Poster M-6

Identification of Microorganisms Clusters by Using Dynamic Programming to Analyze Antimicrobial Susceptibility Sequences

Authors:
Braulio RGM Couto (*UNI-BH*)
Ana P Ladeira (*UNI-BH*)
Carlos EF Starling (*Hospitais Vera Cruz , Life Center e Baleia*)

Short Abstract: Pathogens have their antibiogram inserted in strings of S, R, or U (sensible, resistant or unknown) and analyzed by a dynamic programming algorithm that identifies clusters. Despite the fact that antimicrobial resistance is a phenotype characteristic, the method was capable to identify clusters in absence of automated ribotyping systems.

Long Abstract:
Introduction: automated ribotyping systems and pulsed-field gel eletrophoresis are methods for routine DNA fingerprinting that are not readily available at many institutions in developing countries. In an epidemic situation, these kinds of methods could be replaced by analyzing the similarity of antimicrobial susceptibility of the strains involved with the outbreak. Objective: to use dynamic programming algorithm (the same used to perform DNA sequence alignment) in order to compare antimicrobial susceptibility sequences and to identify possible clusters. Methods: firstly, selected pathogens have their antibiogram for a given number of different antimicrobials inserted in a string of S, R, or U, respectively for sensible, resistant or unknown (not tested) to each drug tested. For example, considering 11 antimicrobials (oxacillin, gentamicin, trimethoprim/sulfamethoxazole, ampicillin, penicillin, cephalothin, ceftriaxone, ciprofloxacin, tetraciclin, eritomicin, clindamycin) an strain-01 could be placed as RRSURSSRSSR, and strain-02 as RRRSSRRSSSR. Secondly, each string or ATB sequence is subject to a pairwise alignments calculated by dynamic programming: each pair of sequence to be compared (S1 and S2) is putted on the two axes of a table and then progressively, the table is filled from upper left to the lower right with integer scores related to the pairwise similarity of each ATB result in the two sequences. The integer scores used are related to the atomic edit functions insertion (score=1), deletion (score=1), substitution (score=1) and match (score=0). The dynamic programming algorithm calculates the minimum edit distance to transform an ATB string S1 in another string S2. An edit distance equal to zero means that both strings are identical. After to calculate a matrix of all edit distance, another matrix of pairwise similarity (0 = dissimilar and 1 = similar) is constructed. Two similar strains are defined according to a cutoff in the edit distance matrix, for example, equal to 0 (too restrictive) or less or equal to 1. Finally, after to calculate the matrix with all pairwise similarity comparisons, all clusters of similar strains can be identified. The method used to identify clusters is a brute-force algorithm that performs all comparisons in the similarity matrix, grouping sequences that are directly or indirectly related. For example, if sequence 01 is similar to 03 and sequence 03 is similar to 05, so the cluster has three items: sequences 01, 03 and 05. Results: we tested the method in two types of infection. Case A involves 14 strains of Methicillin-resistant coagulase-negative staphylococci, of nosocomial

infections diagnosed between April 2003 and September 2004 at a high risk nursery. The antimicrobial susceptibility pattern of each strain was changed in to an ATB string considering 11 antimicrobials: oxacillin, gentamicin, trimethoprim/sulfamethoxazole, ampicillin, penicillin, cephalothin, ceftriaxone, ciprofloxacin, tetraciclin, eritomicin, clindamycin. After to calculate the edit distance matrix and similarity matrix by using one as the cutoff in distance matrix, we identified 8 different clusters: strains 3, 4, 7, 8, 13, and 14 belong to a unique cluster; strains 9 and 10 are in another cluster and strains 1, 2, 5, 6, 11 and 12 are completely separately from every other. Case B involves 55 strains of Pseudmonas aeruginosa resistant to Imipenem, diagnosed in nosocomial infection from January/2003 and April/2005. The antimicrobial susceptibility pattern of each strain considered 11 antimicrobials: imipenem, meropenem, cefepime, ciprofloxacin, amicacin, amoxicillin, ampicillin, aztreonam, ceftazidime, ceftriaxone and gentamicin. Four main clusters were found by applying a cutoff = 1 in the edit distance (the first cluster has 31 strains). Conclusion: despite the fact that antimicrobial resistance is a phenotype characteristic, the method was capable to identify clusters of microorganisms in absence of automated ribotyping systems. With the use of MIC measures, instead of S/R dichotomy, the method could be improved. This method must be validated confronting its results with DNA fingerprints data, which is the actual gold standard. The method and algorithms are implemented in a MATLAB® (The Mathworks Inc.) program that is available by request (bcouto@acad.unibh.br) or for free downloading (www.cc.unibh.br).