Journal of Bioinformatics and Computational Biology

AUTOMATIC PHYLOGENETIC CLASSIFICATION OF BACTERIAL BETA-LACTAMASE SEQUENCES INCLUDING STRUCTURAL AND ANTIBIOTIC SUBSTRATE PREFERENCE INFORMATION

JIANMIN MA

*Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR),*
30 Biopolis Street, #07-01, Matrix, Singapore 138671
majm@bii.a-star.edu.sg

FRANK EISENHABER

*Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR),*
30 Biopolis Street, #07-01, Matrix, Singapore 138671
and
Department of Biological Sciences (DBS), NationalUniversity of Singapore (NUS),
8 Medical Drive, Singapore 117597
and
School of Computer Engineering (SCE), Nanyang Technological University (NTU),
50Nanyang Drive, Singapore 637553
franke@bii.a-star.edu.sg

SEBASTIAN MAURER-STROH

*Bioinformatics Institute (BII), Agency for Science, Technology and Research (A*STAR),*
30 Biopolis Street, #07-01, Matrix, Singapore 138671
and
School of Biological Sciences (SBS), Nanyang Technological University (NTU),
60Nanyang Drive, Singapore 637551
sebastianms@bii.a-star.edu.sg

Beta lactams comprise the largest and still most effective group of antibiotics, but bacteria can gain resistance through different beta lactamases that can degrade these antibiotics. We developed a user friendly tree building web server that allows users to assign beta lactamase sequences to their respective molecular classes and subclasses. Further clinically relevant information includes if the gene is typically chromosomal or transferable through plasmids as well as listing the antibiotics which the most closely related reference sequences are known to target and cause resistance against. This web server can automatically build three phylogenetic trees: the first tree with closely related sequences from a Tachyon search against the NCBI nr database, the second tree with curated reference beta lactamase sequences, and the third tree built specifically from substrate binding pocket residues of the curated reference beta lactamase sequences. We show that the latter is better suited to recover antibiotic substrate assignments through nearest neighbor annotation transfer. The users can also choose to build a structural model for the query sequence and view the binding pocket residues of their query relative to other beta lactamases in the sequence alignment as well as in the 3D structure relative to bound antibiotics. This web server is freely available at http://blac.bii.a-star.edu.sg/.

Keywords: beta lactamase; penicillin; cephalosporin; carbapenems; drug resistance; phylogenetic tree; substrate binding pocket.

1. Introduction

Before the discovery of antibiotics, the mortality rate from bacterial infections was devastatingly high, especially for some acute and severe diseases such as the plague or cholera. In 1928, Sir Alexander Fleming discovered Penicillin and from then on, more and more other beta lactam antibiotics have been and are being discovered and manufactured, for example, different generations of cephalosporins,[1, 2, 3, 4, 5] carbapenems,[6, 7] macrolides,[8] quinolones,[9] and so on.

Till now, beta lactams are still the most effective group of antibiotics. The structure of penicillin is similar to D-alanyl-D-alanine, so it can bind at the active site of the transpeptidase that cross-links the peptidoglycan strands during the formation of bacterial cell walls. It reacts with a serine residue in the transpeptidase and irreversibly inhibits the enzyme and, hence, growth of the bacterial cell wall.[10, 11, 12, 13] But some bacteria can synthesize beta-lactamases which can hydrolyze the amide bond in the beta lactam ring and thereby inactivate beta lactams.

Despite being united by their ability to degrade beta lactams, the beta lactamases are widely diverse and belong to many different sequence and structure protein families. Based on molecular sequence classification, there are four classes of beta lactamases, class A, B, C and D.[14, 15] Class A, C, and D are also called serine-beta lactamases, whose active sites function through a serine based mechanism.[16] Class B enzymes are also called metallo beta lactamases, and their functionality depends on the binding of one or more zinc ions in the functional domain.[17] Class B is further classified into three sub-classes, i.e., B1, B2 and B3.[18, 19] Of these, B1 and B3 need to bind two zinc ions for their maximum activity while B2 just need to bind 1 zinc ion for its maximum activity and the binding of the second zinc ion will lead to its function to be inactivated.[17] The sequence similarity between B1 and B2 is higher than their sequence similarity with B3.[19]

General class designations are also related to the substrate spectrum (beta lactam resistance spectrum). For example, subclass B1 and B3 beta lactamases are generally broad spectrum resistant to beta lactams, while subclass B2 beta lactamases normally specifically hydrolyze carbapenems. In general, class C beta lactamases are resistant to beta lactamase inhibitors such as clavulanic acid, while class A beta lactamases can be inhibited by several beta lactamase inhibitors. The AmpC type beta lactamases (class C) are mainly located on chromosomes, especially those extended spectrum beta lactamases (ESBLs), but some class C beta lactamase genes were found to be located on plasmids as well which could facilitate their transfer between bacteria.[20] Carbapenemases are mainly from class A and B, and class D carbapenemases are also not rare, but increasingly, more class C carbapenemases such as CMY-18[21] are being reported.

Nowadays, the issue of drug resistance to beta lactams becomes more and more widely spread and severe,[22,23] especially the drug degradation caused by class B lactamases, because they will lead to resistance to carbapenems, which usually is the last resort of the therapy against a bacterium conveying broad spectrum drug resistance. Consequently, it is important to investigate the mechanisms, transfer, and evolution of drug resistance through beta lactamases.

With this consideration, we developed the Autotree web server for classification and clinically relevant annotation of beta lactamases. It can automatically generate three phylogenetic trees. The first tree is with its closely related sequences. The second is with a large set of manually curated reference or seed beta lactamase sequences many of which are annotated with clinically important properties such as their substrate drug specificity and their chromosomal/plasmid localization. Sometimes, the evolutionary rate of the active site region differs from the evolutionary rate of the whole sequence. So, a third tree is also built by using the amino acid residues which are adjacent to the substrate binding pocket. When the query sequence is closer to a seed sequence on the substrate binding pocket tree, this means that the substrate binding pocket of the query sequence is similar to that of the seed beta lactamase, and in consequence, the annotation of the respective substrate specificity can be transferred from closely related homologues (i,e., the beta lactams that can be hydrolyzed). Besides these three phylogenetics trees, the users can also choose to build a structural model for the query sequence to gain further insight into the structural and catalytic characteristics of the putative beta lactamase.

2. Materials and Methods

2.1. *Acquiring closely related sequences*

The closely related sequences of the query sequence were determined and retrieved by using the Tachyon service against the NCBI nr database.[24] Tachyon is a peptide-indexing based search method that identifies similar sequences of a query protein at a very fast speed achieved by reducing the searched sequence space through associating each entry in the database only with a small number of representative pentapeptides. Database hits above a specified number of shared pentapeptides with the query are then subjected to a more detailed search over the full length sequences to evaluate the significance of each hit.

2.2. *Manually curated reference sequences of different classes*

We searched the NCBI protein database by using the key words of lactamase, penicillinase, cephalosporinase, carbapenemase, and so on. We then wrote a script to remove all the non-related sequences from the above query result, and the remaining sequences were designated to their corresponding classes according to the annotation information available. For class B, except for these sequences retrieved from NCBI, we also added those sequences from the literature of Hall et al,[19] which have clear subclass designation. After this, CD-HIT [25] was utilized to remove 90 percent of redundancy.

For each sequence, we checked its corresponding literature references to find corresponding information such as coding gene location (chromosome or plasmid), organism, and substrates, if available. The substrate information is according to the original Minimum Inhibitory Concentrations (MIC)[26] test. If available, such information was added to the description lines of the class seed files.

2.3. *Building of tree_tachyon and tree_clustering*

The query sequence is first aligned against the closely related Tachyon hits by using the sequence alignment toolkit of MAFFT[27] with the L-INS-I option (--localpair --maxiterate 1000), and then a maximum likelihood tree with these closely related sequences (tree_tachyon) is built using phyML[28].

For alignment and tree with our curated seed sequences, the query sequence is first searched against sequences of all classes with blast2seq[29] and a conservative E-value threshold of 1e-5 to ensure the query is not an unrelated sequence that may not be a beta lactamase. If it has at least one significant hit, the query is aligned against the four curated reference sequence files using MAFFT[27] with the L-INSI-I option, then the sequence identity with each reference sequence is calculated and an average identity value with each class seed file is acquired, and the class with highest average identity value will be assigned as the class of the query sequence. If a class is successfully identified, the query sequence is aligned against the class reference sequences by using MAFFT[27] with the option of L-INS-I, and a maximum likelihood tree with these seed sequences (tree_clustering) is built using the toolkit of phyML.[28]

2.4. *Building of tree_ligand*

Among the reference seed sequences, there are some which have corresponding experimentally determined crystal structures with substrate or substrate analogue bound. For class A, it is 3BFC;[30] for class B, they are 1JJE, 2YZ3, and 2GKL;[31,32,33] for class C, it is 1LL5;[34] for class D, it is 1H5X.[35] For each class reference alignment, the amino acid residues being within 8 angstroms' distance to the substrate or substrate analogue ligand were determined. All other positions are masked with X and the remaining alignment positions representing the substrate binding pocket residues are used to build a maximum likelihood phylogenetic tree (tree_ligand) accordingly.

2.5. *Building of quick tree*

If the users choose the option “quick tree”, an UPGMA-based guide tree built by MAFFT during the multiple sequence alignment is taken as rough but very quickly built tree. Sometimes this quick tree is useful to get a brief look at the general relationship among the component sequences.

2.6. *Building of structural model*

If the users choose the option of structural model building, a structural model will be built by using the toolkit of MODELLER [36] without the procedure of loop refinement and with the procedure of ligand attachment by using the templates and chain codes designated by the users. If the users do not provide corresponding templates, then the six best PDB hits (if available) of Tachyon[24] will be adopted as the templates for structural modeling. If no such hits were discovered, this procedure will be omitted.

According to the alignment of the query sequence with the corresponding reference seed sequences (alignment_clustering), amino acid conservation values will be calculated

by using the toolkit of SPECS,[37] and the conservation values will be mapped into the B-factor column of the structural model.

3. The Web Server

3.1. *Input*

A compulsory input will be a nucleotide or amino acid sequence of a putative beta lactamase. The users can use the text input area to paste a fasta format query sequence, or choose to upload a text file containing a fasta format query sequence.

3.2. *Options*

There are two checkboxes on the start page, corresponding to two options of the analysis. One is 'structural modeling' and another is 'quick tree'. If the former one is chosen, the process of structural modeling will be performed, and if the latter is chosen, a quick tree instead of the maximum likelihood tree will be built.

3.3. *Workflow*

The general workflow of the web server is demonstrated in Fig. 1.

1. Input check: after the job was submitted, the web server will check whether the input sequence was nucleotide or amino acid sequence.

If the input sequence is a nucleotide sequence, then all the translated amino acid sequences of this sequence and its complement sequence in the frame of N, N+1, and N+2 will be determined. And the longest one will be chosen as the amino acid query sequence for further analysis.

2. The query sequence will be aligned against each seed sequences by using blast2seq[29] to judge whether the query sequence is a beta lactamase or not. If the expectation value does not reach the threshold (1e-5), the query sequence will be considered as not a beta lactamase sequence or too far away from the reference sequences, and an error reporting page will be generated containing the best e-value information of the query sequence with each seed class.

3. Tree type determination: if the users choose to build a quick tree, then the guide tree from MAFFT[27] will be output as the quick result tree. If not, then a maximum likelihood tree will be built by using phyML.[28]

4. Tree_tachyon building: the closely related sequences of the query will be determined and retrieved by using the Tachyon service against the NCBI nr database,[24] the sequences will be aligned, and tree_tachyon will be built.

5. Reference class determination: the query sequence will be aligned against each of the four class seed files, and the one with the maximum average identity value will be determined as the seed file.

6. Tree_clustering building: the alignment of the query sequence against the reference seed sequences will be created and tree_clustering will be built.

7. Structural model building: If the users choose to build a structural model, then the web server will check whether corresponding templates are provided as well. If not, PDB hits of Tachyon will be taken as templates for structural modeling. If Tachyon returns no PDB hits, then this procedure will be omitted.

8. Amino acid conservation value calculation: If structural modeling was performed, then the amino acid conservation values of the query sequence will be calculated according to the alignment_clustering. Then, these values will be mapped to the B-factor column of the built structural model.

9. Tree_ligand building: Amino acid residues that are adjacent to the substrate binding pocket of the representative PDB structures of the reference class will be determined, and positions of these residues inalignment_clustering will be located. All other positions are masked with X (alignment_ligand) and the remaining alignment positions representing the substrate binding pocket residues are used to build a maximum likelihood phylogenetic tree (tree_ligand) accordingly.

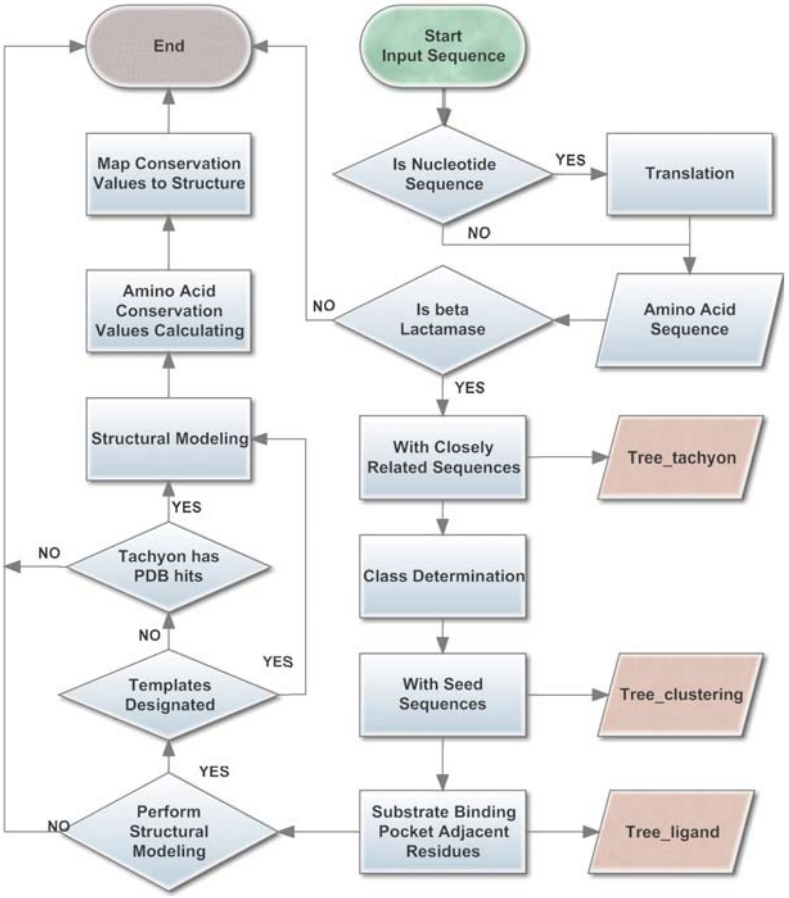

Fig. 1. Flowchart of the web server.

3.4. *Output*

The result tree files (newick format) will be converted to phyloXML format[38] by using the java class of phyloxml_converter in the toolkit of Archaeopteryx.[39] Annotation information such as gene location (chromosome or plasmid), organism, and substrate will be added to the phyloXM format tree file, and then, this phyloXML file will be revised to highlight the branch of the query sequence with red color.

We also wrote two new java classes of exportToPdf and exportToPng to use the phyloXML format tree file as the input and convert the corresponding tree to a publication quality png or pdf format file.

These phyloXML trees are displayed by using the applet version of Archaeopteryx.[39] The users can choose from a drop down menu which tree should be displayed (tree_tachyon, tree_clustering, and tree_ligand). Fig. 2 shows the display of the phylogenetic tree of an example query sequence with its respective identified class of reference sequences.

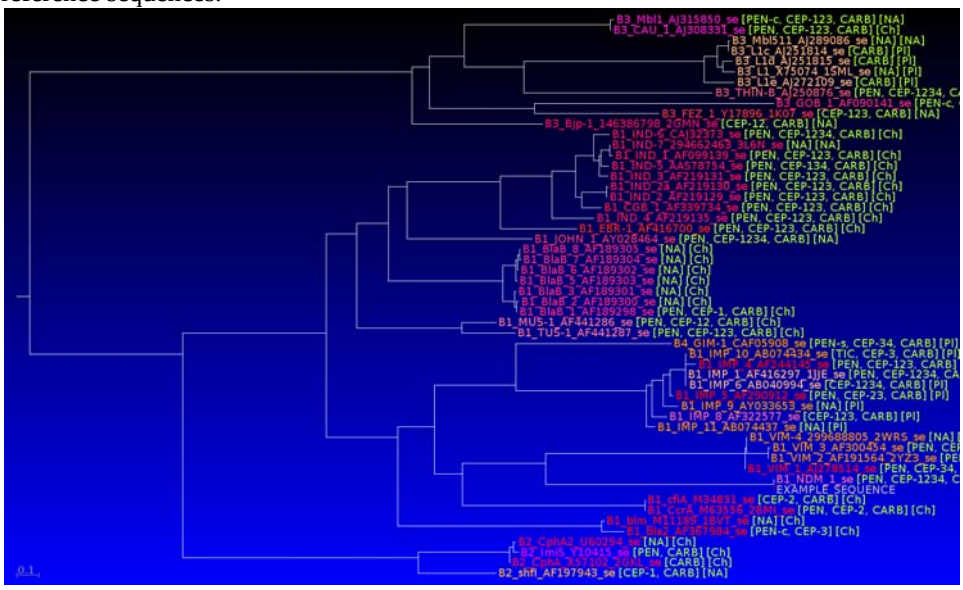

Fig. 2. Display of the phylogenetic tree of an example query sequence (NDM1 in this case) built with the curated class B sequences by using Archaeopteryx.

On the phylogenetic tree, to save display space, we use abbreviation codes, for an example, we use PEN-c to stand for aminopenicillins and carboxypenicillins, which are prone to resistance by the bacteria, and use PEN-s to stand for ureidopenicillins, which are less prone to resistance by the bacteria, and use PEN to stand for aminopenicillins, carboxypenicillins, and ureidopenicillins together. Same applies to the situation of CEP-x,

which stands for some generations of cephalosporins or all. The abbreviation codes used in the phylogenetic tree are listed in Table 1.

Table 1. Some abbreviation code of the substrate (beta lactams) of beta lactamase

Abbreviation code	Full name or meanings	Abbreviation code	Full name or meanings
PEN	Penicillins	CARB	Carbapenems
PEN-c	aminopenicillins and carboxypenicillins,	ATM	Aztreonam
PEN-s	ureidopenicillins	MBM	other monobactams
CEP	Cephalosporins	TZB	beta lactamase inhibitor tazobactam
CEP-i	cephalosporin generation i	NA	not available

The corresponding alignment files (alignment_tachyon, alignment_clustering, and alignment_ligand) are displayed by using the applet version of Jalview.[40]

The structural model built is displayed by using the applet version of Jmol.[41] For this purpose, several Jmol functions were written to fulfill specific functions, such as highlighting conserved amino acid residues above certain threshold values, highlighting amino acid residues close to the substrate binding pocket in the structural model (if available, cyan if between 3 and 5 angstroms, and magenta if within 3 angstroms).Fig. 3 shows the display of the structural model of the example sequence built by using the applet version of Jmol. In Fig. 3, amino acid residues within 3 angstroms to the ligands are colored magenta, and residues between 3 to 5 angstroms to the ligands are colored cyan.

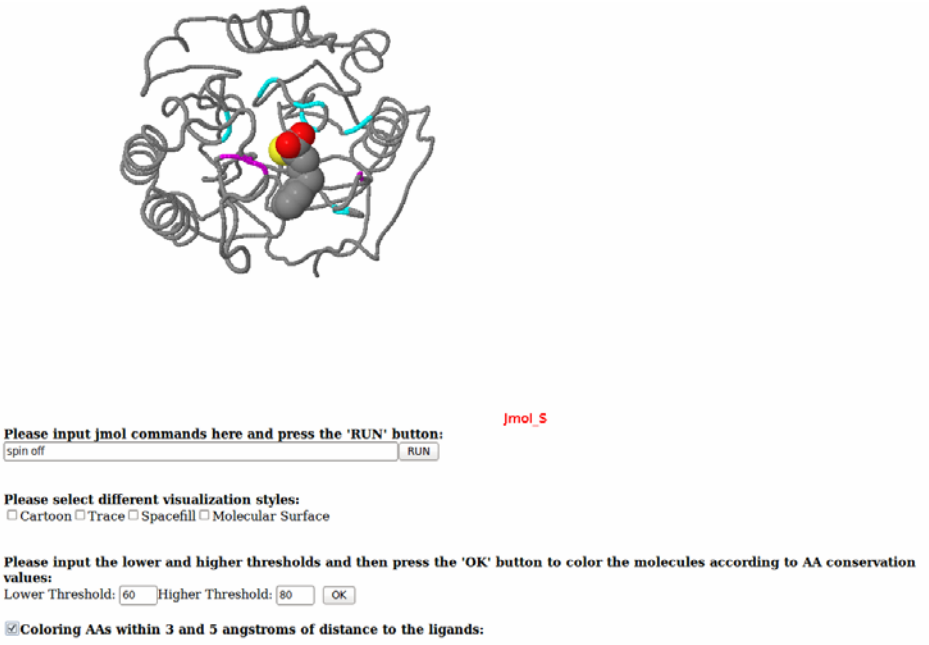

Fig. 3. Visualization of the structural model of the example sequence by using Jmol including typical structure manipulation options.

There is also a download link on the output page of a tar format compressed file containing the alignment files (alignment_tachyon, alignment_clustering, and alignment_ligand), newick format of phylogenetic tree files (tree_newick_tachyon, tree_newick_clustering, and tree_newick_ligand), phyloXML format of tree files (tree_xml_tachyon, tree_xml_clustering, and tree_xml_ligand), pdf format of the tree files (tree_xml_pdf _tachyon, tree_xml_pdf _clustering, and tree_xml_pdf _ligand), and pdb file of the structural model if available. The users can choose to download these files to their local computer for further local manipulation later.

The running time is about 1 minute with 'quick tree option' and 20 minutes or more otherwise, depending on whether a structural model is to be built or not. During this time, the result page will be refreshed in a regular interval. Whenever the analysis process is finished, the result will be displayed accordingly. The users can also copy the result page address and visit it later.

3.5. *Application purpose*

This web server has been developed following feedback of the needs of infectious disease surveillance labs and clinicians in Singapore which should be representative for most developed countries with active surveillance programs. While the completely automated selection of the best suited drugs for patients using computational predictions alone is still not reliable enough, a server as presented here nevertheless has great value as it overtakes several necessary analysis steps otherwise requiring advanced bioinformatics knowledge and taking considerable amount of time. Simpler tasks such as sequence similarity search and multiple alignments are complemented with advanced phylogenetic analysis and structural modeling and the non-trivial combination of the latter two. At the same time, with manually curated reference families for the different beta lactamase classes with substrate annotation and risk of transferability of resistance (being on plasmid etc.) we aim to create a unique novel resource for clinicians and surveillance labs requiring sequence analysis of beta lactamase sequences.

4. Benchmark of the performance of the server

4.1. *Performance of the classification of beta lactamase and non-beta lactamase sequences*

To check the ability of the server to correctly assign class labels to potential beta lactamase sequences and correctly recognize sequences not related to beta lactamases, we identified suitable positive and negative sets for performance testing. Given the exhaustive database and literature curation effort described above, our seed sequences represent the current best set of known and highly likely beta lactamases and were hence adopted as the positive dataset (altogether 215 sequences). Sequences of the SCOP ASTRAL subset with known 3D structures but unrelated to beta lactamase folds were

adopted as the negative dataset.[42,43] In detail, after the Astral SCOP 1.75b nr40 sequences were fetched from the website of SCOP, sequences belonging to the 'Metallo-hydrolase/oxidoreductase' super-family and 'beta-lactamase/transpeptidase-like' super-family were removed, leaving 11152 sequences. The latter includes sequences of class A, C and D beta lactamases, and the former includes sequences of class B beta lactamase which comprise a different structural fold compared to the other classes.

For the positive data set, when a sequence is blasted against the seed file of its own class, this sequence will be removed from the corresponding seed file and only the remaining sequences will be used to determine whether this sequence should be classified as a member of this class or not following the style of a leave-one-out or jack knife cross validation. All 215 beta lactamase sequences were correctly assigned to the corresponding classes, and no false negative cases occurred. For the negative data set, also all 11152 sequences were correctly labeled as not being beta lactamases, and no false positive cases occurred. The number of cases of true positive, false negative, true negative, and false positive were listed in Table 2, and the values of accuracy, sensitivity, and specificity were all 100% respectively. This literally perfect performance indicates that more sophisticated methods such as HMM- or machine learning-based approaches are not needed for this specific classification task.

Table 2. Accuracy, sensitivity and specificity of the server

True positive (TP)	False negative (FN)	True negative (TN)	False positive (FP)	Accuracy (%)	Sensitivity (%)	Specificity (%)
205	0	11152	0	100	100	100

4.2. *Antibiotic Substrate Recovery through Nearest Neighbor Annotation Transfer*

While the assignment of beta lactamase sequences into their correct classes appears trivial, assigning the correct substrate range of a beta lactamase is not. Within one class, even remote relatives in the phylogenetic tree can share at least some substrates. However, since substrate recognition depends on features of the binding pocket, we hypothesize that a phylogenetic tree built only using substrate binding pocket residues (tree_ligand) could be more accurate than the whole sequence tree (tree_clustering) in the ability to recover substrate information through neighbor annotation transfer. To test this hypothesis, we calculated the number of shared substrates between the nearest nodes in our annotated reference alignments.

The nearest node pairs are determined through the comparison of distances according to branch lengths information of the newick tree and applying a distance cutoff threshold <0.3 to exclude long branch pairs. The shared substrates between nearest neighbor nodes of the whole sequence tree and the substrate binding pocket tree were calculated respectively for the well annotated class B beta lactamases, and the result is shown in Table 3.

Table 3. Comparison of substrate recovery through nearest neighbors

Tree type	Whole sequence tree	Substrate binding pocket tree
Node pairs both with substrate information	17	21
Shared substrates (absolute number)	76	95

As can be seen, there seems to be an advantage for the substrate binding pocket tree. At the same time, it is difficult to bring this comparison into more classical performance parameters. One possibility is to consider identical substrate range matches among nearest pairs as true positives and those for farthest pairs in the respective tree as false positives. This allows calculating sensitivity, specificity and accuracy as shown in Table 4. In all these parameters, the substrate binding pocket tree performs better than the classical whole sequence tree. It should be noted that this performance measurement is not ideal as also the most remote hits within a class can still sensibly share some substrates. Also, performance of any nearest neighbor transfer method also strongly depends on availability of substrate range annotation etc.

Table 4. Comparison of substrate recovery of nearest neighbor node and most remote nodes

	TP	FN	FP	TN	Accuracy (%)	Sensitivity (%)	Specificity (%)
Whole sequence tree	15	2	20	19	60.7	88.2	48.7
Substrate binding pocket tree	19	2	11	19	74.5	90.5	63.3

5. Conclusions

The AutoTree web server for beta lactamases allows determining whether a query sequence is a beta lactamase, and if so, which class or subclass it belongs to, what are its possible antibiotic substrates, and whether the resistance ability it stands for can be transferred easily as plasmid.

Besides an optional structural model, a tree with closest Tachyon hits against the NCBI nr database, a tree with manually curated full length reference sequences and a tree built from substrate binding pocket residues are automatically built. The latter is more suitable to deduce the possible antibiotic resistance profile of a putative beta lactamase through nearest neighbor annotation transfer.

This web server provides a free one-stop Bioinformatics solution for clinicians and surveillance labs, overtaking otherwise time consuming manual steps and providing state-of-the-art computational analysis of sequences including structure and substrate prediction from the diverse functional and structural families of protein beta lactamases.

Acknowledgments

We acknowledge support by A*STAR and the IMaGIN platform grant.

References

1. Boniece WS, Wick WE, Holmes DH, Redman CE, In vitro and in vivo laboratory evaluation of cephalothin, a new broad spectrum antibiotic, J Bacteriol 84: 1292-1296, 1962.
2. Kosmidis J, Hamilton-Miller JM, Gilchrist JN, Kerry DW, and Brumfitt W, Cefoxitin, a new semi-synthetic cephamycin: an in-vitro and in-vivo comparison with cephalothin, Br Med J 4(5893): 653-655, 1973.
3. Kamimura T, Kojo H, Matsumoto Y, Mine Y, Goto S, Kuwahara S, In vitro and in vivo antibacterial properties of FK 027, a new orally active cephem antibiotic, Antimicrob Agents Chemother 25(1): 98-104, 1984.
4. Bodey GP, Ho DH, LeBlanc B, In vitro studies of BMY-28142, a new broad-spectrum cephalosporin, Antimicrob Agents Chemother 27(2): 265-269, 1985.
5. Yancey RJ Jr, Kinney ML, Roberts BJ, Goodenough KR, Hamel JC, Ford CW, Ceftiofur sodium, a broad-spectrum cephalosporin: evaluation in vitro and in vivo in mice, Am J Vet Res 48(7): 1050-1053, 1987.
6. Kesado T, Hashizume T, Asahi Y, Antibacterial activities of a new stabilized thienamycin, N-formimidoylthienamycin, in comparison with other antibiotics, Antimicrob Agents Chemother, 17(6): 912-917, 1980.
7. Edwards JR, Turner PJ, Wannop C, Withnell ES, Grindey AJ, Nairn K, In vitro antibacterial activity of SM-7338, a carbapenem antibiotic with stability to dehydropeptidase I, Antimicrob Agents Chemother, 33(2): 215-222, 1989.
8. HAIGHT TH, FINLAND M, Laboratory and clinical studies on erythromycin, N Engl J Med 14; 247(7): 227-232, 1952.
9. Rohlfing SR, Gerster JR, KvamDC, Bioevaluation of the antibacterial flumequine for urinary tract use, Antimicrob Agents Chemother 10(1): 20-24, 1976.
10. FEW AV, COOPER PD, ROWLEY D, Reaction of penicillin with the staphylococcal cell wall, Nature 169(4294): 283-284, 1952.
11. Izaki K, Matsuhashi M, Strominger JL, Glycopeptidetranspeptidase and D-alanine carboxypeptidase: penicillin-sensitive enzymatic reactions, ProcNatlAcadSci U S A 55(3): 656-663, 1966.
12. Highton PJ, Hobbs DG, Penicillin and cell wall synthesis: a study of Bacillus licheniformis by electron microscopy, J Bacteriol 106(2): 646-658, 1971.
13. Lee B, Conformation of penicillin as a transition-state analog of the substrate of peptidoglycan transpeptidase, J MolBiol 61(2): 463-469, 1971.
14. Ambler RP, The structure of beta-lactamases, Philos Trans R SocLond B BiolSci 289(1036): 321-231, 1980.
15. Hall BG, Barlow M, Revised Ambler classification of {beta}-lactamases, J AntimicrobChemother 55(6): 1050-1105, 2005.
16. Ghuysen JM, Bacterial active-site serine penicillin-interactive proteins and domains: mechanism, structure, and evolution, Rev Infect Dis 10(4): 726-732, 1988.
17. Wang Z, Fast W, Valentine AM, Benkovic SJ, Metallo-beta-lactamase: structure and mechanism, CurrOpinChemBiol 3(5): 614-622, 1999.
18. Rasmussen BA, Bush K, Carbapenem-hydrolyzing beta-lactamases, Antimicrob Agents Chemother 41(2): 223-232, 1997.
19. Hall BG, Salipante SJ, Barlow M, The metallo-beta-lactamases fall into two distinct phylogenetic groups, J MolEvol 57(3): 249-254, 2003.
20. Philippon A, Arlet G, Jacoby GA, Plasmid-determined AmpC-type beta-lactamases, Antimicrob Agents Chemother 46(1): 1-11, 2002.
21. Lee K, Lee M, Shin JH, Lee MH, Kang SH, Park AJ, Yong D, Chong Y, Prevalence of plasmid-mediated AmpC beta-lactamases in Escherichia coli and Klebsiella pneumoniae in Korea, Microb Drug Resist 12(1): 44-49, 2006.

22. Livermore DM, Bacterial resistance to carbapenems, AdvExp Med Biol 390: 25-47, 1995.
23. Bonnin RA, Cuzon G, Poirel L, Nordmann P, Multidrug-Resistant Acinetobacterbaumannii Clone, France, Emerg Infect Dis 19(5): 822-823, 2013.
24. Tan J, Kuchibhatla D, SirotaFL, ShermanWA, Gattermayer T, Kwoh CY, Eisenhaber F, Schneider G, Maurer-Stroh S, Tachyon search speeds up retrieval of similar sequences by several orders of magnitude, Bioinformatics 28(12): 1645-1646, 2012.
25. Li W, Godzik A, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, Bioinformatics 22(13): 1658-1659, 2006.
26. Andrews JM, Determination of minimum inhibitory concentrations, J AntimicrobChemother 48(Suppl 1): 5-16, 2001.
27. Katoh K, Misawa K, Kuma K, Miyata T, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, Nucleic Acids Res 30(14): 3059-3066, 2002.
28. Guindon S, Gascuel O, A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood,SystBiol 52(5): 696-704, 2003.
29. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, Basic local alignment search tool, J Mol Biol 215(3): 403-410, 1990.
30. http://www.rcsb.org/pdb/explore/explore.do?structureId=3BFC
31. Toney JH, Hammond GG, Fitzgerald PM, Sharma N, Balkovec JM, Rouen GP, Olson SH, Hammond ML, Greenlee ML, Gao YD, Succinic acids as potent inhibitors of plasmid-borne IMP-1 metallo-beta-lactamase, J BiolChem 276(34): 31913-31918, 2001.
32. Yamaguchi Y, Jin W, Matsunaga K, Ikemizu S, Yamagata Y, Wachino J, Shibata N, Arakawa Y, Kurosaki H, Crystallographic investigation of the inhibition mode of a VIM-2 metallo-beta-lactamase from Pseudomonas aeruginosa by a mercaptocarboxylate inhibitor, J Med Chem 50(26): 6647-6653, 2007.
33. Horsfall LE, Garau G, Liénard BM, Dideberg O, Schofield CJ, Frère JM, Galleni M, Competitive inhibitors of the CphAmetallo-beta-lactamase from Aeromonashydrophila, Antimicrob Agents Chemother 51(6): 2136-2142, 2007.
34. Beadle BM, Shoichet BK, Structural basis for imipenem inhibition of class C beta-lactamases, Antimicrob Agents Chemother 46(12): 3978-3980, 2002.
35. http://www.rcsb.org/pdb/explore/explore.do?structureId=1H5X
36. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A, Comparative protein structure modeling using Modeller, CurrProtoc Bioinformatics Chapter 5: Unit 5.6, 2006.
37. Mihalek I, Res I, Lichtarge O, A family of evolution-entropy hybrid methods for ranking protein residues by importance, J MolBiol, 336(5): 1265-1282, 2004.
38. Han MV, Zmasek CM, phyloXML: XML for evolutionary biology and comparative genomics, BMC Bioinformatics, 10: 356, 2009.
39. Zmasek CM, Eddy SR, ATV: display and manipulation of annotated phylogenetic trees, Bioinformatics 17(4): 383-384, 2001.
40. Clamp M, Cuff J, Searle SM, Barton GJ, The Jalview Java alignment editor, Bioinformatics 20(3): 426-427, 2004.
41. Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org/
42. Murzin AG, Brenner SE, Hubbard TJP, Chothia C, SCOP: a structural classification of proteins database for the investigation of sequences and structures, J Mol Biol 247, 536-540, 1995.
43. Chandonia JM, Hon G, Walker NS, Lo Conte L, Koehl P, Levitt M, Brenner SE, The ASTRAL compendium in 2004, Nucleic Acids Research 32: D189-D192, 2004.