An hierarchical classification system for beta-lactamases

Melise Chaves Silveira[1], Fábio Mota[1], Rodrigo Jardim[1], Rangeline Azevedo daSilva[1], Marcos Paulo Catanho de Souza[2], Ana Carolina Ramos Guimarães[2], Antônio Basílio de Miranda[1]

Laboratório de Biologia Computacional e Sistemas[1]; Laboratório de Genômica Funcional e Bioinformática[2]

NGS sequencing projects are revealing a growing number of unique and naturally occurring beta-lactamases, enzymes able to irreversibly inactivate beta-lactams antibiotics, the major option to treat bacterial infections. This diversity and the major clinical impact of beta-lactamases led to several attempts to achieve a representative classification system. Ambler's structural classification (1980) is the most used, but it does not represent the evolutionary relationships between these enzymes, the reason Hall and Barlow suggested a hierarchical organization of it (2005). In this work, we propose a system to identify and hierarchically classify beta-lactamases, considering structure and sequence characteristics, prioritizing confidence when attributing function and class for a given protein. Beta-lactamases (EC 3.5.2.6) primary and tertiary structures were downloaded from PDB. Structural and sequence hierarchical clustering tests using MaxCluster and BLASTCLUST programs were performed to achieve Ambler's classification as reviewed by Hall and Barlow. According to it, serine beta-lactamase classes A, C and D were renamed to SA, SC and SD, and the metallo beta-lactamases were divided in MB and ME, and in the next level MB is divided in subclasses B1 and B2. After achieving the five initial clusters, we constructed Hidden Markov Models profiles for each one using the HMMER package and the protocol HMM-ModE. The profiles were tested for class specificity (CE), beta-lactamase function specificity (FE) and beta-lactamase function sensibility (FS) through in house scripts, using a beta-lactamase dataset constructed by us, the CATH database and the Swiss-Prot database, respectively. Single linkage hierarchical clustering using beta-lactamase structures formed the five clusters expected; sequence clustering under a 50% similarity threshold separated groups B1 from B2; and using a 60% similarity threshold a novel level is suggested, where classes SA, SD, B1 and ME were each divided in two, corroborating previous studies by other groups. All profiles obtained 100% CE. Together, the profiles have 87% FS using the Gene Ontology/Swiss-Prot annotation. We noticed that the remaining 13% is actually comprised by proteins without beta-lactamase activity. MB and ME profiles displayed 100% FE, as either SA, SC and SD after to apply phylogenetic methods to determine new score thresholds and establishing “gray zones” for searches. With these procedures we have achieved a system for the identification and classification of beta-lactamases exclusively on the basis of different levels of protein structures, that also reflects their evolutionary relationships. The system was curated and proved to be efficient, able to be applied in large scale.

Funding support: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)