Supplementary Figures

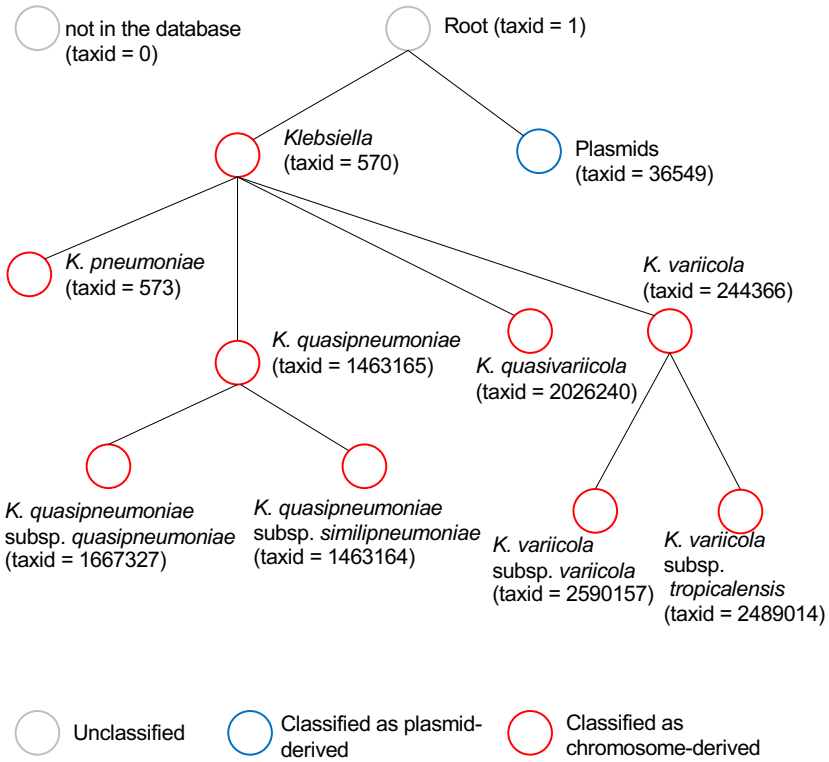


Figure S1. The taxonomic tree used in the Kraken analysis. This is a taxonomy tree from the NCBI taxonomy downloaded by "kraken-build --download-taxonomy --db database". Although there are other taxa in the tree, they are not shown here because no chromosomal sequences other than *Klebsiella* spp. were included in the database. One of the taxonomic IDs in the tree was assigned to each contig. Contigs with a taxonomic ID 0 or 1 were labeled as unclassified, those with an ID 36549 were classified as plasmid-derived, and those with the remaining IDs were classified as chromosome-derived. Note that the taxon *Klebsiella variicola* subsp. *tropicalensis* has been renamed to *K. variicola* subsp. *tropica*, but has not yet been updated in NCBI (https://www.sciencedirect.com/science/article/pii/S0923250819300956).

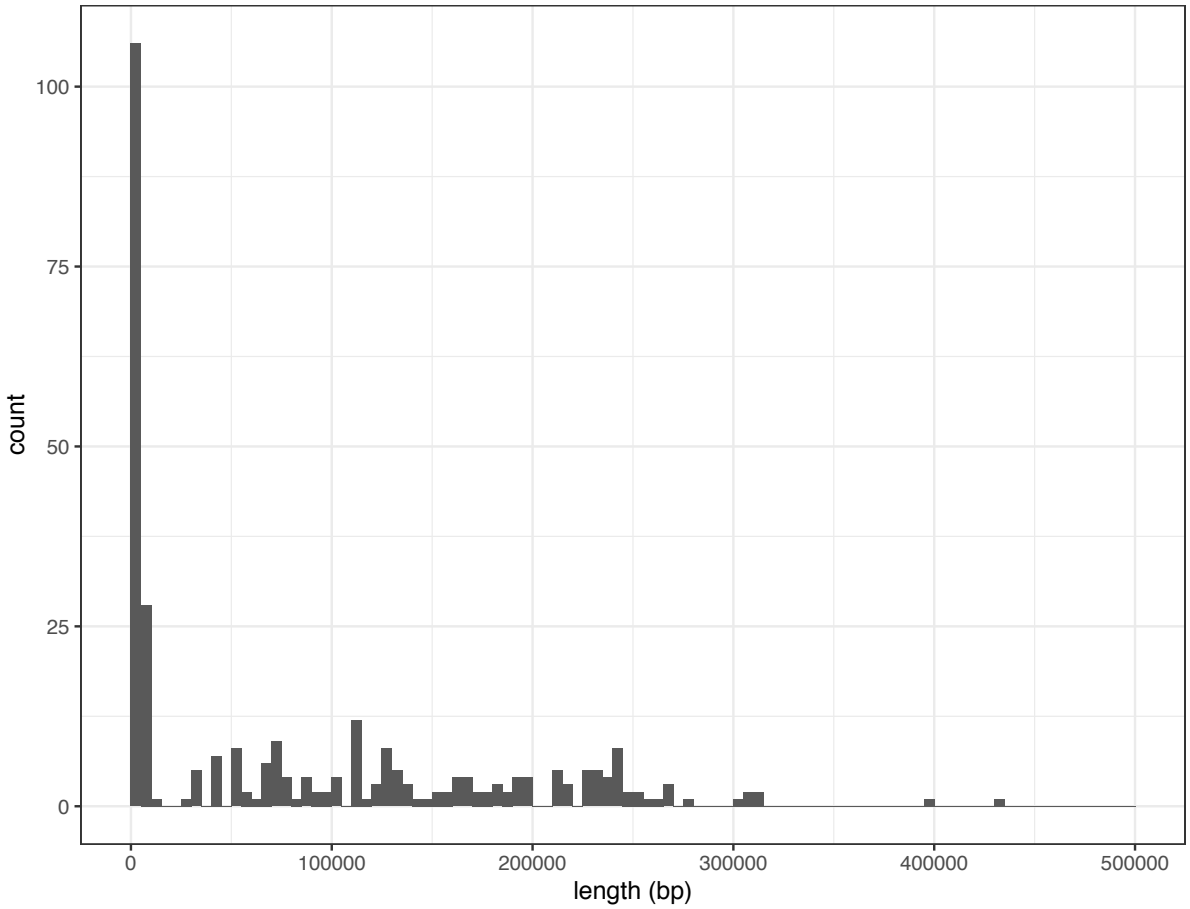
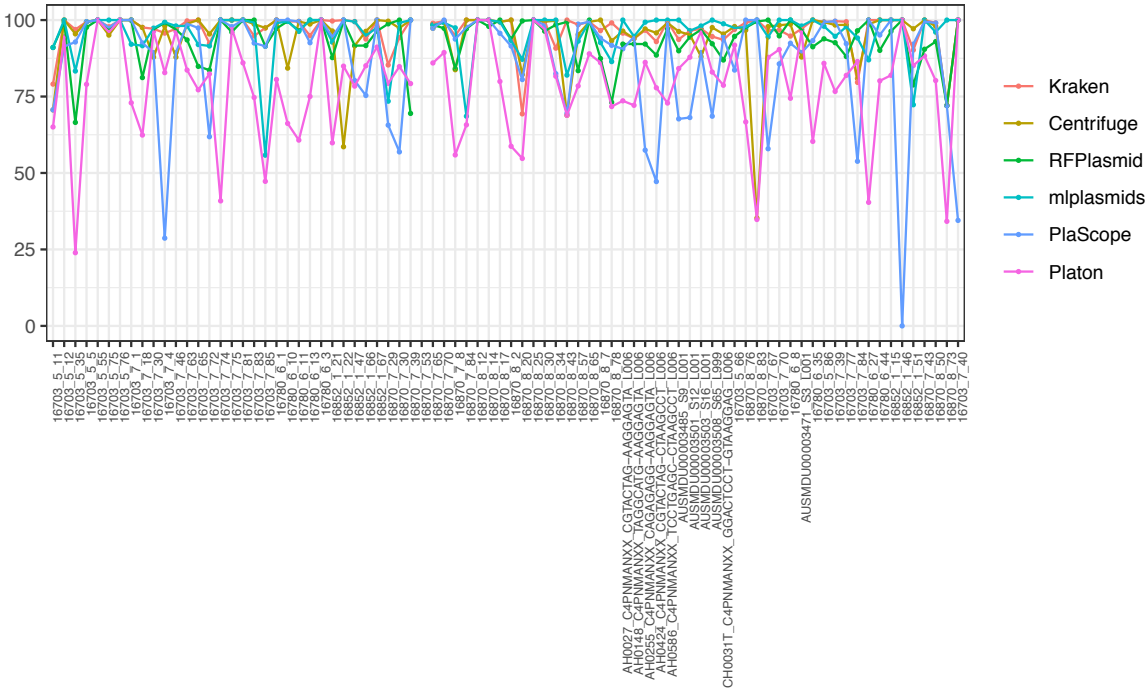
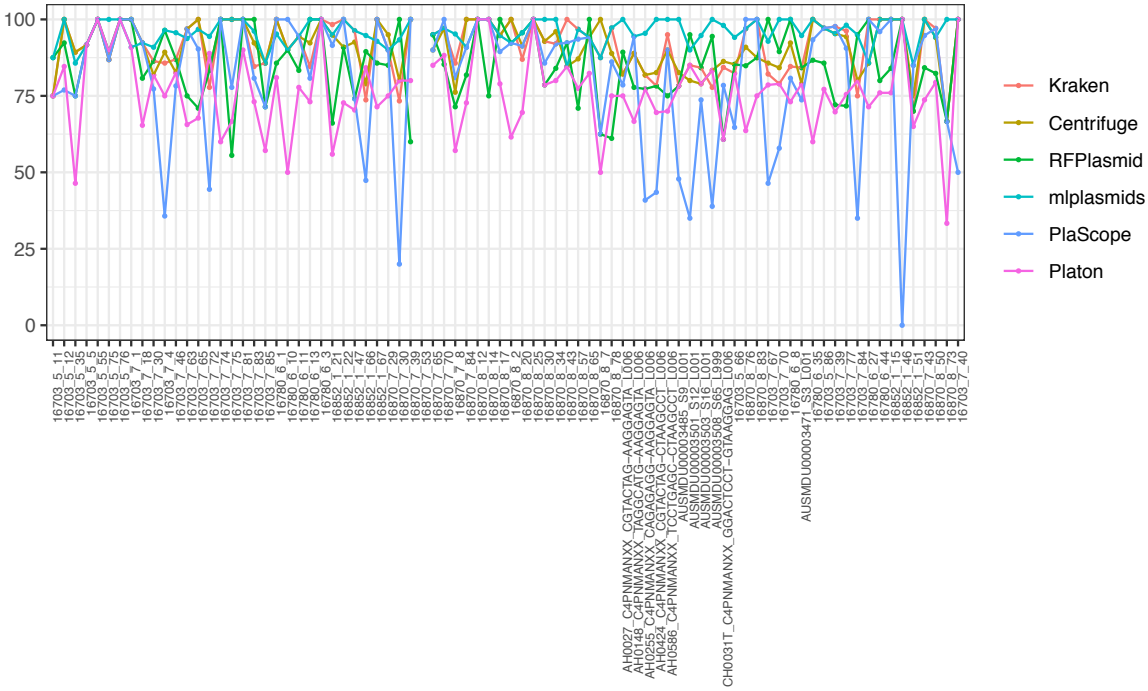


Figure S2. Length distribution of the completed plasmids (n=301).

Sensitivity (contig length)

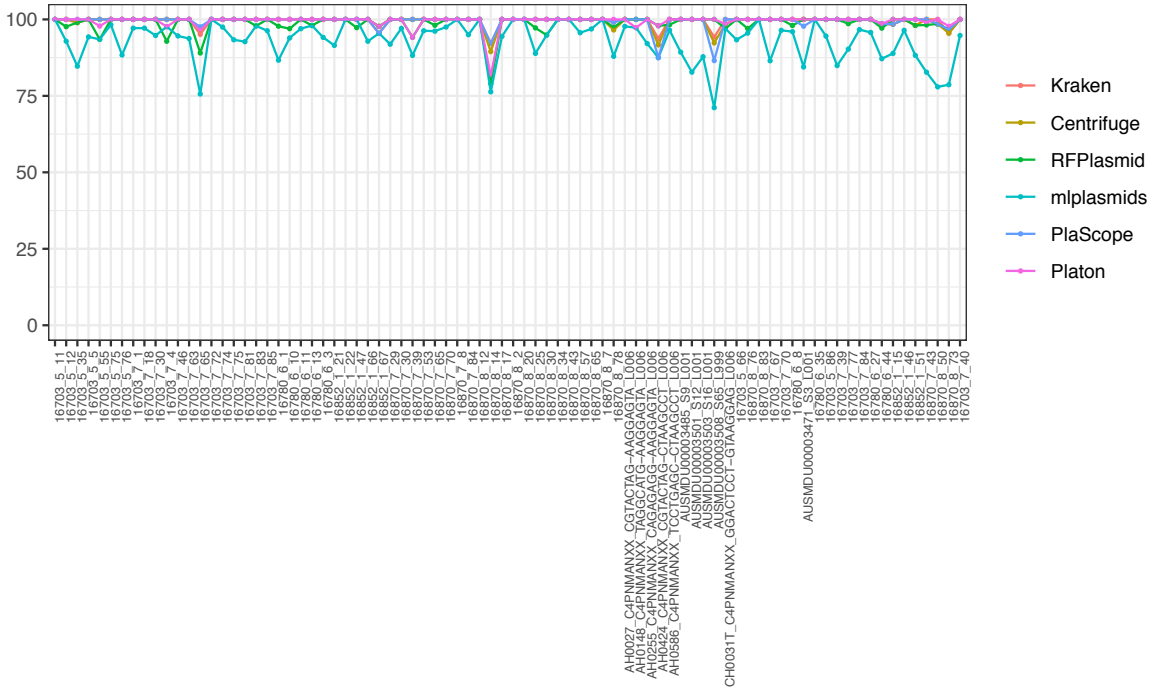


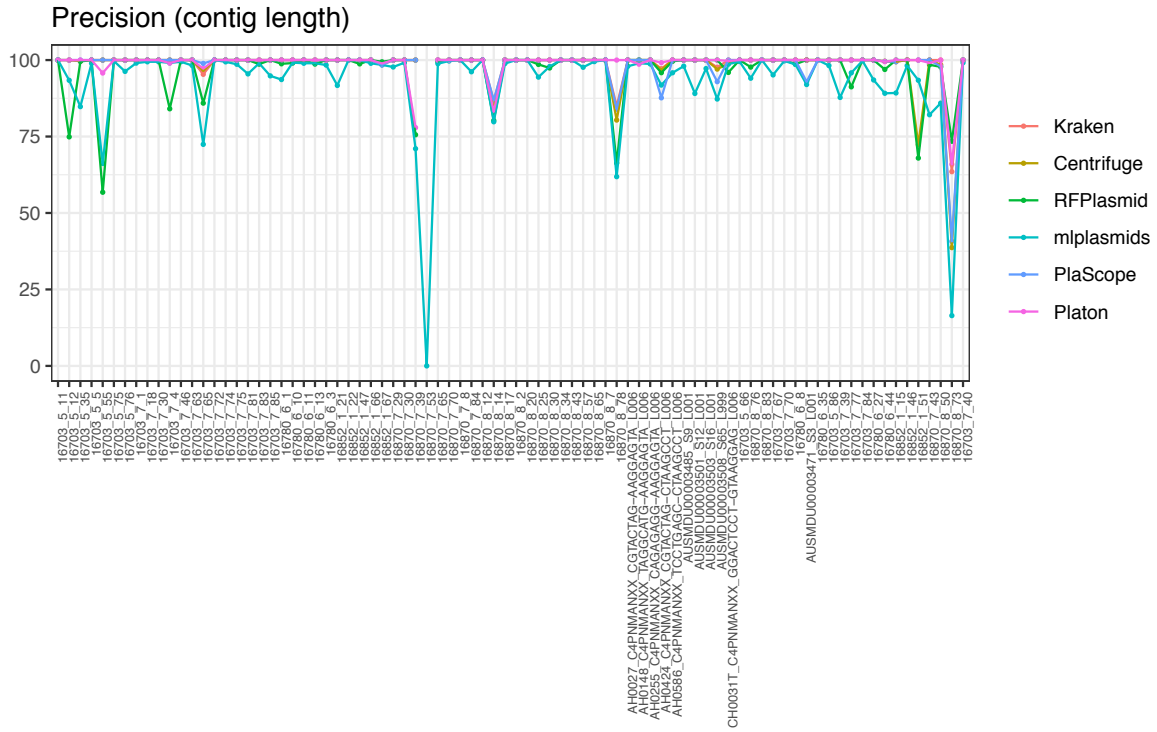
Sensitivity (contig number)

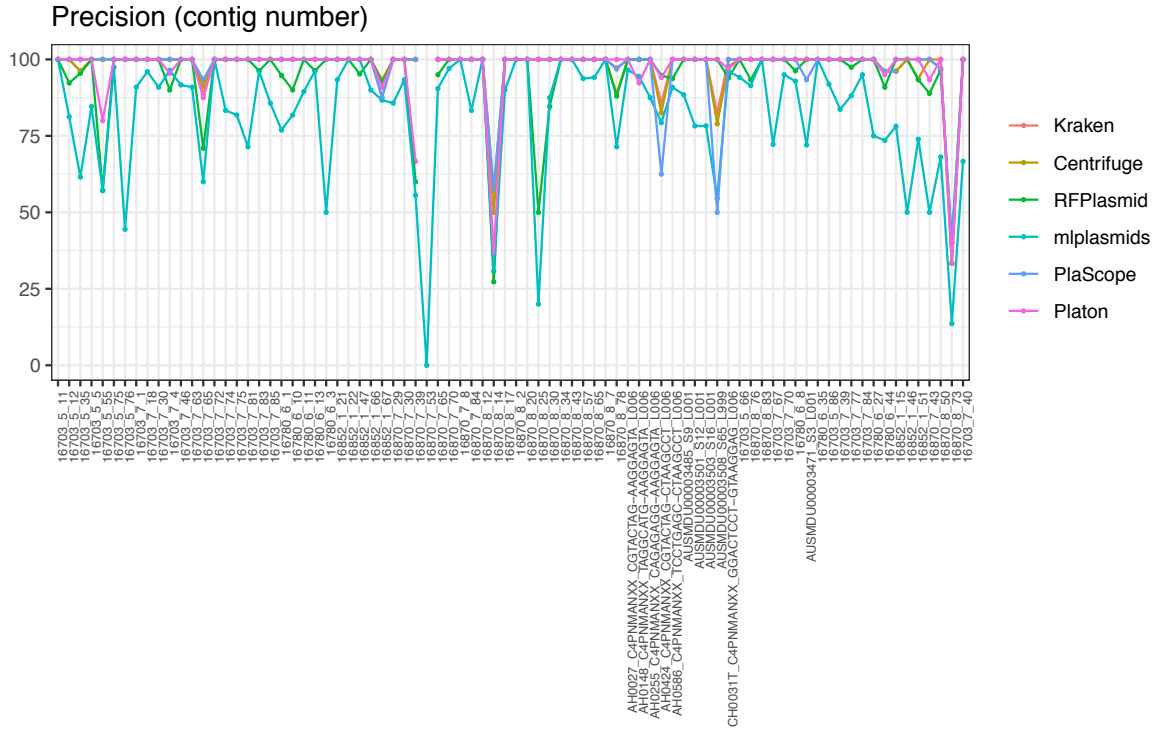


Specificity (contig length)

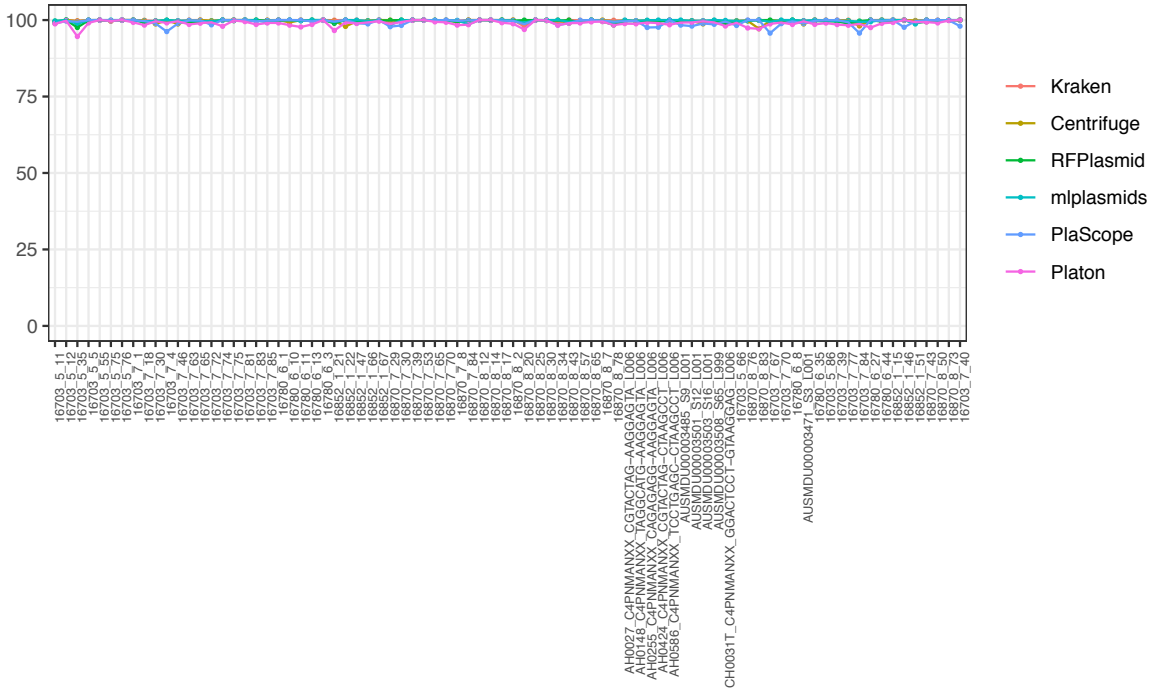
Specificity (contig number)



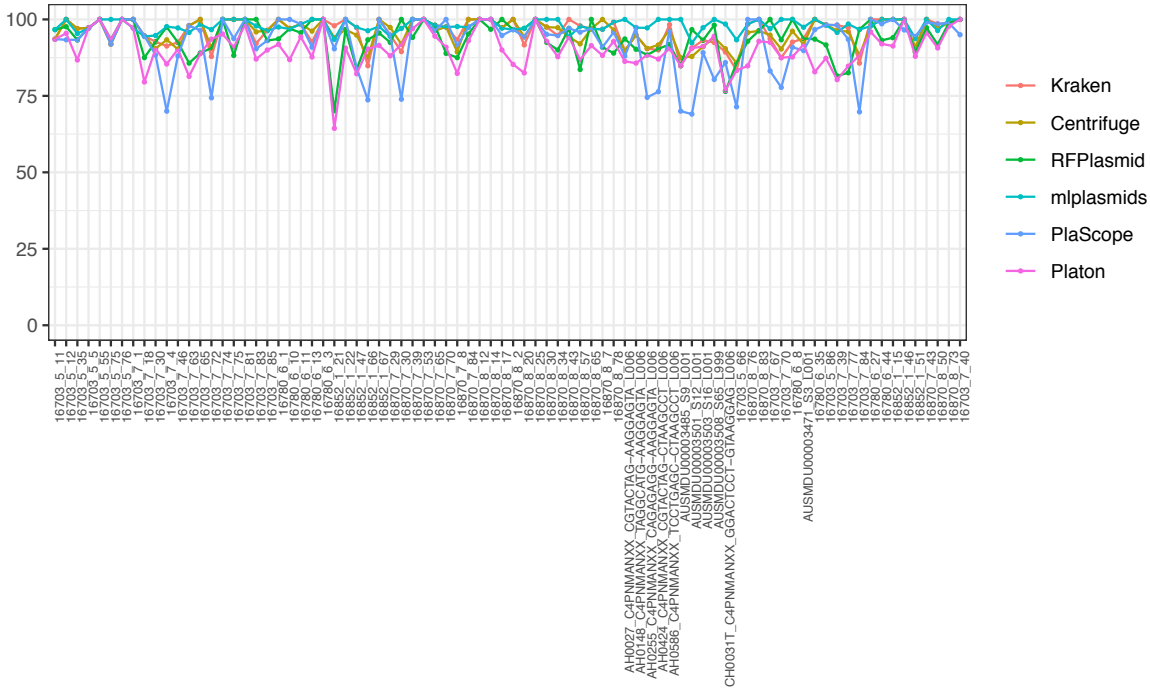

Precision (contig length)
100
75
50
25
0
Kraken
Centrifuge
RFPlasmid
mlplasmids
PlaScope
Platon


Precision (contig number)
100
75
50
25
0
Kraken
Centrifuge
RFPlasmid
mlplasmids
PlaScope
Platon

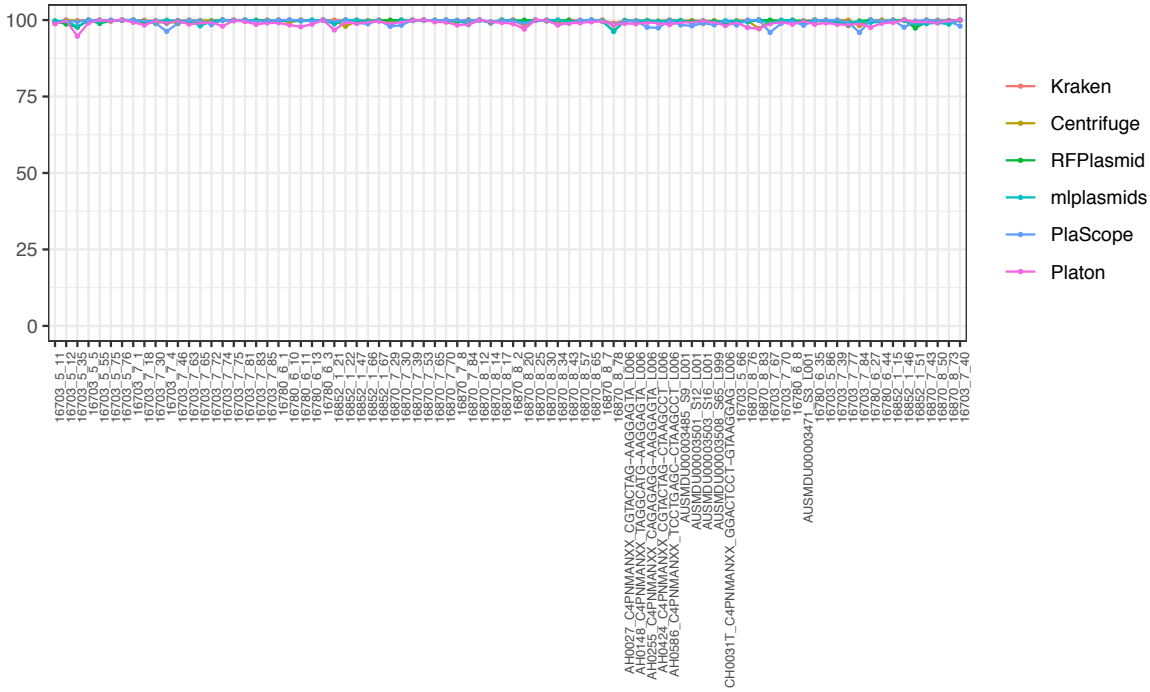
Negative predictive value (contig length)



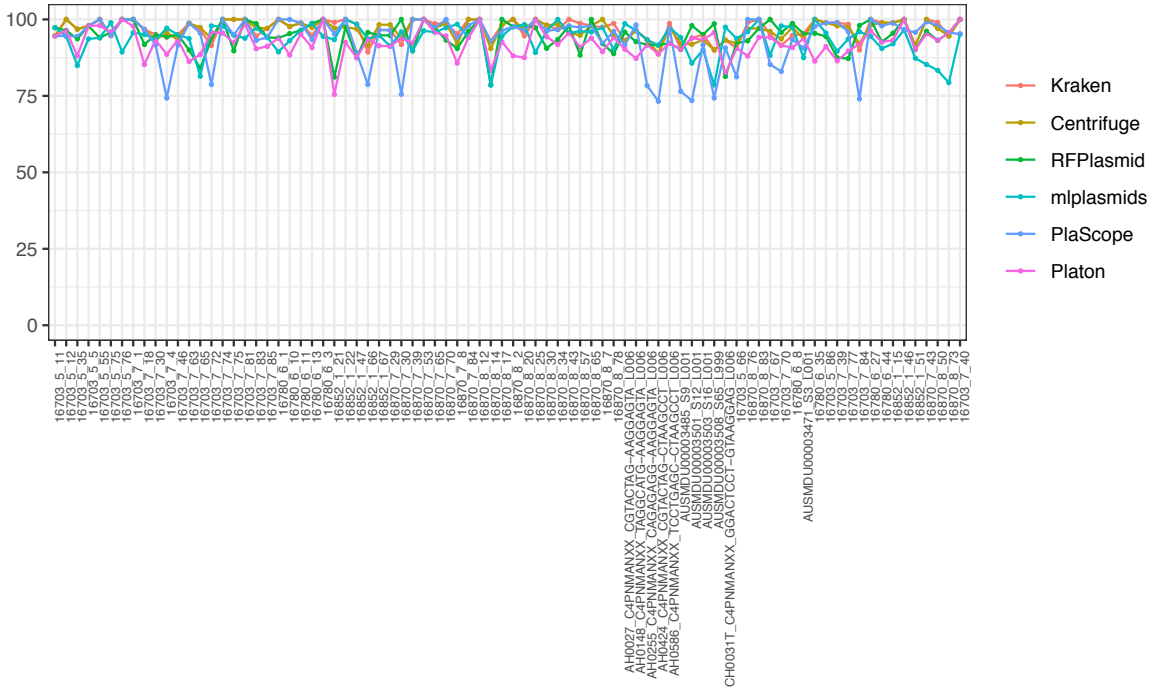
Negative predictive value (contig number)



Accuracy (contig length)



Accuracy (contig number)



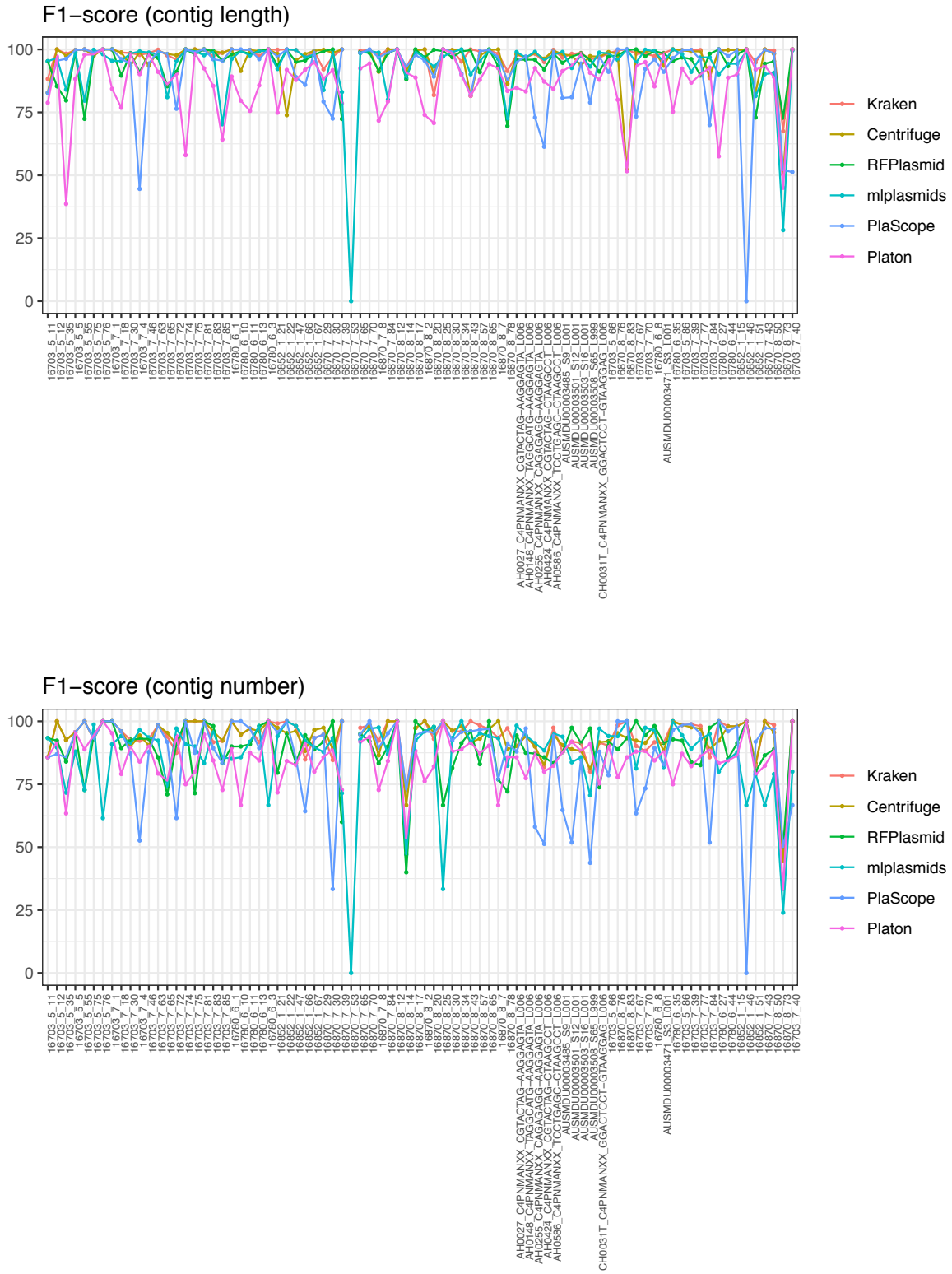
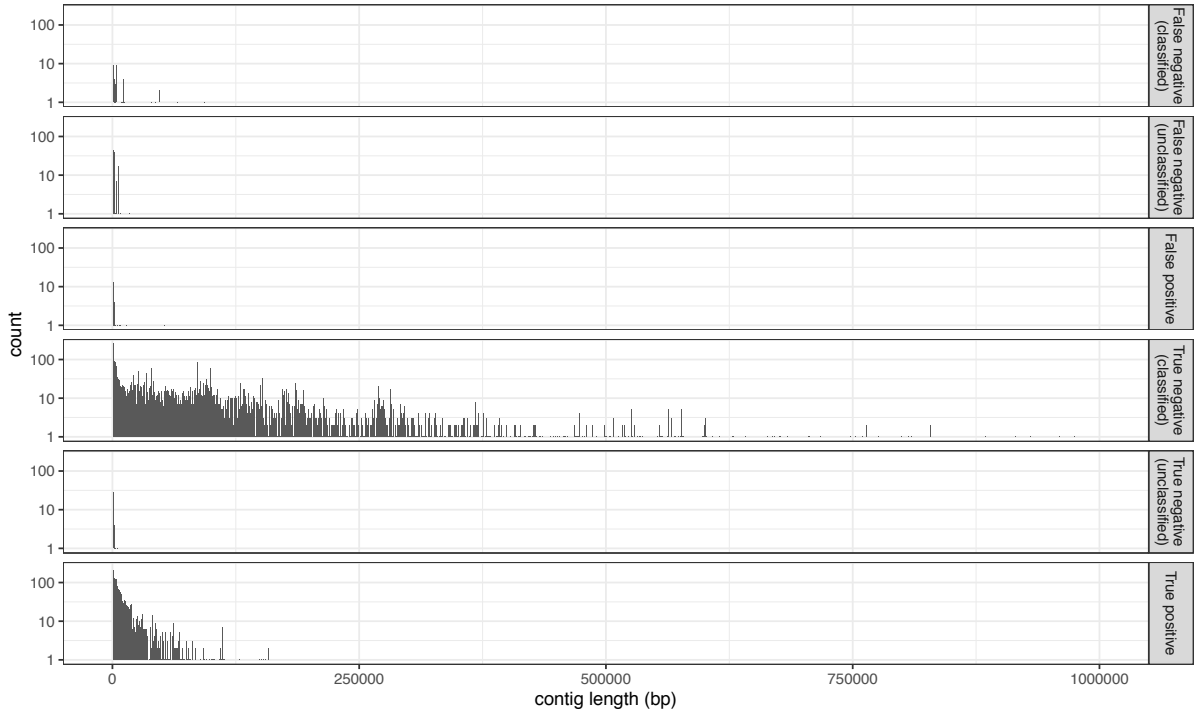
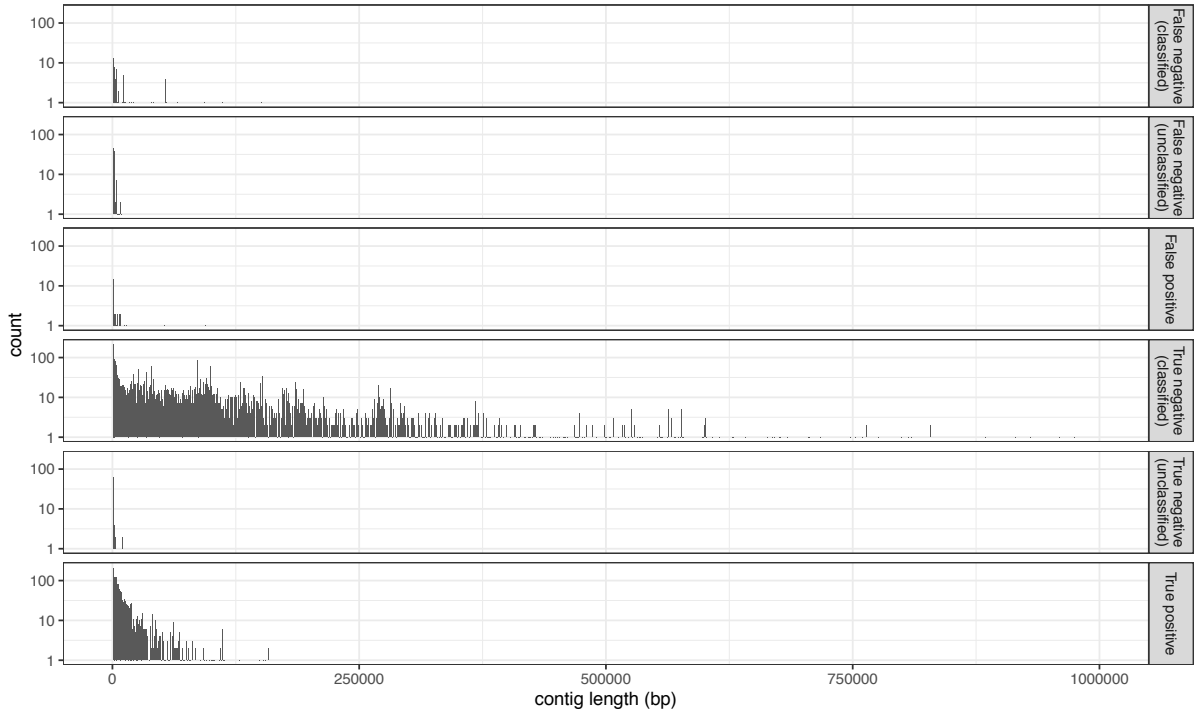


Figure S3. Performance metrics for Kraken, Centrifuge, RFPlasmid, mlplasmids, PlaScope, and Platon for each genome**.** Sensitivity, precision, and F1-score are not shown for 16870_7_53 except mlplasmids because this genome does not include plasmids and thus the true positive category is not applicable. (mlplasmids predicted one contig as plasmid-derived for this genome.)

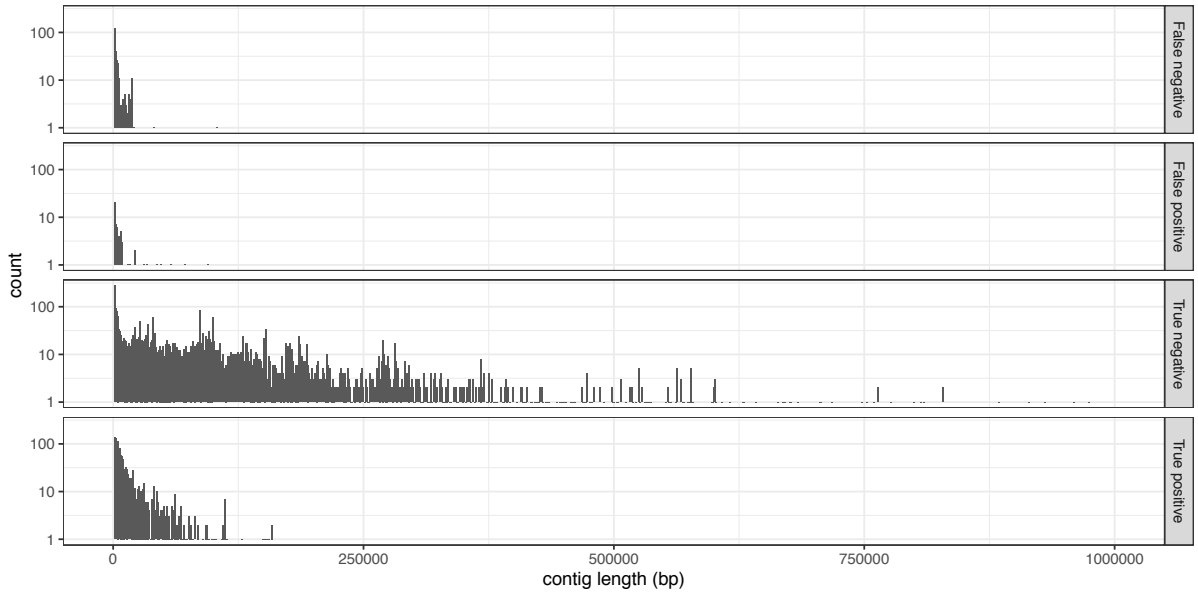
Kraken



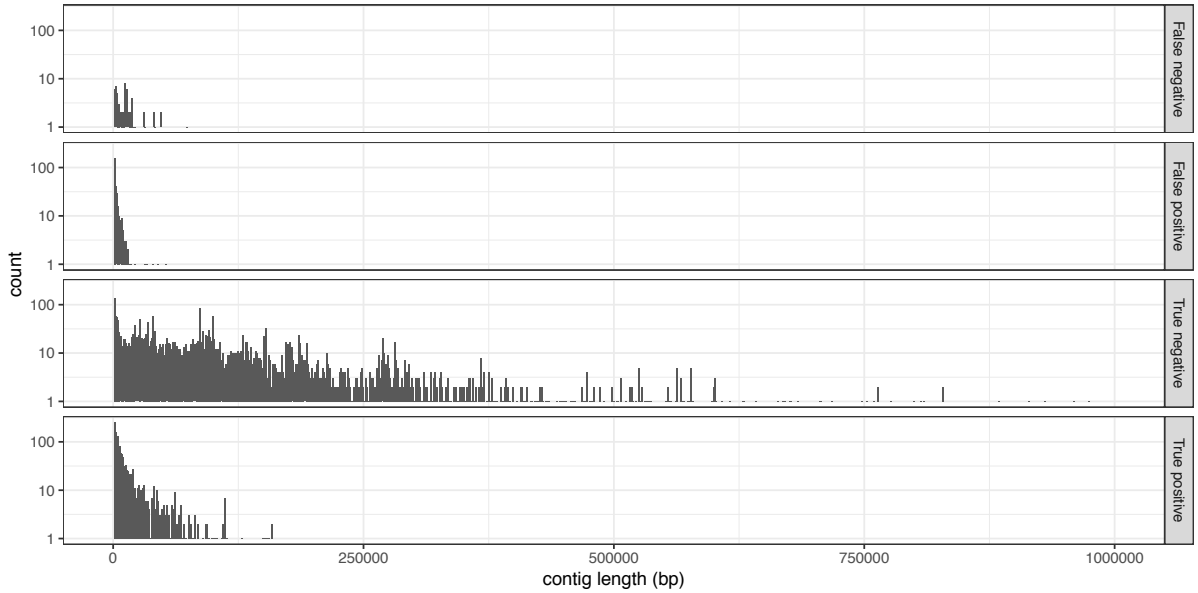
Centrifuge



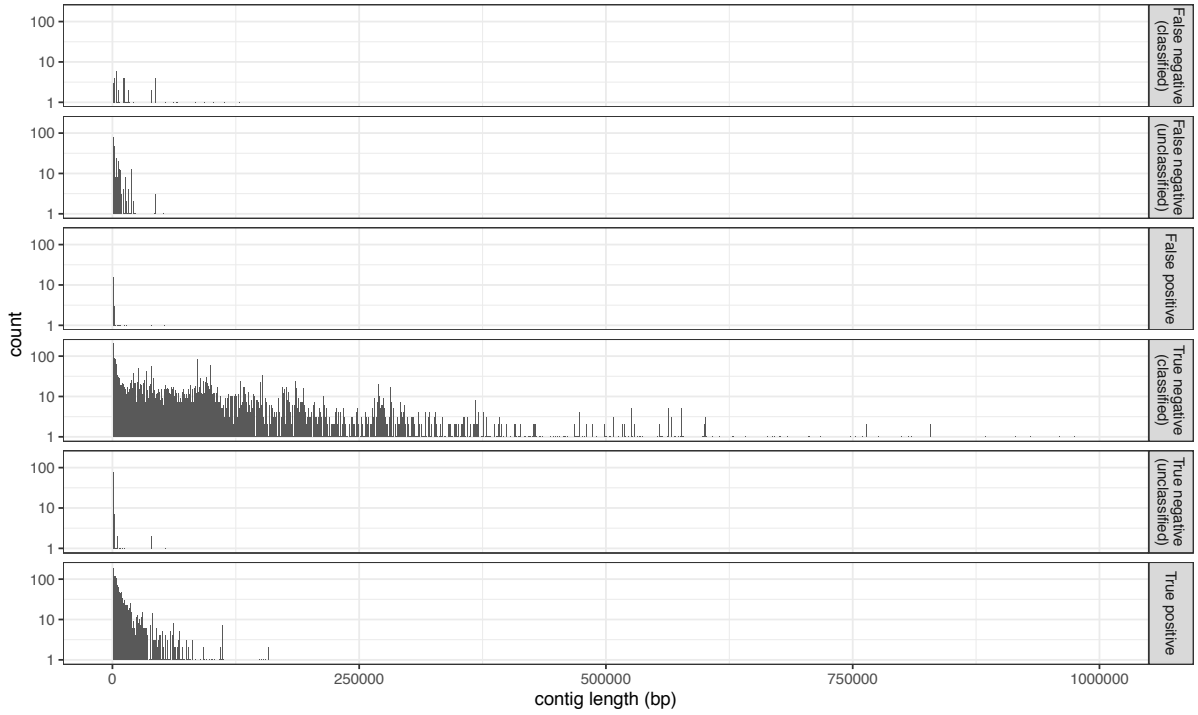
RFPlasmid



mlplasmids



PlaScope



Platon

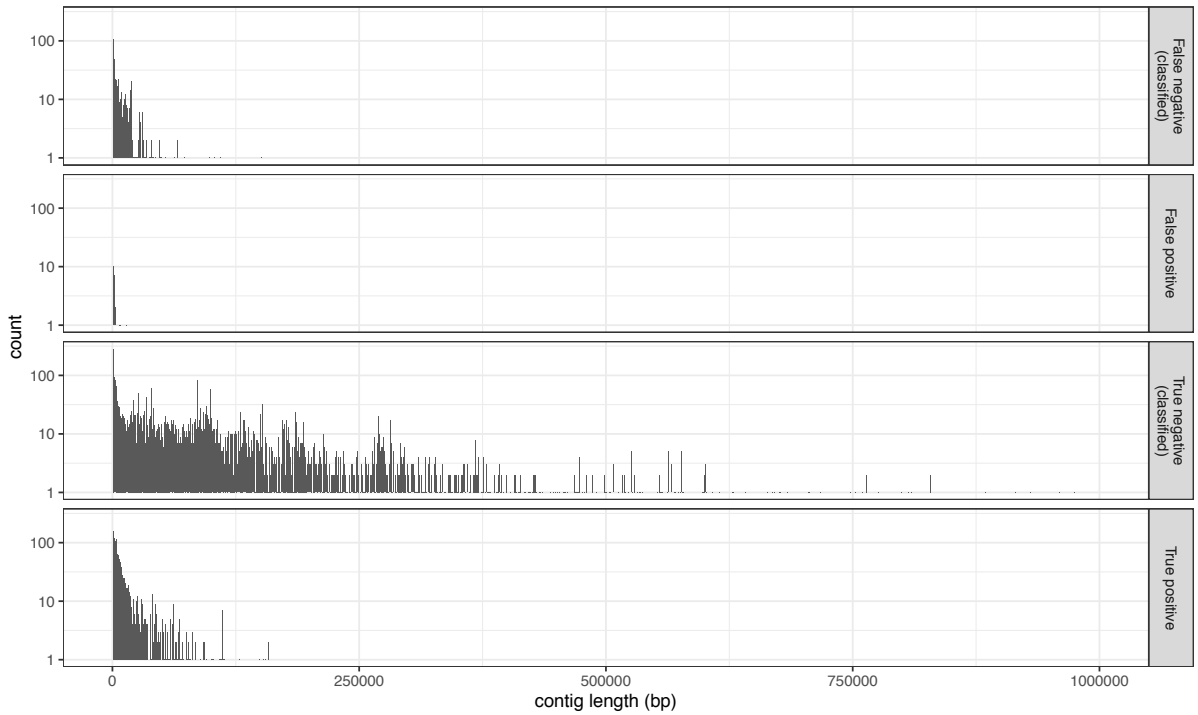


Figure S4. Relationships between contig length and classification categories. The y axes are on a log10 scale.

Supplementary Tables

Table S1. See the excel file.

Table S2. See the excel file.

Table S3. Total contig sizes and contig counts assigned to each category.

	Kraken	Centrifuge	RFPlasmid	mlplasmids	PlaScope	Platon
True positive						
Contig length (bp)	22,223,924	21,776,743	21,692,165	22,104,317	20,080,424	18,472,253
Contig count	1,460	1,455	1,333	1,541	1,305	1,193
True negative (classified)						
Contig length (bp)	428,515,693	428,319,115	428,083,646	427,516,067	428,203,081	428,620,951
Contig count	3,640	3,595	3,640	3,394	3,573	3,675
True negative (unclassified)						
Contig length (bp)	54,468	115,685	-	-	306,627	-
Contig count	34	72	-	-	96	-
False positive						
Contig length (bp)	117,854	253,215	604,369	1,171,948	178,307	67,064
Contig count	23	30	57	303	28	22
False negative (classified)						
Contig length (bp)	486,738	1,025,913	1,336,170	924,018	1,297,404	4,556,082
Contig count	38	57	275	67	47	415
False negative (unclassified)						
Contig length (bp)	317,673	225,679	-	-	1,650,507	-
Contig count	110	96	-	-	256	-

Kraken, Centrifuge, and PlaScope can label contigs as unclassified. As described in **Methods** section in the main text, these unclassified contigs were assigned to true negative if they were mapped to true chromosome or assigned to false negative if they were mapped to true plasmids.

Table S4. Consistency of performance between genomes in terms of interquartile range (IQR).

	Kraken	Centrifuge	RFPlasmid	mlplasmids	PlaScope	Platon
IQR of sensitivity (%)						
Contig	**4.27**	4.55	8.15	5.36	15.88	16.10
length	**(99.32)**	(98.38)	(96.51)	(99.05)	(94.08)	(81.91)
Contig	15.38	14.71	17.22	**5.88**	23.07	12.84
count	(94.74)	(92.31)	(85.71)	**(96.30)**	(90.00)	(77.14)
IQR of specificity (%)						
Contig	**0**	**0**	0.06	0.27	**0**	**0**
length	**(100)**	**(100)**	**(100)**	(99.93)	**(100)**	**(100)**
Contig	**0**	**0**	1.99	8.27	**0**	**0**
count	**(100)**	**(100)**	**(100)**	(94.80)	**(100)**	**(100)**
IQR of precision (%)						
Contig	**0**	**0**	1.32	5.91	**0**	**0**
length	**(100)**	**(100)**	**(100)**	(98.39)	**(100)**	**(100)**
Contig	**0**	**0**	5.00	21.57	**0**	**0**
count	**(100)**	**(100)**	**(100)**	(88.97)	**(100)**	**(100)**
IQR of negative predictive value (%)						
Contig	0.27	**0.26**	0.42	0.27	1.10	0.93
length	**(99.96)**	(99.91)	(99.82)	(99.95)	(99.76)	(99.07)
Contig	7.38	6.61	7.41	**3.10**	8.65	7.25
count	(97.53)	(96.64)	(93.89)	**(97.85)**	(95.48)	(90.65)
IQR of accuracy (%)						
Contig	**0.28**	0.29	0.42	0.55	1.01	0.87
length	**(99.95)**	(99.90)	(99.70)	(99.75)	(99.76)	(99.10)
Contig	5.26	**4.36**	5.74	5.61	6.56	5.17
count	**(98.04)**	(97.36)	(95.33)	(94.85)	(96.40)	(92.68)
IQR of F1-score (%)						
Contig	2.60	**2.53**	5.03	5.63	9.33	9.99
length	**(99.48)**	(99.01)	(97.60)	(97.37)	(96.79)	(90.06)
Contig	8.33	**7.08**	12.00	11.75	12.45	10.32
count	**(96.30)**	(95.52)	(91.43)	(91.78)	(94.44)	(86.02)

Median values are shown in parentheses. Bold font indicates the lowest IQR and highest median value(s) in each category.

Table S5. Wilcoxon signed-rank test results (p-values) for Kraken vs other methods.

	Centrifuge	RFPlasmid	mlplasmids	PlaScope	Platon
Sensitivity					
Contig length	1.000	**<0.001**	1.000	**<0.001**	**<0.001**
Contig count	1.000	**0.00350**	**<0.001**	**<0.001**	**<0.001**
Specificity					
Contig length	0.137	**<0.001**	**<0.001**	1.000	1.000
Contig count	0.488	**0.00327**	**<0.001**	1.000	1.000
Precision					
Contig length	0.122	**<0.001**	**<0.001**	0.527	1.000
Contig count	0.488	**0.00430**	**<0.001**	1.000	1.000
Negative predictive value					
Contig length	1.000	**<0.001**	1.000	**<0.001**	**<0.001**
Contig count	1.000	**0.0262**	**0.00468**	**<0.001**	**<0.001**
Accuracy					
Contig length	0.458	**<0.001**	**<0.001**	**<0.001**	**<0.001**
Contig count	1.000	**0.00191**	**<0.001**	**<0.001**	**<0.001**
F1-score					
Contig length	0.495	**<0.001**	**<0.001**	**<0.001**	**<0.001**
Contig count	1.000	**<0.001**	**<0.001**	**<0.001**	**<0.001**

Wilcoxon signed-rank test was performed for each metric using values shown in **Figure 1**and **Figure S3**. Bonferroni correction was applied across each row. p-values < 0.05 are shown in bold.

Table S6. Performance of Centrifuge using the database created without excluding chromosomal sequences containing integrated plasmid sequences.

	Contig length	Contig count
Sensitivity (true positive rate, %)	80.030 (94.565)	78.109 (90.485)
Specificity (true negative rate, %)	99.943 (99.941)	99.216 (99.189)
Precision (positive predictive value, %)	98.685 (98.851)	97.743 (97.980)
Negative predictive value (%)	98.938 (99.709)	91.244 (95.995)
Accuracy (%)	98.928 (99.667)	92.818 (96.550)
F1-score ‘%)	88.384 (96.660)	86.830 (94.083)

Values in **Table 2** i.e. those calculated when excluding chromosomes with integrated plasmids are shown in parentheses for comparison purposes.

Table S7. Performance of mlplasmids for *K. pnuemoniae* genomes (n=69).

	Contig length	Contig count
Sensitivity (true positive rate, %)	96.158	95.761
Specificity (true negative rate, %)	99.771	93.032
Precision (positive predictive value, %)	95.770	86.054
Negative predictive value (%)	99.792	97.995
Accuracy (%)	99.585	93.878
F1-score (%)	95.963	90.649

Table S8. mlplasmids performance with a requirement of a minimum posterior probability of 0.7 for classification.

	Contig length	Contig count
Sensitivity (true positive rate, %)	90.241	78.794
Specificity (true negative rate, %)	99.882	98.323
Precision (positive predictive value, %)	97.618	95.335
Negative predictive value (%)	99.478	91.424
Accuracy (%)	99.390	92.403
F1-score (%)	93.785	86.279

Table S9. Performance metrics of Kraken in the cross validation analysis.

	Group A	Group B	Group C	Overall	Base database*
Sensitivity (true positive rate, %)					
Contig length	96.699	97.107	98.861	97.547	96.507
Contig count	89.984	89.883	92.784	90.796	90.796
Specificity (true negative rate, %)					
Contig length	99.952	99.996	99.975	99.975	99.973
Contig count	99.584	99.718	99.497	99.594	99.378
Precision (positive predictive value, %)					
Contig length	99.103	99.934	99.505	99.517	99.472
Contig count	98.917	99.355	98.684	98.983	98.449
Negative predictive value (%)					
Contig length	99.820	99.839	99.942	99.868	99.813
Contig count	95.922	95.320	97.138	96.136	96.128
Accuracy (%)					
Contig length	99.784	99.844	99.921	99.851	99.796
Contig count	96.732	96.510	97.558	96.927	96.777
F1-score (%)					
Contig length	97.886	98.500	99.182	98.522	97.967
Contig count	94.239	94.382	95.643	94.713	94.468

*Values in this column are from **Table 2** and shown for comparison purposes.

Table S10. See the excel file.

Supplementary Methods

Commands used for running Kraken.

1. Download the NCBI taxonomy.

```
kraken-build --download-taxonomy --db database
```

2. Add genomes to the library.

```
for file in reference/*.fasta; do
    kraken-build --add-to-library ${file} --db database
done
```

#Note that taxonomy information was added to the headers of fasta files. For example, the fasta file of *K. pneumoniae* chromosome NZ_CP026130.1 begins with ">NZ_CP026130.1|kraken:taxid|573", the fasta file of *K. variicola* subsp. *variicola* chromosome NZ_CP020847.1 begins with ">NZ_CP020847.1|kraken:taxid|2590157", and the fasta file of plasmid AP014611 begins with ">AP014611.1|kraken:taxid|36549".

3. Build the database.

```
kraken-build --build --db database
```

4. Classify sequences.

```
for file in assemblies/*.fasta; do
    kraken --preload --db database ${file} > $(basename ${file/.fasta})_output.txt
done
```

Commands used for running Centrifuge.

1. Build the database.

```
centrifuge-build -p 10 --conversion-table seqid_to_taxid.map --taxonomy-tree nodes.dmp --name-table names.dmp database.fna chromosome_plasmid_db
```

#Note that the corresponding line in the nodes.dmp file was modified as follows:

from

```
36549 | 28384 | no rank | | 0 | 0 | 11 | 1 | 0 | 1 | 0 | 0 | |
```

to

```
36549 | 28384 | species | | 0 | 0 | 11 | 1 | 0 | 1 | 0 | 0 | |
```

2. Classify sequences.

```
for file in assemblies/*.fasta; do
    centrifuge -f -p 10 --reorder -x chromosome_plasmid_db -U ${file} -k 1 --report-file $(basename ${file/.fasta})_summary.txt -S $(basename ${file/.fasta})_output.txt
```

done

Commands used for running RFPlasmid.

python3 rfplasmid.py --species Enterobacteriaceae --input assemblies/ --jelly --threads 8 --out output_directory

R commands used for running mlplasmids.

plasmid_classification(path_input_file = fasta_file, full_output = TRUE, species = "Klebsiella pneumoniae")

Commands used for running PlaScope.

plaScope.sh --fasta my_fastafile.fasta -o output_directory --db_dir path/to/DB --db_name Klebsiella_PlaScope --sample name_of_my_sample

Commands used for running Platon.

platon --db path/to/DB --threads 10 --output results/ my_fastafile.fasta