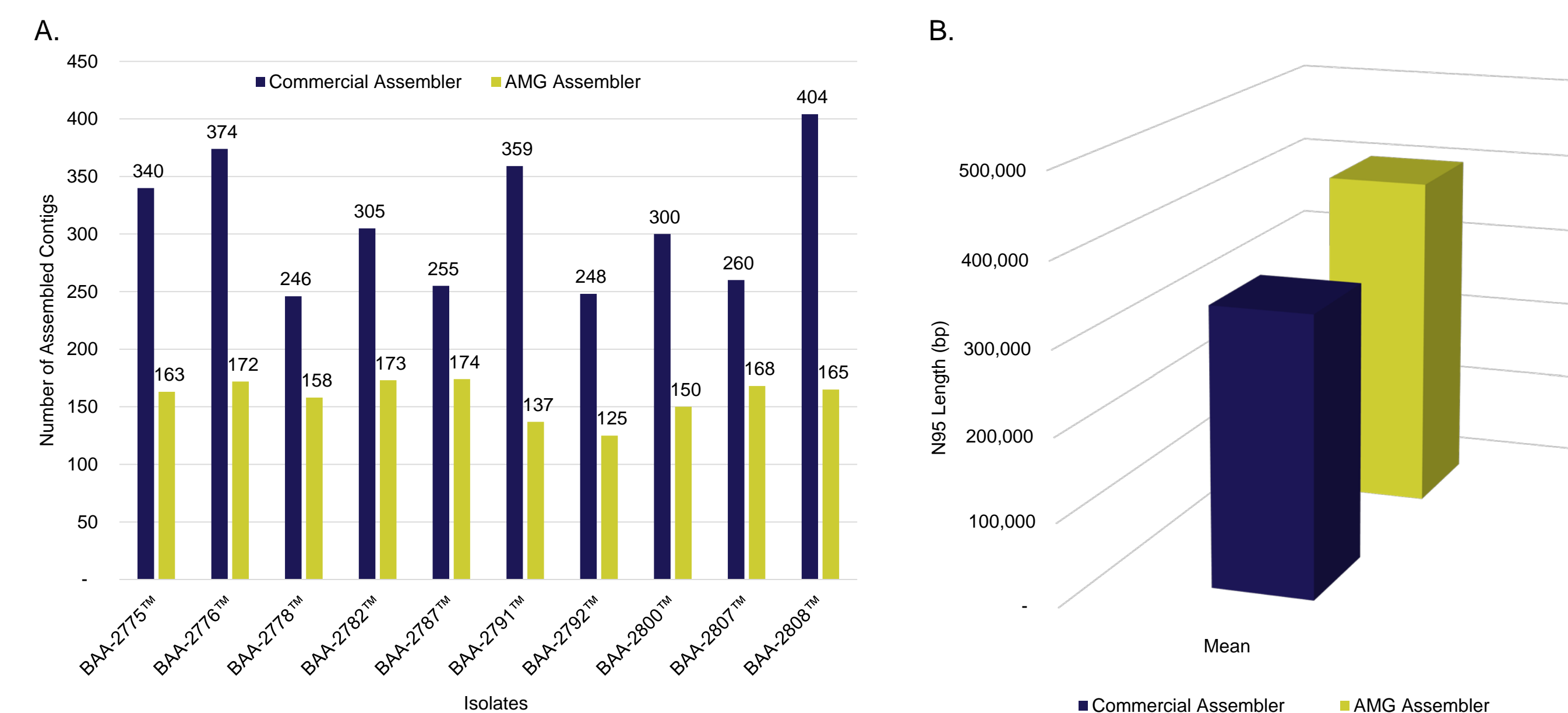


Figure 2. Efficiency and accuracy of genome assembly. The A) number of assembled contigs and the B) length of contigs generated were assessed for both the commercial assembler and the Advanced Microbial Genomics (AMG) assembler. The bioinformatics pipeline was able to successfully assemble multiple datasets per day compared with a licensed workbench. The assembly from the AMG platform produced a lower number of contigs that were longer in length, which improves the quality during ORFs call and AMR gene annotation.

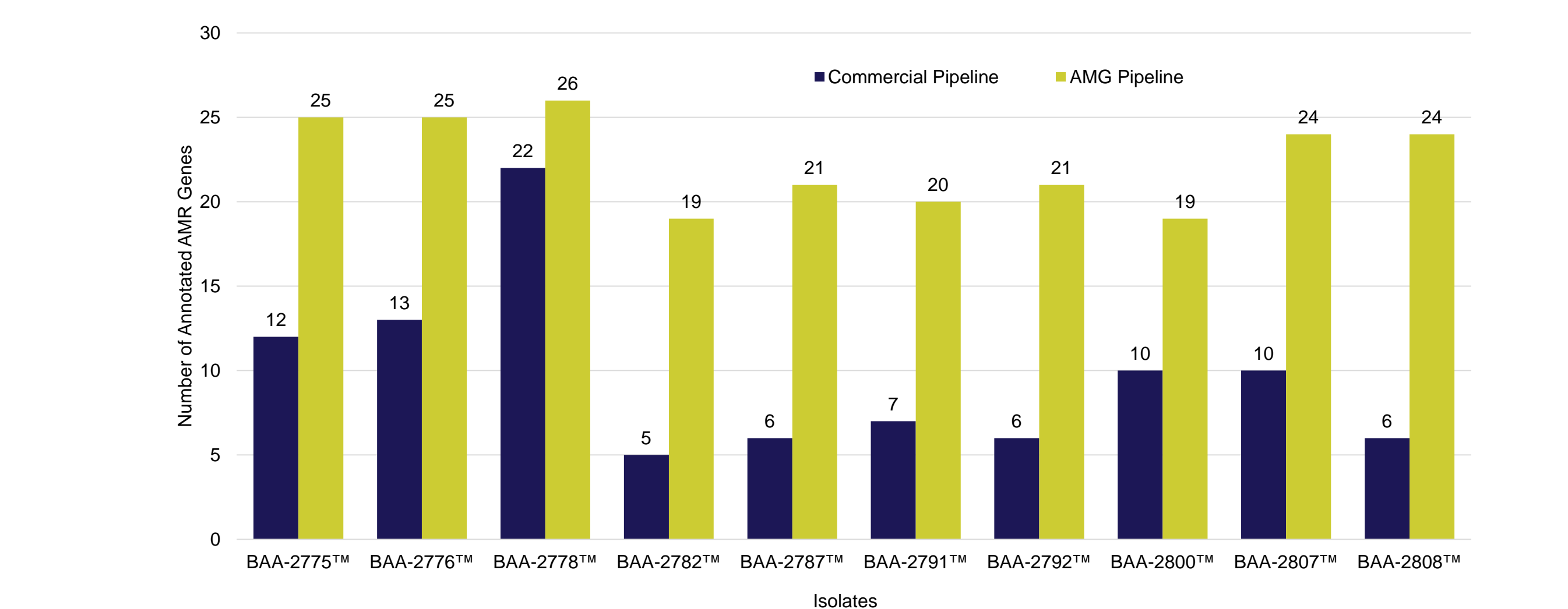


Figure 3. Predicted AMR genes identified with the AMG bioinformatics platform. A comparison of the number of AMR genes (perfect match) between public databases and the AMG platform analysis showed major efficiency in AMR identification using our proprietary bioinformatics pipeline.

Disclaimers

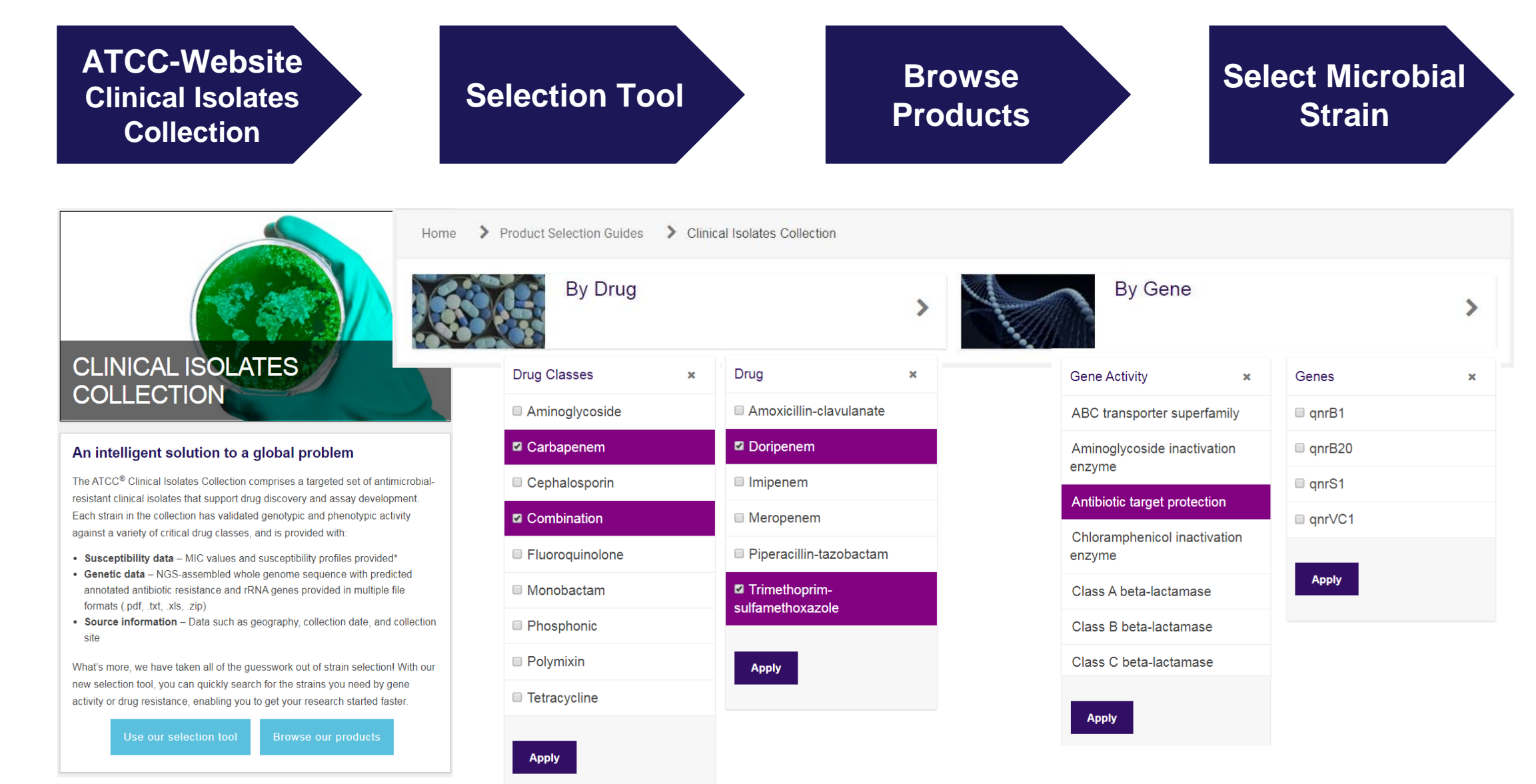
© 2017 American Type Culture Collection. The ATCC trademark and trade name, and any other trademarks listed in this publication are trademarks owned by the American Type Culture Collection unless indicated otherwise.

Predicted AMR Genes in Gram-Negative Species:

Table 2. Composition of predicted AMR genes identified within Gram-negative strains

<i>E. coli</i>	<i>K. pneumoniae</i>	<i>C. freundii, P. mirabilis, E. cloacae, S. marcescens</i>	<i>P. aeruginosa</i>	<i>A. baumannii</i>
AAC(3)-IIa	AAC(3)-IIa	AAC(1)	aadA	ADC-2
AAC(6)-Ib-cr	AAC(3)-IV	AAC(3)-IIa	aadA6	AAC(3)-IIa
aadA	AAC(6)-Ib	AAC(6)-Ib7	amrA	AAC(6)-Ib7
aadA5	AAC(6)-Ib-cr	AAC(6)-Ib-cr	amrB	aadA
<i>E. coli</i> acrA	aadA	aadA7	ANT(2)-Ia	adeA
acrB	aadA2	aadA3	APH(3)-Ib	adeB
acrD	acrA	acrB	APH(3)-Ib	adeC
acrE	<i>K. pneumoniae</i> acrA	<i>E. cloacae</i> acrA	APH(3)-VI	adeF
acrF	acrB	<i>E. coli</i> acrA	APH(3)-VIIa	adeG
APH(3)-Ib	acrD	acrB	APH(6)-Id	adeI
APH(6)-Id	acrE	acrD	catB7	adeJ
baeR	acrF	acrE	catI	adeK
catB3	adeB	acrF	catII	adeR
CMY-2	APH(3)-Ia	ACT-25	flaR	adeS
CMY-47	APH(3)-Ib	APH(3)-Ia	GES-5	APH(3)-Ia
CTX-M-14	APH(6)-Id	APH(3)-Ib	KPC-2	APH(3)-Ib
CTX-M-15	armA	APH(3)-VIIa	LCR-1	APH(3)-VIIa
ermB	catB3	APH(6)-Id	macB	APH(6)-Id
<i>E. coli</i> ermE	catI	armA	armA	armA
flaR	CTX-M-15	cat	mdsB	catB8
KPC-2	CTX-M-2	catII	macA	macA
macA	ermB	crnB	mexB	macB
macB	macB	CMY-16	mexC	mexB
marA	KPC-3	CMY-4	mexD	mexK
mdtB	macB	CMY-48	mexE	mexN
mdtC	mdsB	CTX-M-15	mexF	msrE
mdtF	mdtB	CTX-M-3	mexH	NDM-1
mexB	mdtC	ermB	mexI	OXA-23
msbA	mexB	flaR	mexK	OXA-64
OXA-1	msbA	KPC-2	mexL	OXA-66
qnrB20	qnrB20	msrE	mexP	PER-1
TEM-1	oxqA	mdtB	mexO	TEM-1
tet(A)	oxqB	mdtC	mexV	tet(A)
tet(G)	OXA-1	mdtE	mexW	
tolC	OXA-2	mdtF	msbA	
	OXA-9	mexB	msrE	
	SHV-1	rmxB		
	SHV-11	mexC		
	SHV-66	msbA		
	TEM-1	NDM-1		
	qnrB1	oprM		
	qnrS1	OXA-10		
	ramA	OXA-129		
	robA	PDC-2		
	SRT-2	PDC-3		
	TEM-1	PDC-7		
	TEM-2	PER-1		
	tet(41)	qnrVC1		
	tet(G)	tet(G)		
	tolC	VEB-9		
		VIM-2		
		VIM-5		

Running the Selection Tool on the ATCC Website



Summary

- We developed an accurate and efficient tool for microbial genome sequencing and annotation
- The genetic sequence and annotation of strains provide researchers crucial genomic information needed to identify potential modes of antibiotic resistance
- This set of fully characterized reference strains are ideal for facilitating antibacterial drug discovery and diagnostic assay development

References

- World Health Organization. Global Priority List of Antibiotic-Resistant Bacteria to Guide Research, Discovery, and Development of New Antibiotics. Accessed online August 2017 <http://www.who.int/medicines/publications/WHO-PPL-Short_Summary_25Feb-ET_NM_WHO.pdf>.
- Centers for Disease Control and Prevention. Antibiotic/Antimicrobial Resistance. Accessed online August 2017 <https://www.cdc.gov/drugresistance/index.html>.