BacAnt: A Combination Annotation Server for Bacterial DNA Sequences to Identify Antibiotic Resistance Genes, Integrons, and Transposable Elements.

Xiaoting Hua[1,2,3†], Qian Liang[2,4†‡], Min Deng[5†], Jintao He[1,2,3], Meixia Wang[2,4], Wenjie Hong[2,4], Jun Wu[6], Bian Lu[7], Sebastian Leptihn[1,8], Yunsong Yu[1,2,3*], Huan Chen[1,2,4*‡]

[1]Department of Infectious Diseases, Sir Run Run Shaw Hospital, College of Medicine, Zhejiang University, Hangzhou, China

[2]Key laboratory of microbial technology and bioinformatics of Zhejiang Province, Zhejiang Institute of Microbiology, Hangzhou 310012, China

[3]Regional Medical Center for National Institute of Respiratory Diseases, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Hangzhou, China

[4]NMPA Key laboratory for Testing and Risk Warning of Pharmaceutical Microbiology, Zhejiang Institute of Microbiology, Hangzhou 310012, China

[5]Department of Infectious Diseases, The First Hospital of Jiaxing, The First Affiliated Hospital of Jiaxing University, Jiaxing 314000, China

[6]Lin'an Center for Disease Control and Prevention, Lin'an, 311300, China

[7]Xiaoshan Center for Disease Control and Prevention, Hangzhou, 311201, China

[8]Zhejiang University-University of Edinburgh Institute, School of Medicine, Zhejiang University, Hangzhou, China

[†]These authors contributed equally to this work and should be considered as co-first authors.

[‡] Present address: Hangzhou Digital Micro Biotech Co., Ltd., Hangzhou, 311215, China

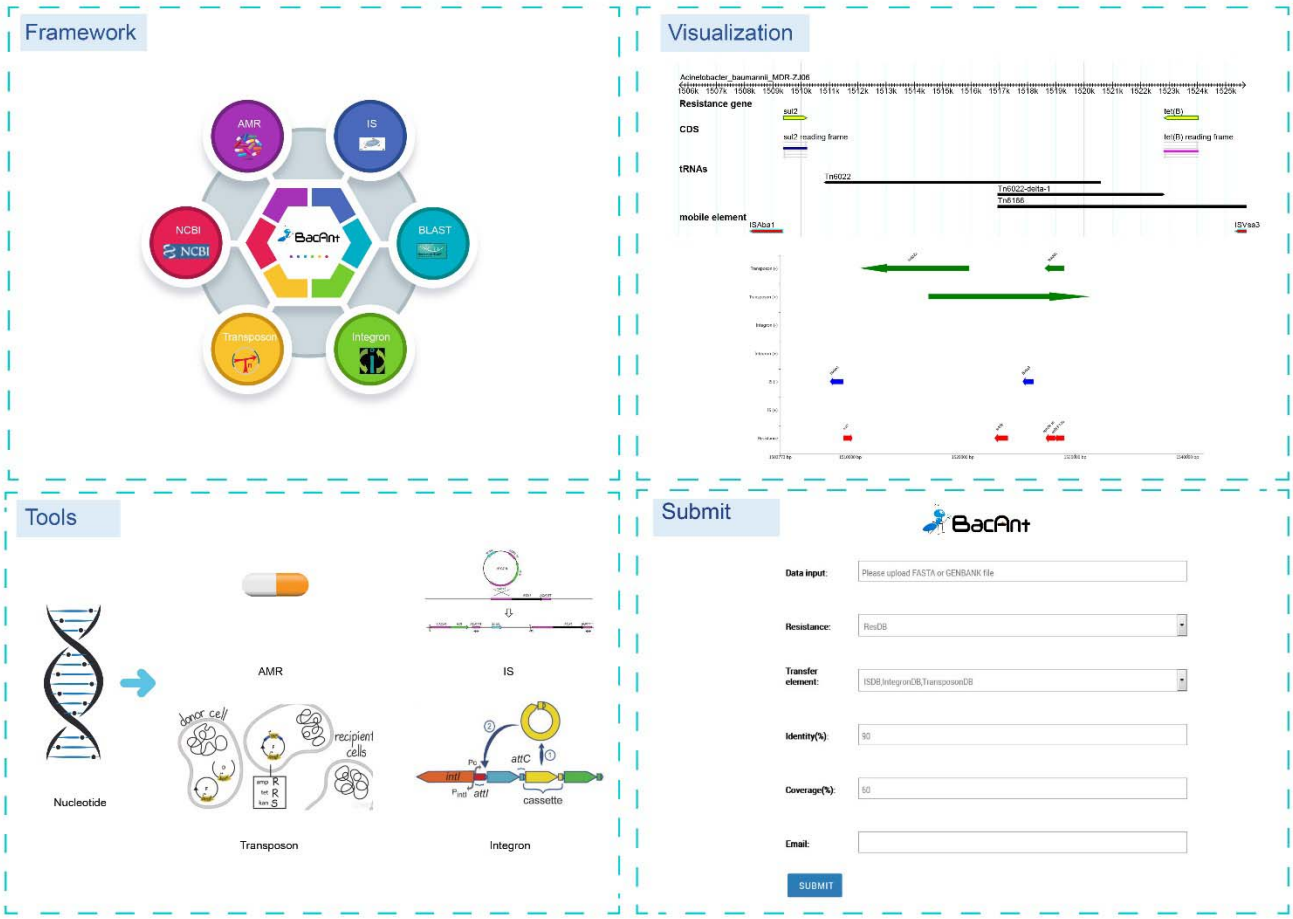
***Correspondence:**
Yunsong Yu, yvys119@zju.edu.cn, or Huan Chen, 10607068@zju.edu.cn

Key words: Antimicrobial Resistance, Resistance genes, Gene Annotation, Insertion Sequences, IS, integron, Transposons, Tn

Abstract

Whole genome sequencing (WGS) of bacteria has become a routine method in diagnostic laboratories. One of the clinically most useful advantages of WGS is the ability to predict antimicrobial resistance genes (ARGs) and mobile genetic elements (MGEs) in bacterial sequences. This allows comprehensive investigations of such genetic features but can also be used for epidemiological studies. A plethora of software programs have been developed for the detailed annotation of bacterial DNA sequences, such as RAST (Rapid Annotation using Subsystem Technology), Resfinder, ISfinder, INTEGRALL and The Transposon Registry. Unfortunately, to this day, a reliable annotation tool of the combination of ARGs and MGEs is not available, and the generation of genbank files requires much manual input. Here, we present a new webserver which allows the annotation of ARGs, integrons, and transposable elements at the same time. The pipeline generates genbank files automatically, which are compatible with easyfig for comparative genomic analysis. Our BacAnt code and standalone software package are available at https://github.com/xthua/bacant with an accompanying web application at http://bacant.net.

Graphical Abstract



Introduction

The era of next-generation sequencing (NGS) took off in 2005 with the commercial release of massively parallel pyrosequencing (Margulies et al., 2005). The NGS technology developed rapidly in the past years and has made substantial improvements in terms of quality and yield. With the rapid decrease of sequencing costs, falling by as much as 80% year over year, whole genome sequencing of bacteria has become a routine method in diagnostic laboratories (Didelot et al., 2012). NGS applications include whole genome sequencing (WGS), targeted NGS and metagenomic NGS. Among them, the most common use of WGS is for simultaneous identification, typing, and/or antimicrobial susceptibility prediction of pathogens (Mitchell and Simner, 2019). One of the most exciting advantages of NGS is the ability to predict antimicrobial resistance genes (ARGs) and mobile genetic elements (MGEs) in bacteria, which allows the investigation of both, the organization and structure of such genetic features, and the epidemiology for the distribution of bacterial strains or virulence genes, including the spread and distribution of antibiotic-resistant bacteria as part of surveillance programs (Zhou et al., 2015; Mitchell and Simner, 2019). Every day, a massive number of bacterial genomes is being sequenced using NGS technology in laboratories across the globe, with genomes released at remarkable rates. With this huge amount of data available, it is important to extract project-relevant information easily. However, in publicly available databases, most of bacterial genomes are available as contigs which have been constructed employing auto-annotation algorithms. Over the years, highly efficient methods for bacterial genome annotation have been developed that do not require much user input.

Rapid Annotation using Subsystem Technology (RAST) is a widely used webserver for genome annotations of microbial species (Aziz et al., 2008). Although the performance using RAST-based annotation is very useful, several important limitations remain. For example, RAST will label many Open Reading Frames (ORFs) as “hypothetical proteins”, and the performance to identify ARGs and label them as such, is fairly limited as the algorithm is not tailored towards this purpose. Based on the RAST system, the Pathosystems Resource Integration Center (PATRIC) improved the data collection of ARGs, and provided users a more powerful analysis for both genomes and individual genes (Wattam et al., 2018). Another available annotation server is Resfinder which is managed by the Center for Genomic Epidemiology; it provides a convenient way of identifying acquired ARGs in sequenced bacterial isolates (Zankari et al., 2012). In addition to annotations of ARGs, some databases specifically designed to annotate MGEs such as insertion sequences (ISfinder) and integrons (INTEGRALL) and transposable elements (The Transposon Registry) have been created (Siguier et al., 2006; Moura et al., 2009; Tansirichaiya et al., 2019). ISs are abundant mobile elements in bacteria, which are responsible for the mobilization of many genes, including those mediating ARG (Razavi et al., 2020). Such ARGs are often found in the genetic context of specific ISs, while ISs flanking regions are diverse (Razavi et al., 2020). For example, a clear association of ARGs with class 1 integrons can be observed (Partridge et al., 2018). The analysis of which ISs are associated with ARG genes would help to discover novel AMGs (Razavi et al., 2020). In addition, there is a major interest to explore how ARGs spread via MGEs (Che et al., 2021). The early identification of ARGs in bacteria would facilitate surveillance and molecular diagnostics (Razavi et al., 2020). Also, inter/intra-species genetic transfer events of MGEs are responsible for the emergence and rapid spread of resistance (Subedi et al., 2018). Therefore, the knowledge of MGE-associated drug resistance is crucial for the monitoring of resistance of microbial species. Unfortunately, up to now, a rapid annotation tool of the combination of ARGs and MGEs is not available, and the generation of genbank files has to be done manually. Therefore, we created a new

program/pipeline called BacAnt, which rapidly and efficiently annotates ARGs, integrons, and transposable elements in a single step and generates a genbank file automatically which is compatible with the easyfig program for comparative genomic analysis.

Materials and Methods

Reference sequences

Three curated databases for BacAnt tool are used, including ResDB (resistance gene sequence database), IntegronDB (integron sequence database), and TransposonDB (transposon sequence database). We collected 5029 sequences from NCBI Bacterial Antimicrobial Resistance Reference Gene Database (PRJNA313047, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA313047, version: 2019-09-06.1) into ResDB at 2019-12-01. In addition, we collected 1094 sequences from INTEGRALL (http://integrall.bio.ua.pt, version: 2017-11-30) to be included into IntegronDB by 2019-12-01. We also collected 234 sequences from THE TRANSPOSON REGISTRY (https://transposon.lstmed.ac.uk, version: 2019-07-23) into the TransposonDB by 2019-12-01.

Program for identifying resistance genes and mobile elements

We created a python program (BacAnt) to identify resistance genes and mobile elements for bacteria nucleotide sequences with BLAST analysis. We first used the BLASTN program comparing input sequences with the reference database with an -evalue 10^{-5}. For the detection of integrons, we used integron_finder to predict possible integrons and used the BLASTN program comparing the integron sequence with integronDB database for the best match sequence (Didelot et al., 2012). We then filtered the raw results by identities and coverage (blast align match length/subject length). All results that pass the identity and coverage filter are retained for further analysis. The default threshold was set to 90% for identities and to 60% for coverage. Finally, we display the filtered results in text and genbank format, while also providing a visual output. Three types of annotations for the sequences are displayed in the same figure to guide the analysis of the genomic sequence.

BacAnt has six parameters. The user has two choices regarding the input sequence file: --nucleotide (-n), fasta format or --genbank (-g), genbank format. The required output path: --resultdir (-o). --databases (-d), reference databases, select all by default. --coverages (-c), coverage threshold, 60% by default. --identities (-i), identities threshold, 90% by default. In average, it takes about 2 minutes for each run (number of available cores: 6; used threads: 24; memory 64G).

Website for BacAnt

For the analysis to be performed online, we developed a website which we call http://bacant.net. The pipeline running on the server is based on Python/Django, which allows the user to upload sequence files for the rapid identification of ARGs and MGEs. The output format allows the display of graphic representations of the results. A demo report can be seen here: http://bacant.net/BacAnt/demo.

Datasets for validation of BacAnt

BacAnt was validated with 1100 genomes (Table S1) from eight species (*A. baumannii*, *Bacillus cereus*, *Clostridioides difficile*, *E coli*, *Listeria monocytogenes*, *S. enterica*, *S. aureus*, *Vibrio parahaemolyticus*). The BacAnt output was analyzed and compared with the results of NCBI AMRFinder . The parameters of BacAnt used in the study was: identity 0.9, coverage 0.6.

Examples analysis using BacAnt

Four genome sequences from NCBI, including *Acinetobacter baumannii* (2019.11.17), *Escherichia coli* (2019.5.7), *Salmonella enterica* (2019.6.4) and *Staphylococcus aureus* (2019.6.4) were used to illustrate the capabilities of our program. The accession number of the genomes used in this study were listed in Table S2 and Table S3. The number of drug-resistance genes and insertion sequences of each sample were identified in the BacAnt analysis and plotted as horizontal and vertical coordinates respectively. Scatter plots were created using the ggplot2 package in R (V3.6.2), together with the trend line (Wickham, 2016).

We created a network diagram by Cytoscape (v3.7.2) using the analysis results of BacAnt from four species genome sequences from NCBI (Shannon et al., 2003). Only resistance gene pairs with a distance of less than 10 kb and frequency of occurrence no less than 10 are extracted from the data to construct the network map. In addition, we also created a map from resistance genes and insertion sequence pairs with a distance of less than 10 kb and the frequency of occurrence no less than 10.

Results and Discussion

BacAnt is a browser-based platform to annotate DNA sequences, and to visualize the annotation results. When using the web interface, the user first has the choice to upload DNA sequences as Fasta, Seq or GenBank files (Figure 1). The user then has the option to select one or multiple databases, which include ResDB, IntegronDB or TransponsonDB for sequence annotation. After the DNA sequence is submitted, the python-based program BacAnt will identify ARGs and MGEs in the bacterial nucleotide sequence.

The output of BacAnt commences with a summary of the annotation, followed by up to four tables including the annotation result from each database; should they have been selected in the first step. The final part of the BacAnt output visualizes an annotation result which is combined from all three databases. All annotation results obtained by running BacAnt, including figures and genbank files, can then be downloaded. The genebank files generated by BacAnt are compatible with Easyfig (Sullivan et al., 2011). As an example we used *A. baumannii* MDR-ZJ06 (NC_017171.2) to display the result of the annotation by BacAnt (Figure 1). The annotation output of the MDR-ZJ06 strain shows that the isolate harbors 19 resistance genes, 124 integrons and 17 transposons.

BacAnt was validated with NCBI AMRFinder using 1100 selected genomes. The file output "AMR.possible.tsv" in the BacAnt result was used for analysis in NCBI AMRFinder to test which of the programs is able to identify a larger number of ARGs with high accuracy. Both programs reported similar results regarding the number of resistance genes (Table S4). However, the number of ARGS in BacAnt is slightly larger than that of NCBI AMRFinder. Some resistance genes were absent in the output of NCBI AMRFinder: aac(6')-Iaa(NC_003197,aminoglycoside N-acetyltransferase) in *S. enterica*, $bla_{\text{EC-15}}$(NG_049081,class C extended-spectrum beta-lactamase EC-15) in *E. coli*, BcII(NG_056058,BcII family subclass B1 metallo-beta-lactamase in *B. cereus*.

To investigate whether a relationship between ARGs and ISs exists, we extracted the number of ARGs and ISs from the results of the BacAnt analysis of four species genome sequences from NCBI. The results show that the number of ARGs does not correlate with the number of ISs in the four species (Pearson's R2 < 0.8, Figure 2). Previously, it was reported that at least eight MGEs were detected together with ARGs in *A. baumannii* (Leal et al., 2020). The plasmids were grouped into

three categories based on the DNA transfer machinery: conjugative, mobilizable and nonmobilizable (Smillie et al., 2010). Che *et al.* showed that most ARG genes are located in conjugative plasmids, which -together with Insertion Sequences- play the most important role in mediating the horizontal transfer of ARGs (Che et al., 2021). When the relationship between ARG and ISs were investigated, we did not analyze the location of ARGs which might explain why no significant correlation between the two was observed in our study.

We also calculated the pairwise association between ISs and ARGs. The results were subjected to a permutation test to differentiate between statistically significant associations and random chance (Razavi et al., 2020). Only statistically significant associations ($P < 0.001$) of ISs and ARGs were analyzed (Table S5). We identified commonly found ARG pairs that were in close proximity to the insertion sequence (< 10 kb apart), which allows the detection of gene cassettes that may play an important role in evolution, regulation and ARG exchange. ARG cassettes are generally small (2-7) and specific to the species we investigated (Figure 3). The ARG cassettes for *A. baumannii* and *E. coli* were larger and more stable than the cassettes in *S. aureus*, which is consistent with a previously published observation (Chng et al., 2020). In the case of *A. baumannii*, we identified two ARG cassettes, with one containing the genes *mph(E), msr(E), armA, aadA1, sul1, aac(6')-Ib* and *catB8*. The ARG cassette which contained *mph(E), msr(E)* and *armA* was described previously (Chng et al., 2020). For *E. coli*, the program identified a stable small cassette that shows overlap with that of *A. baumannii*, including *sul2, aph(3")-Ib* and *aph(6)-Id*. When genomes of *S. aureus* were analyzed, the program BacAnt found two ARG cassettes; the first one contained the genes *bla*PC1 and *bla*R1, while the second one encoded for *mecR1* and erm(A).

We also identified commonly found ARG-IS pairs that were in close proximity to the insertion sequence (< 10 kb apart). For the ARG-IS pairs, IS*Vsa3* containing the genes *aph(6)-Id*, *aph(3")-Ib* and *tet(B)* comprised the top three ARG-IS pairs in *A. baumannii* (Figure 4A). For ARGs number, IS*26*, IS*Aba1* and IS*Vsa3* were the top three active ISs (Figure 5A). IS*26* was the most abundant IS in *E. coli* and *S. enterica* (Figure 5B, Figure 5C). In *S. aureus*, *mecA* with diverse IS (including IS*257-3*, IS*431mec*, IS*257R1*, IS*257-1*, IS*257R2*, IS*431L* and IS*431R*) are the top seven ARG-IS pairs (Figure 4D, Figure 5D). The result of this study confirmed that IS*6* family elements IS*26* and IS*257* play an important role in the dissemination of ARGs in *A. baumannii*, *E. coli*, *S. enterica* and *S. aureus* (Partridge et al., 2018). As previously reported, we also observed notable differences between important MGEs in *A. baumannii*, *E. coli*, *S. enterica* and *S. aureus* (Partridge et al., 2018).

Using BacAnt, we also explored the relationship between ARGs and transposons. Tn*6292* was the most commonly observed transposon containing ARGs in *A. baumannii*, *E. coli* and *S. enterica* (Figure 6A, Figure 6B, Figure 6C). In *S. aureus* the most prevalent transposon with ARGs was identified to be Tn*552* (Figure 6D). Tn*6292* belongs to the Tn*3*-family and harbored an IS*26* at the right end (Chen et al., 2020). Tn*6292* also contained a quinolone resistance region *qnrS1* (Feng et al., 2016). Multidrug-resistance bacteria containing Tn*6292* are commonly observed in China (Li et al., 2018), and possibly accelerate the emergence and spread of multidrug-resistant pathogens. Tn*552* belonged to Tn*7* family, comprised of BlaZ, BlaR1 and BlaI proteins. BlaR1 is the sensor protein for the extracellular β-lactam antibiotics. The overproduction of the beta-lactamase BlaZ were responsible of β-lactam resistance. Tn*552*-like element was thought as the origin of the all β-lactamase genes in staphylococci (Gregory et al., 1997).

In order to be able to extract the maximum amount of information from whole genome sequence data, we need the improve annotation and analysis methods for MGEs (Partridge et al., 2018). In this work, we created a webserver that is easy to use and allows the annotation of ARGs, ISs, integron,

and transposable elements at the same time. The pipeline generates genbank files automatically, which are compatible with easyfig for comparative genomic analysis, which will accelerate the bioinformatics analysis of ARG-related sequences.

Data Availability Statement

BacAnt code and standalone software package are available at https://github.com/xthua/bacant with an accompanying web application at http://bacant.net.

Author Contributions

HC, YY and XH designed the study. QL, MW and WH established the BacAnt. XH, QL, MD, WH, JW and BL analyzed the bioinformatics data. XH, JH, WH and SL wrote the manuscript.

Funding

This work was supported by the grants from the National Natural Science Foundation of China [31970128, 31770142, 81861138054, 31670135], the National Key Research and Development Project [2018YFE0101800], Zhejiang Province Medical Platform [2020RC075, 2017KY571] and 2019 Jiaxing Key Discipline of Medicine--infection disease (Supporting Subject, no 2019-zc-02).

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References：

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9**,** 75. doi: 10.1186/1471-2164-9-75.

Che, Y., Yang, Y., Xu, X., Brinda, K., Polz, M.F., Hanage, W.P., et al. (2021). Conjugative plasmids interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes. *Proc Natl Acad Sci U S A* 118(6). doi: 10.1073/pnas.2008731118.

Chen, Q., Lin, Y., Li, Z., Lu, L., Li, P., Wang, K., et al. (2020). Characterization of a New Transposon, Tn6696, on a bla NDM- 1-Carrying Plasmid From Multidrug-Resistant Enterobacter cloacae ssp. dissolvens in China. *Front Microbiol* 11**,** 525479. doi: 10.3389/fmicb.2020.525479.

Chng, K.R., Li, C., Bertrand, D., Ng, A.H.Q., Kwah, J.S., Low, H.M., et al. (2020). Cartography of opportunistic pathogens and antibiotic resistance genes in a tertiary hospital environment. *Nat Med* 26(6)**,** 941-951. doi: 10.1038/s41591-020-0894-4.

Didelot, X., Bowden, R., Wilson, D.J., Peto, T.E.A., and Crook, D.W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet* 13(9)**,** 601-612. doi: 10.1038/nrg3226.

Feng, W., Zhou, D., Wang, Q., Luo, W., Zhang, D., Sun, Q., et al. (2016). Dissemination of IMP-4-encoding pIMP-HZ1-related plasmids among Klebsiella pneumoniae and Pseudomonas aeruginosa in a Chinese teaching hospital. *Sci Rep* 6**,** 33419. doi: 10.1038/srep33419.

Leal, N.C., Campos, T.L., Rezende, A.M., Docena, C., Mendes-Marques, C.L., de Sa Cavalcanti, F.L., et al. (2020). Comparative Genomics of Acinetobacter baumannii Clinical Strains From Brazil Reveals Polyclonal Dissemination and Selective Exchange of Mobile Genetic Elements Associated With Resistance Genes. *Front Microbiol* 11**,** 1176. doi: 10.3389/fmicb.2020.01176.

Li, B., Feng, J., Zhan, Z., Yin, Z., Jiang, Q., Wei, P., et al. (2018). Dissemination of KPC-2-Encoding IncX6 Plasmids Among Multiple Enterobacteriaceae Species in a Single Chinese Hospital. *Front Microbiol* 9**,** 478. doi: 10.3389/fmicb.2018.00478.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437(7057)**,** 376-380. doi: 10.1038/nature03959.

Mitchell, S.L., and Simner, P.J. (2019). Next-Generation Sequencing in Clinical Microbiology: Are We There Yet? *Clin Lab Med* 39(3)**,** 405-418. doi: 10.1016/j.cll.2019.05.003.

Moura, A., Soares, M., Pereira, C., Leitao, N., Henriques, I., and Correia, A. (2009). INTEGRALL: a database and search engine for integrons, integrases and gene cassettes. *Bioinformatics* 25(8)**,** 1096-1098. doi: 10.1093/bioinformatics/btp105.

Partridge, S.R., Kwong, S.M., Firth, N., and Jensen, S.O. (2018). Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin Microbiol Rev* 31(4). doi: 10.1128/CMR.00088-17.

Razavi, M., Kristiansson, E., Flach, C.F., and Larsson, D.G.J. (2020). The Association between Insertion Sequences and Antibiotic Resistance Genes. *mSphere* 5(5). doi: 10.1128/mSphere.00418-20.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11)**,** 2498-2504. doi: 10.1101/gr.1239303.

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 34(Database issue)**,** D32-36. doi: 10.1093/nar/gkj014.

Smillie, C., Garcillan-Barcia, M.P., Francia, M.V., Rocha, E.P., and de la Cruz, F. (2010). Mobility of plasmids. *Microbiol Mol Biol Rev* 74(3)**,** 434-452. doi: 10.1128/MMBR.00020-10.

Subedi, D., Vijay, A.K., and Willcox, M. (2018). Overview of mechanisms of antibiotic resistance in Pseudomonas aeruginosa: an ocular perspective. *Clin Exp Optom* 101(2)**,** 162-171. doi: 10.1111/cxo.12621.

Sullivan, M.J., Petty, N.K., and Beatson, S.A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27(7)**,** 1009-1010. doi: 10.1093/bioinformatics/btr039.

Tansirichaiya, S., Rahman, M.A., and Roberts, A.P. (2019). The Transposon Registry. *Mob DNA* 10**,** 40. doi: 10.1186/s13100-019-0182-3.

Wattam, A.R., Brettin, T., Davis, J.J., Gerdes, S., Kenyon, R., Machi, D., et al. (2018). Assembly, Annotation, and Comparative Genomics in PATRIC, the All Bacterial Bioinformatics Resource Center. *Methods Mol Biol* 1704**,** 79-101. doi: 10.1007/978-1-4939-7463-4_4.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*. doi: https://ggplot2.tidyverse.org.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012). Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67(11)**,** 2640-2644. doi: 10.1093/jac/dks261.

Zhou, K., Lokate, M., Deurenberg, R.H., Arends, J., Lo-Ten Foe, J., Grundmann, H., et al. (2015). Characterization of a CTX-M-15 Producing Klebsiella Pneumoniae Outbreak Strain Assigned to a Novel Sequence Type (1427). *Front Microbiol* 6**,** 1250. doi: 10.3389/fmicb.2015.01250.

Figure legends:

Figure 1

Screenshots of the BacAnt web interface. Users upload an assembled file from their local personal computer and select the desired annotation database. “Framework” lists the public database and tools integrated in BacAnt; “Tools” functions for whole genome sequence annotation based on user uploaded sequence(s); “Visualization” allows to visualize the annotation results for uploaded sequence(s).

Figure 2

The relationship between the number of drug-resistance genes (ARGs) and insertion sequences (IS) from four species (A: *A. baumannii*, B: *E. coli*, C: *S. enterica*, D: *S. aureus*). The number of ARGs and ISs of each sample were extracted from the results of the BacAnt analysis, and the number was plotted as the horizontal and vertical coordinates respectively. The scatter plot was created using the ggplot2 package in R (V3.6.2), and a trend line generated.

Figure 3

Genomic proximity of antibiotic resistance genes (ARGs) in four species (A: *A. baumannii*, B: *E. coli*, C: *S. enterica*, D: *S. aureus*). ARG pairs with a distance of less than 10 kb and frequency of occurrence no less than 10 were screened to construct the network. Diamonds are used to represent ARGs, solid lines are used to represent gene pairs whose frequency is not less than 80% of the total number and labeled as pink, dotted lines represent gene pairs with frequency less than 80% of the total number and are marked as gray, and the width of the line indicates the frequency. Different antibiotics are displayed in different colors, corresponding to the ARGs.

Figure 4

Genomic proximity of antibiotic resistance genes (ARGs) and Insertion sequences (ISs) in four species: A) *A. baumannii*, B) *E. coli*, C) *S. enterica* and D) *S. aureus*. ARGs and ISs pairs with a distance of less than 10 kb and frequency of occurrence no less than 10 were screened to construct the network. Diamond symbols are used to represent ARGs, circles for ISs. The solid line is the gene pair whose frequency is not less than 80% of the total number and labeled as orange, the dotted line represents the gene pair whose frequency is less than 80% of the total number and labeled as sky blue, and the width of the line indicates the frequency. Different colors are assigned to the antibiotics corresponding to the drug-resistant genes, the insertion sequences are displayed in gray.

Figure 5

Distribution of Insertion sequences (ISs) with statistically significant association with different types of antibiotic resistance genes (ARGs) within a 10 kb distance. Different colors show various IS, and the size of the circles indicate the presence of their associations with ARGs in A) *A. baumannii*, B) *E. coli*, C) *S. enterica* and D) *S. aureus*.

Figure 6

Distribution of transposons with statistically significant association with different types of ARGs within a 10 kb distance. Different colors show various transposons, the size of the circles indicate the presence of their associated ARGs in A) *A. baumannii*, B) *E. coli*, C) *S. enterica* and D) *S. aureus*.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9**,** 75. doi: 10.1186/1471-2164-9-75.

Che, Y., Yang, Y., Xu, X., Brinda, K., Polz, M.F., Hanage, W.P., et al. (2021). Conjugative plasmids interact with insertion sequences to shape the horizontal transfer of antimicrobial resistance genes. *Proc Natl Acad Sci U S A* 118(6). doi: 10.1073/pnas.2008731118.

Chen, Q., Lin, Y., Li, Z., Lu, L., Li, P., Wang, K., et al. (2020). Characterization of a New Transposon, Tn6696, on a bla NDM- 1-Carrying Plasmid From Multidrug-Resistant Enterobacter cloacae ssp. dissolvens in China. *Front Microbiol* 11**,** 525479. doi: 10.3389/fmicb.2020.525479.

Chng, K.R., Li, C., Bertrand, D., Ng, A.H.Q., Kwah, J.S., Low, H.M., et al. (2020). Cartography of opportunistic pathogens and antibiotic resistance genes in a tertiary hospital environment. *Nat Med* 26(6)**,** 941-951. doi: 10.1038/s41591-020-0894-4.

Didelot, X., Bowden, R., Wilson, D.J., Peto, T.E.A., and Crook, D.W. (2012). Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet* 13(9)**,** 601-612. doi: 10.1038/nrg3226.

Feng, W., Zhou, D., Wang, Q., Luo, W., Zhang, D., Sun, Q., et al. (2016). Dissemination of IMP-4-encoding pIMP-HZ1-related plasmids among Klebsiella pneumoniae and Pseudomonas aeruginosa in a Chinese teaching hospital. *Sci Rep* 6**,** 33419. doi: 10.1038/srep33419.

Gregory, P.D., Lewis, R.A., Curnock, S.P., and Dyke, K.G. (1997). Studies of the repressor (BlaI) of beta-lactamase synthesis in Staphylococcus aureus. *Mol Microbiol* 24(5)**,** 1025-1037. doi: 10.1046/j.1365-2958.1997.4051770.x.

Leal, N.C., Campos, T.L., Rezende, A.M., Docena, C., Mendes-Marques, C.L., de Sa Cavalcanti, F.L., et al. (2020). Comparative Genomics of Acinetobacter baumannii Clinical Strains From Brazil Reveals Polyclonal Dissemination and Selective Exchange of Mobile Genetic Elements Associated With Resistance Genes. *Front Microbiol* 11**,** 1176. doi: 10.3389/fmicb.2020.01176.

Li, B., Feng, J., Zhan, Z., Yin, Z., Jiang, Q., Wei, P., et al. (2018). Dissemination of KPC-2-Encoding IncX6 Plasmids Among Multiple Enterobacteriaceae Species in a Single Chinese Hospital. *Front Microbiol* 9**,** 478. doi: 10.3389/fmicb.2018.00478.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437(7057)**,** 376-380. doi: 10.1038/nature03959.

Mitchell, S.L., and Simner, P.J. (2019). Next-Generation Sequencing in Clinical Microbiology: Are We There Yet? *Clin Lab Med* 39(3)**,** 405-418. doi: 10.1016/j.cll.2019.05.003.

Moura, A., Soares, M., Pereira, C., Leitao, N., Henriques, I., and Correia, A. (2009). INTEGRALL: a database and search engine for integrons, integrases and gene cassettes. *Bioinformatics* 25(8)**,** 1096-1098. doi: 10.1093/bioinformatics/btp105.

Partridge, S.R., Kwong, S.M., Firth, N., and Jensen, S.O. (2018). Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin Microbiol Rev* 31(4). doi: 10.1128/CMR.00088-17.

Razavi, M., Kristiansson, E., Flach, C.F., and Larsson, D.G.J. (2020). The Association between Insertion Sequences and Antibiotic Resistance Genes. *mSphere* 5(5). doi: 10.1128/mSphere.00418-20.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11)**,** 2498-2504. doi: 10.1101/gr.1239303.

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 34(Database issue)**,** D32-36. doi: 10.1093/nar/gkj014.

Smillie, C., Garcillan-Barcia, M.P., Francia, M.V., Rocha, E.P., and de la Cruz, F. (2010). Mobility of plasmids. *Microbiol Mol Biol Rev* 74(3)**,** 434-452. doi: 10.1128/MMBR.00020-10.

Subedi, D., Vijay, A.K., and Willcox, M. (2018). Overview of mechanisms of antibiotic resistance in Pseudomonas aeruginosa: an ocular perspective. *Clin Exp Optom* 101(2)**,** 162-171. doi: 10.1111/cxo.12621.

Sullivan, M.J., Petty, N.K., and Beatson, S.A. (2011). Easyfig: a genome comparison visualizer. *Bioinformatics* 27(7)**,** 1009-1010. doi: 10.1093/bioinformatics/btr039.

Tansirichaiya, S., Rahman, M.A., and Roberts, A.P. (2019). The Transposon Registry. *Mob DNA* 10**,** 40. doi: 10.1186/s13100-019-0182-3.

Wattam, A.R., Brettin, T., Davis, J.J., Gerdes, S., Kenyon, R., Machi, D., et al. (2018). Assembly, Annotation, and Comparative Genomics in PATRIC, the All Bacterial Bioinformatics Resource Center. *Methods Mol Biol* 1704**,** 79-101. doi: 10.1007/978-1-4939-7463-4_4.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*. doi: https://ggplot2.tidyverse.org.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012). Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67(11)**,** 2640-2644. doi: 10.1093/jac/dks261.

Zhou, K., Lokate, M., Deurenberg, R.H., Arends, J., Lo-Ten Foe, J., Grundmann, H., et al. (2015). Characterization of a CTX-M-15 Producing Klebsiella Pneumoniae Outbreak Strain Assigned to a Novel Sequence Type (1427). *Front Microbiol* 6**,** 1250. doi: 10.3389/fmicb.2015.01250.

BacAnt

BacAnt is a consolidation tool for finding the Resistance genes, Insertion elements and Transposon regions from genome sequences. It provide curated database for transfer elements search and annotation.Visualization of result will work well for your job.You can upload the genome fasta sequences and Getting Start ↓

Submit Job

Data input: Please upload FASTA or GENBANK file

Resistance: ResDB

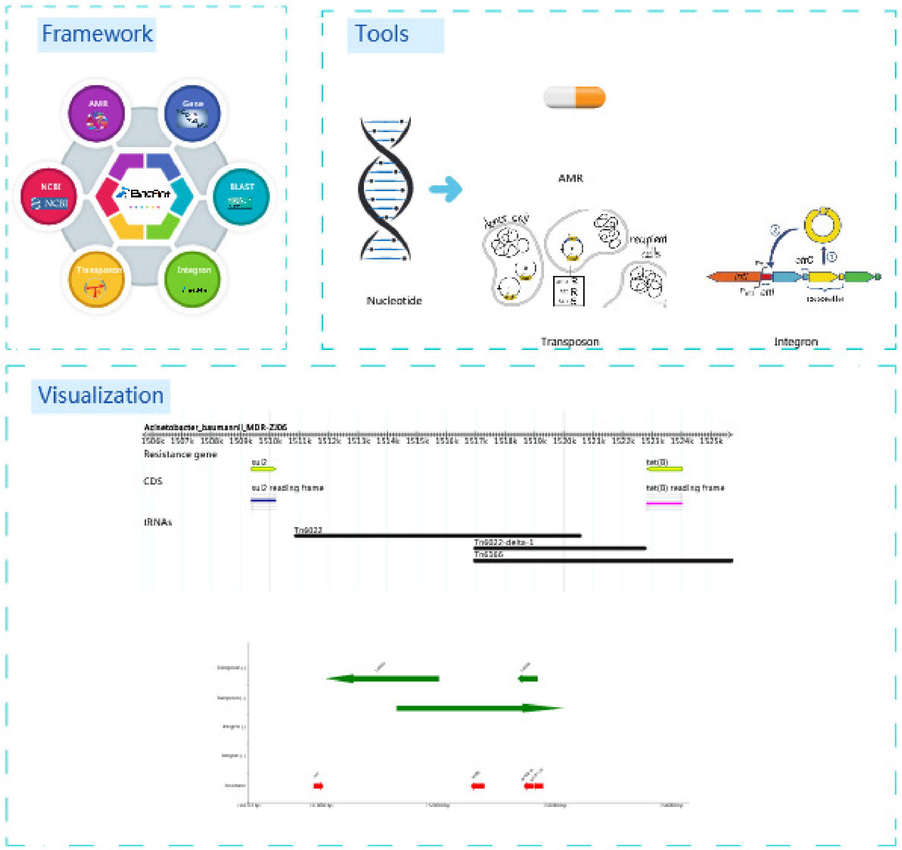
Transfer element: IntegronDB,TransposonDB

Identity(%): 90

Coverage(%): 60

Email:

SUBMIT

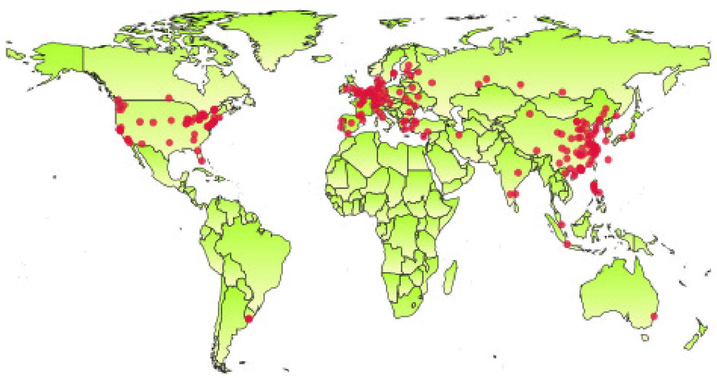


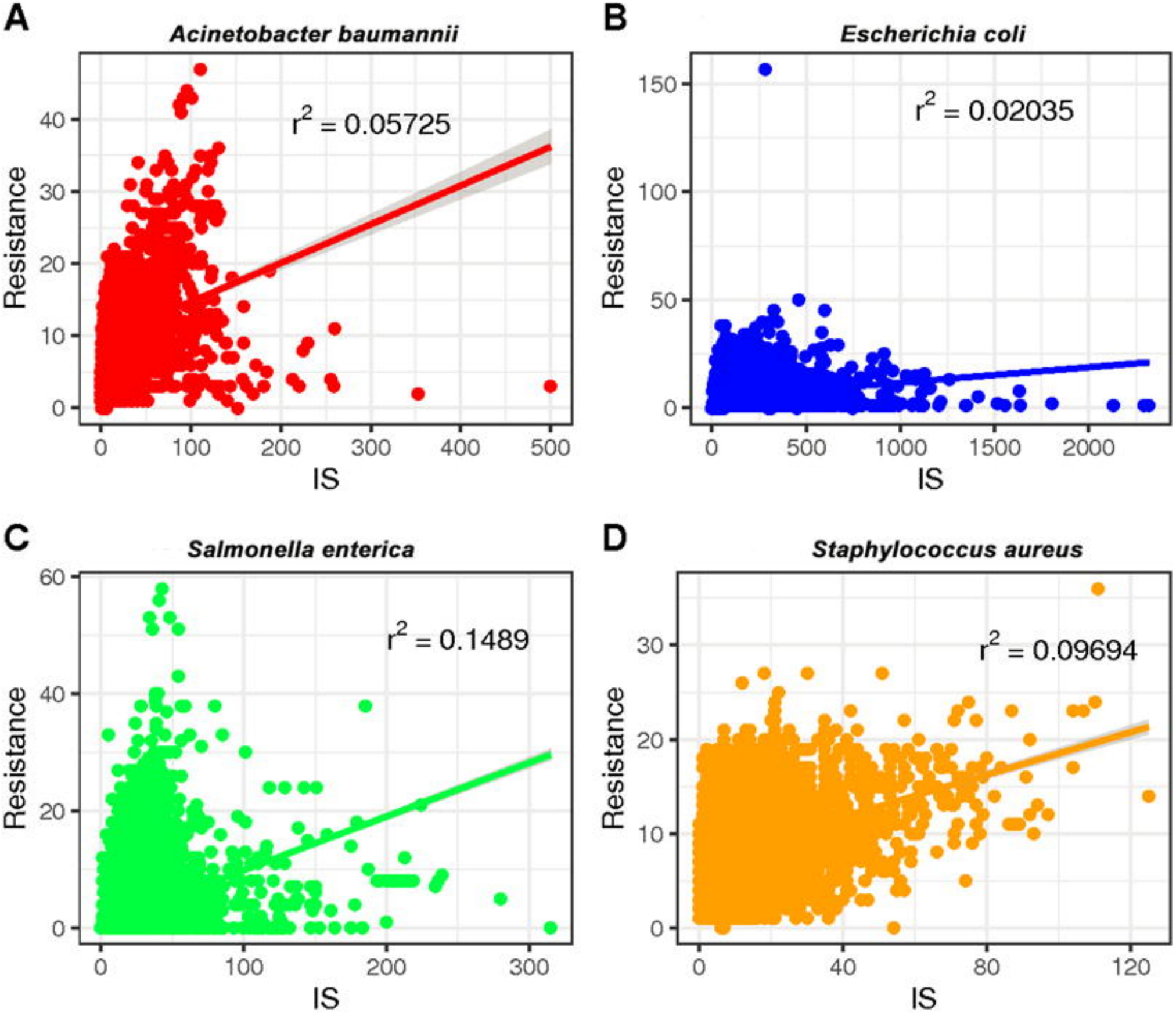
Upload the genome assembly fasta file(Uncompressed format) example.fasta, you can see demo report here.

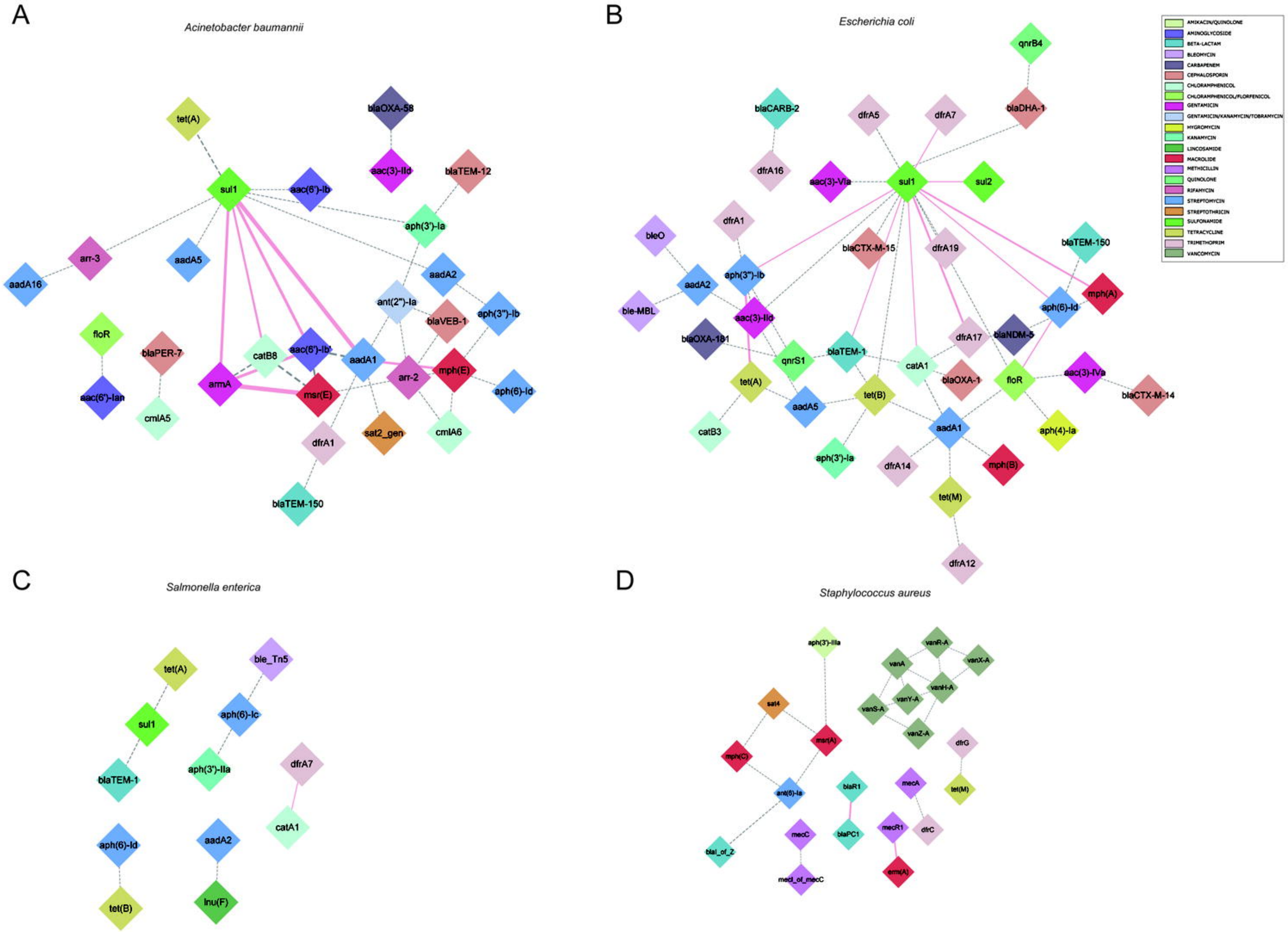
If you have any questions or advice about our website, you can contact us. For publication of results,please cite here

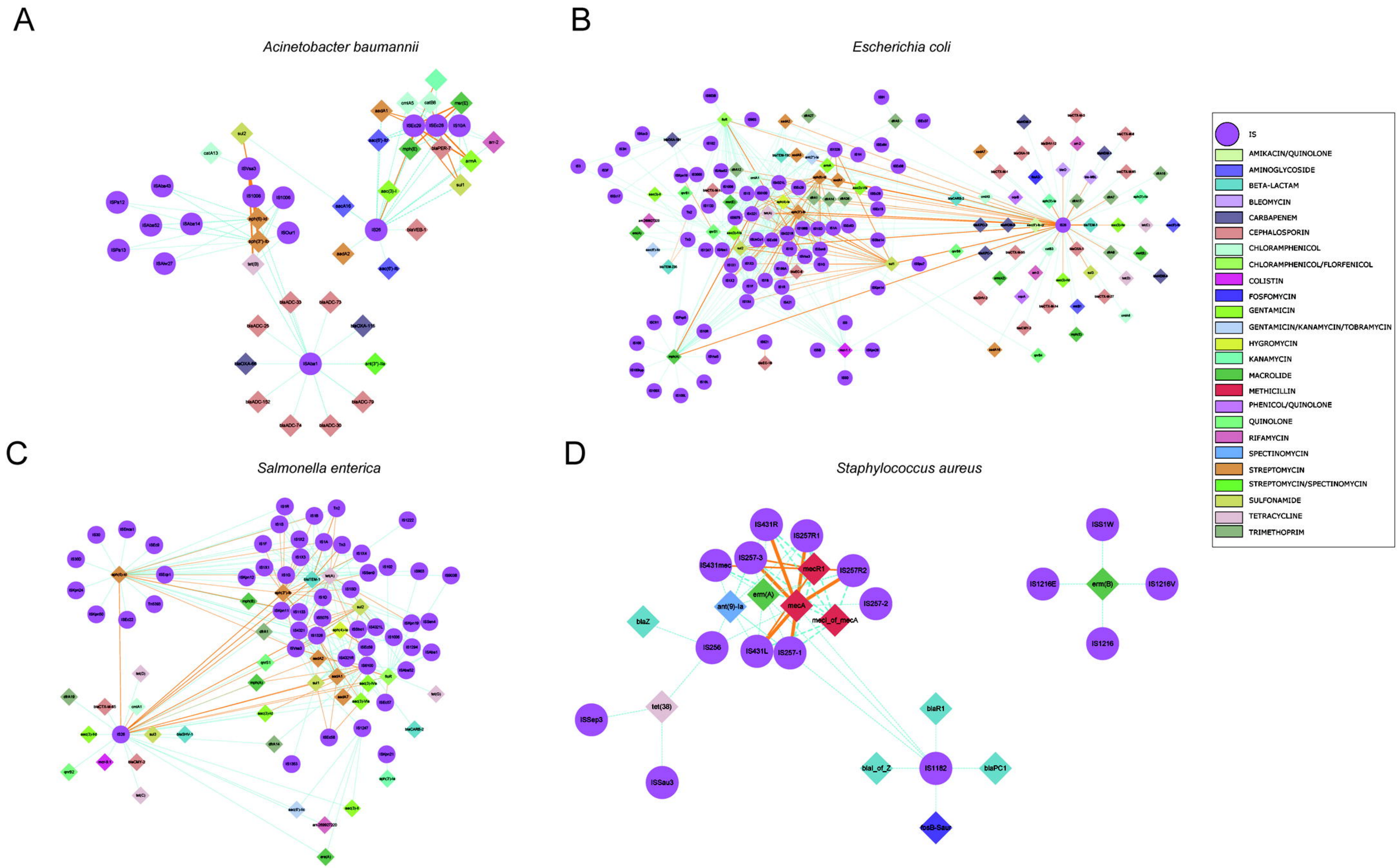
If you want to know more, fetch source code on github

Copyright © 2020 Key laboratory of Microbiol technology and Bioinformatics of Zhejiang Province All Rights Reserved. 备案序号：浙ICP备09014071号-10

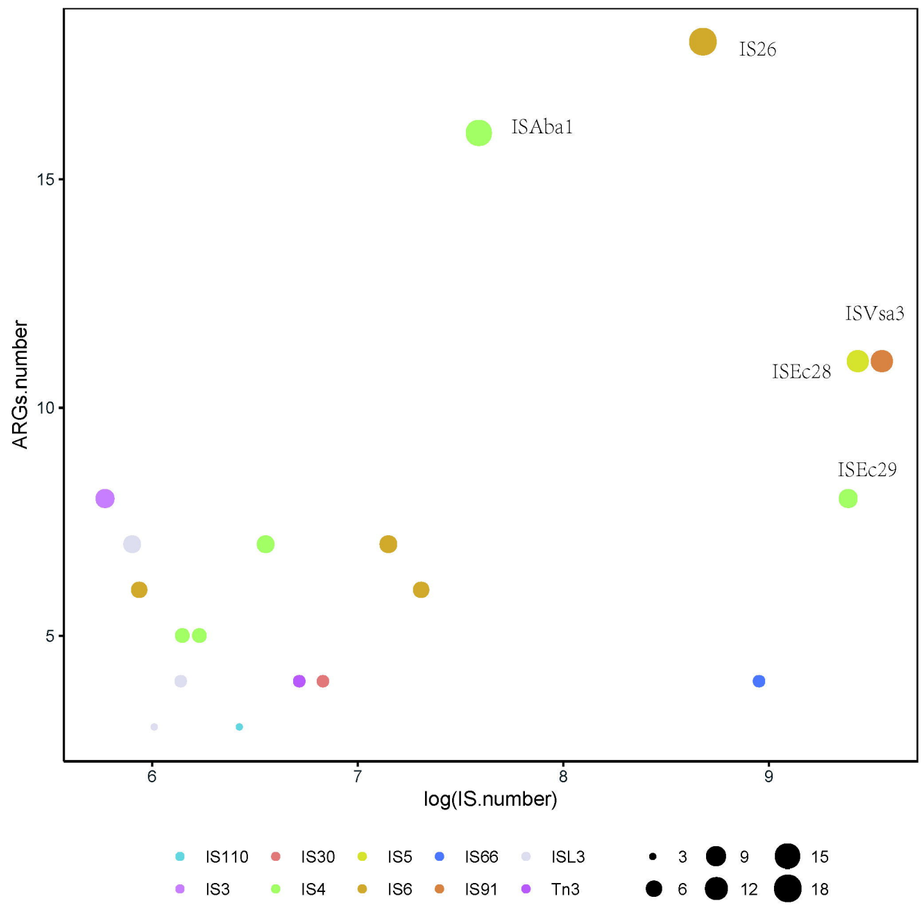



A
Acinetobacter baumannii
r2 = 0.05725
Resistance
IS
B
Escherichia coli
r2 = 0.02035
Resistance
IS
C
Salmonella enterica
r2 = 0.1489
Resistance
IS
D
Staphylococcus aureus
r2 = 0.09694
Resistance
IS

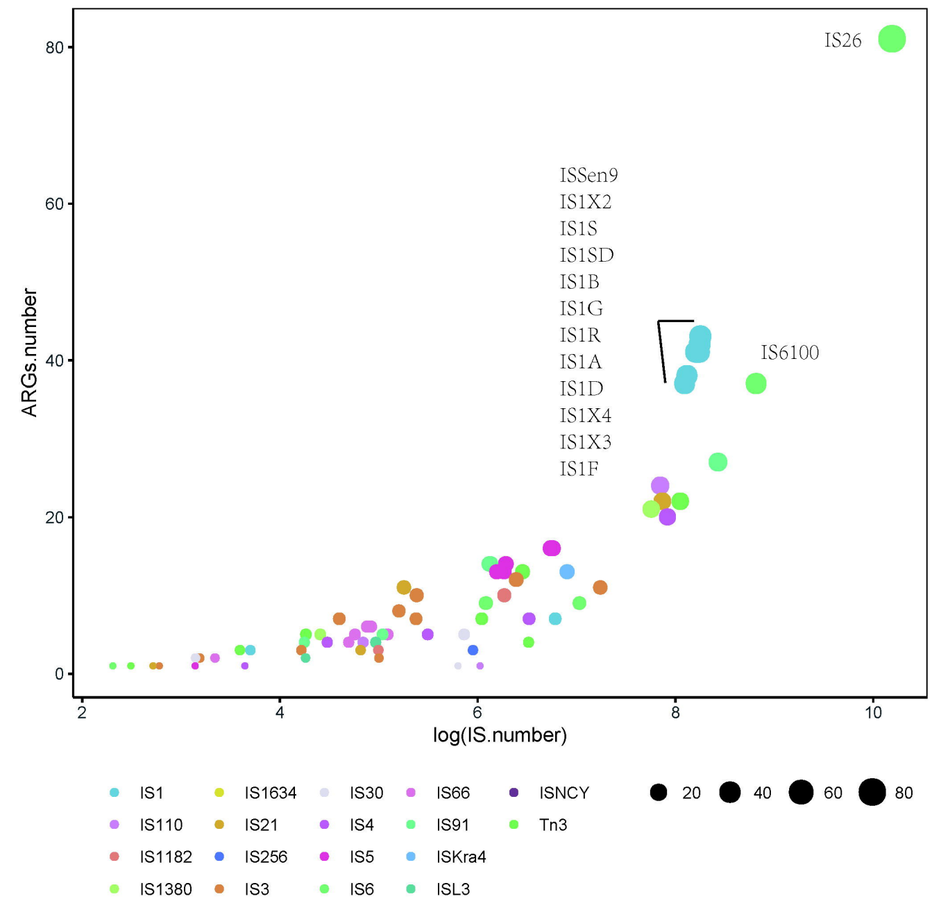

A
Acinetobacter baumannii
tet(A)
blaOXA-58
aac(3)-IId
blaTEM-12
sul1
aac(6')-Ib
aph(3')-Ia
arr-3
aadA5
aadA16
aadA2
ant(2")-Ia
blaVEB-1
aph(3")-Ib
floR
blaPER-7
catB8
aac(6')-Ib'
aadA1
arr-2
mph(E)
armA
msr(E)
aac(6')-Iah
aph(6)-Id
cmlA5
dfrA1
sat2_gen
cmlA6
blaTEM-150
B
Escherichia coli
AMIKACIN/QUINOLONE
AMINOGLYCOSIDE
BETA-LACTAM
BLEOMYCIN
CARBAPENEM
CEPHALOSPORIN
CHLORAMPHENICOL
CHLORAMPHENICOL/FLORFENICOL
GENTAMICIN
GENTAMICIN/KANAMYCIN/TOBRAMYCIN
HYGROMYCIN
KANAMYCIN
LINCOSAMIDE
MACROLIDE
METHICILLIN
QUINOLONE
RIFAMYCIN
STREPTOMYCIN
STREPTOTHRICIN
SULFONAMIDE
TETRACYCLINE
TRIMETHOPRIM
VANCOMYCIN
qnrB4
blaCARB-2
dfrA5
dfrA7
blaDHA-1
dfrA16
aac(3)-VIa
sul1
sul2
dfrA1
bleO
blaCTX-M-15
dfrA19
blaTEM-150
aadA2
aph(3")-Ib
mph(A)
ble-MBL
aac(3)-IId
aph(6)-Id
blaOXA-181
dfrA17
blaNDM-5
qnrS1
blaTEM-1
catA1
aac(3)-IVa
tet(A)
blaOXA-1
floR
blaCTX-M-14
tet(B)
aadA5
catB3
aadA1
aph(4)-Ia
aph(3')-Ia
dfrA14
mph(B)
tet(M)
dfrA12
C
Salmonella enterica
tet(A)
ble_Tn5
sul1
aph(6)-Ic
blaTEM-1
aph(3')-IIa
dfrA7
catA1
aph(6)-Id
aadA2
tet(B)
lnu(F)
D
Staphylococcus aureus
aph(3')-IIIa
vanR-A
vanA
vanX-A
vanH-A
sal4
vanY-A
vanS-A
vanZ-A
msr(A)
dfrG
mph(C)
ant(6)-Ia
blaR1
mecA
tet(M)
blaI_of_Z
mecC
blaPC1
mecR1
dfrC
mecI_of_mecC
erm(A)


A
Acinetobacter baumannii
B
Escherichia coli
C
Salmonella enterica
D
Staphylococcus aureus
IS
AMIKACIN/QUINOLONE
AMINOGLYCOSIDE
BETA-LACTAM
BLEOMYCIN
CARBAPENEM
CEPHALOSPORIN
CHLORAMPHENICOL
CHLORAMPHENICOL/FLORFENICOL
COLISTIN
FOSFOMYCIN
GENTAMICIN
GENTAMICIN/KANAMYCIN/TOBRAMYCIN
HYGROMYCIN
KANAMYCIN
MACROLIDE
METHICILLIN
PHENICOL/QUINOLONE
QUINOLONE
RIFAMYCIN
SPECTINOMYCIN
STREPTOMYCIN
STREPTOMYCIN/SPECTINOMYCIN
SULFONAMIDE
TETRACYCLINE
TRIMETHOPRIM
IS431R
IS257R1
IS431mec
IS257-3
mecR1
IS257R2
erm(A)
ant(9)-Ia
mecA
IS257-2
mecI_of_mecA
blaZ
IS256
IS431L
IS257-1
tet(38)
ISSep3
ISSau3
blaR1
blaI_of_Z
IS1182
blaPC1
dsrB-Sau
ISS1W
IS1216E
erm(B)
IS1216V
IS1216

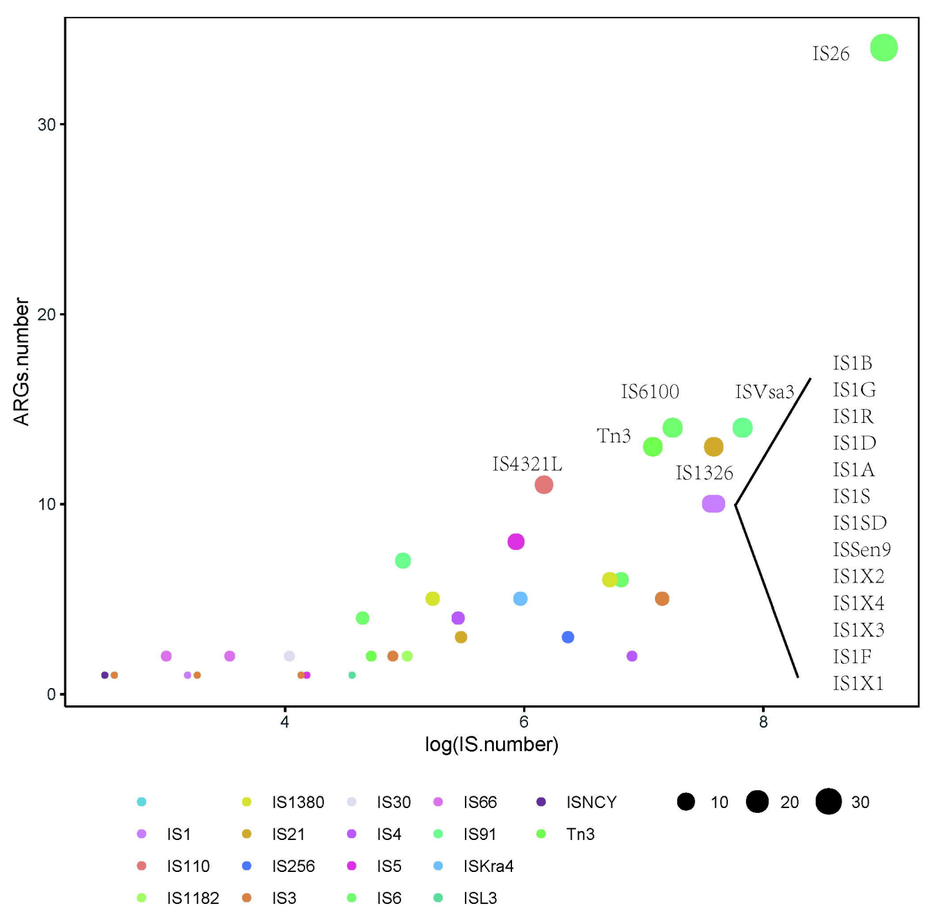
A


IS26
ISAba1
ISVsa3
ISEc28
ISEc29
ARGs.number
log(IS.number)
IS110
IS30
IS5
IS66
ISL3
IS3
IS4
IS6
IS91
Tn3
3
9
15
6
12
18

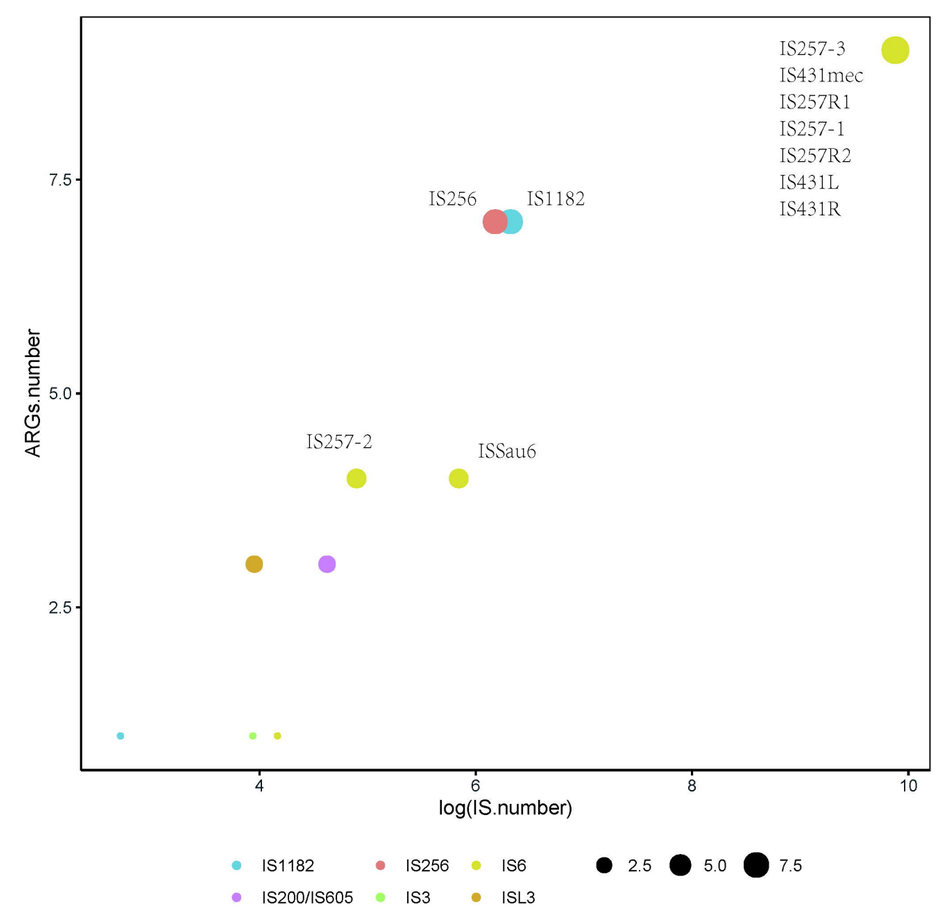
B

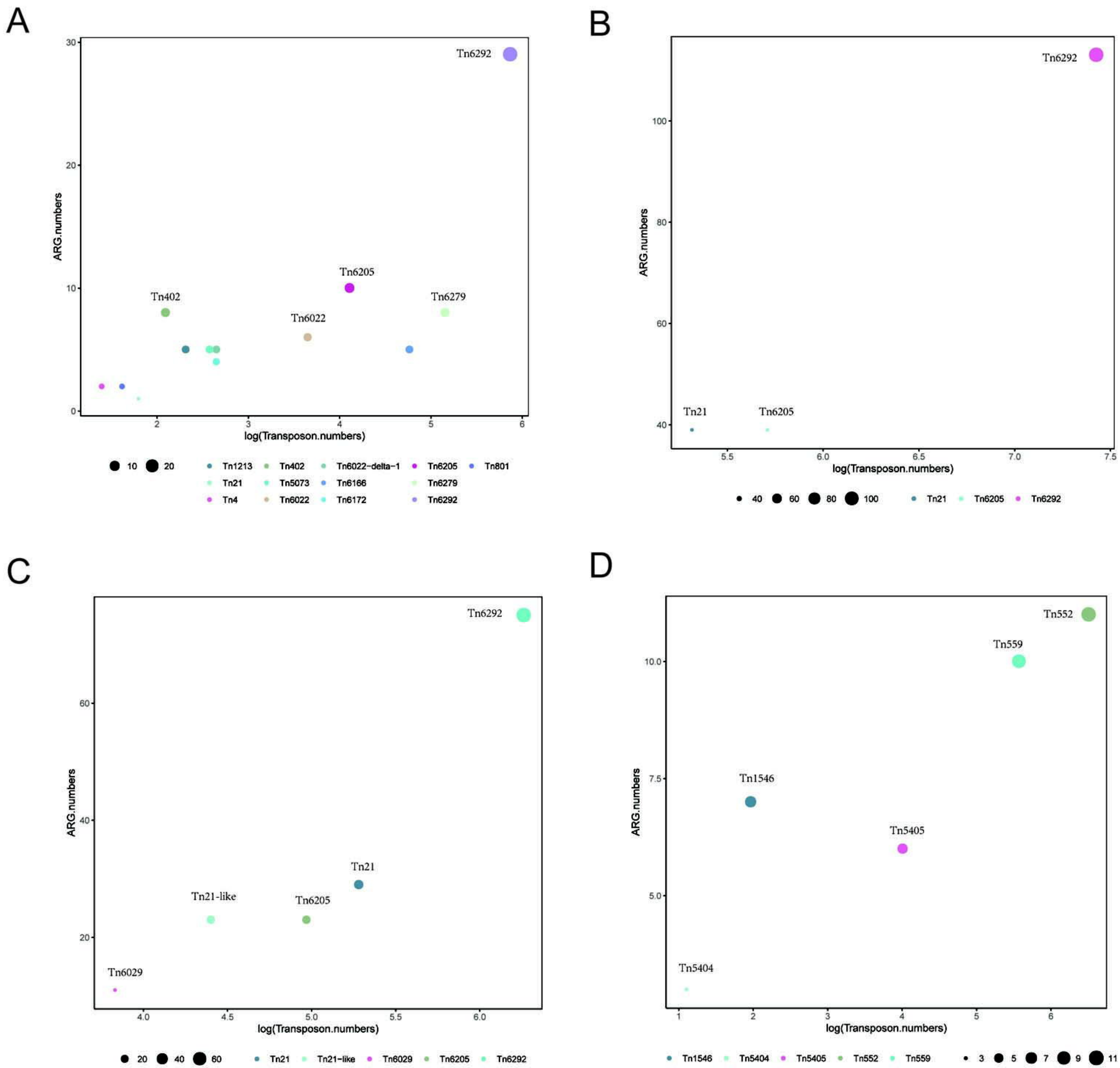

IS26
ISSen9
IS1X2
IS1S
IS1SD
IS1B
IS1G
IS1R
IS1A
IS1D
IS1X4
IS1X3
IS1F
IS6100
ARGs.number
log(IS.number)
IS1
IS1634
IS30
IS66
ISNCY
IS110
IS21
IS4
IS91
Tn3
IS1182
IS256
IS5
ISKra4
IS1380
IS3
IS6
ISL3
20
40
60
80

C


IS26
IS6100
ISVsa3
Tn3
IS4321L
IS1326
IS1B
IS1G
IS1R
IS1D
IS1A
IS1S
IS1SD
ISSen9
IS1X2
IS1X4
IS1X3
IS1F
IS1X1
ARGs.number
log(IS.number)
IS1380
IS30
IS66
ISNCY
IS1
IS21
IS4
IS91
Tn3
IS110
IS256
IS5
ISKra4
IS1182
IS3
IS6
ISL3
10
20
30

D


IS257-3
IS431mec
IS257R1
IS257-1
IS257R2
IS431L
IS431R
IS256
IS1182
IS257-2
ISSau6
ARGs.number
log(IS.number)
IS1182
IS256
IS6
IS200/IS605
IS3
ISL3
2.5
5.0
7.5


A
Tn6292
Tn6205
Tn402
Tn6279
Tn6022
ARG.numbers
log(Transposon.numbers)
10
20
Tn1213
Tn402
Tn6022-delta-1
Tn6205
Tn801
Tn21
Tn5073
Tn6166
Tn6279
Tn4
Tn6022
Tn6172
Tn6292
B
Tn6292
Tn21
Tn6205
ARG.numbers
log(Transposon.numbers)
40
60
80
100
Tn21
Tn6205
Tn6292
C
Tn6292
Tn21
Tn21-like
Tn6205
Tn6029
ARG.numbers
log(Transposon.numbers)
20
40
60
Tn21
Tn21-like
Tn6029
Tn6205
Tn6292
D
Tn552
Tn559
Tn1546
Tn5405
Tn5404
ARG.numbers
log(Transposon.numbers)
Tn1546
Tn5404
Tn5405
Tn552
Tn559
3
5
7
9
11