

Aplikasi Quaternion dalam pengkodean DNA dan penentuan skor pada pensejajaran barisan DNA = Quaternion application in DNA encoding and DNA sequence alignment scoring

Teti Novianti, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20330231&lokasi=lokal>

Abstrak

Barisan DNA dapat diartikan sebagai permutasi dari empat kode basa DNA yaitu A, T, G, dan C. Pada hasil sekuensing DNA, kadang kala ada basa DNA yang sulit terbaca dengan jelas apakah A, T, G, atau C. Untuk mengatasi masalah ini, maka diberikan kode-kode lain yang merupakan probabilitas munculnya A, T, G, atau C pada setiap kode basa DNA. Sehingga secara keseluruhan terdapat kode basa DNA yang dapat dibentuk dari kode A, T, G, dan C. Enam belas kode basa DNA yang telah terbentuk dapat dinyatakan dalam quaternion berbentuk. Dengan menggunakan perkalian titik antara keenambelas kode basa DNA tersebut, diperoleh matriks skoring. Matriks skoring dibutuhkan pada pensejajaran barisan DNA untuk memberikan skor kecocokan atau ketidakcocokan antara dua kode basa DNA.

Algoritma pensejajaran barisan DNA yang diimplementasikan pada penelitian ini adalah Algoritma Needleman-Wunsch untuk pensejajaran global dan Algoritma Smith-Waterman untuk pensejajaran lokal, algoritma ini diimplementasikan menggunakan bahasa pemrograman berbasis open source (Octave). Kemudian program pensejajaran yang telah dibuat diaplikasikan untuk mensejajarkan barisan DNA dari bakteri *Streptococcus pneumoniae* yang diambil dari pangkalan data gen (GeneBank) dengan barisan DNA hasil sekuensing dari bakteri yang diduga *Streptococcus pneumoniae*. Dari hasil pensejajaran, diketahui bahwa kedua barisan mempunyai kemiripan yang maksimal.

DNA sequence can be defined as a permutation of four DNA base codes: A, T, G, and C. From the result of DNA sequencing, sometime there were difficulties to determine whether the DNA base code of A, T, G, or C. the Probability of other DNA base codes were given to solve this umbigue of DNA sequence reading with form of DNA base code which can be obtained from base code of A, T, G, and C. Sixteen of DNA base code can be represented with a quaternion form: . The scoring matrix was obtained from those sixteen of DNA base code using a dot product method. This scoring matrix can be applied in the DNA alignment for match and mismatch between two DNA base code.

In this study, we applied the Needleman-Wunsch Algorithm for gobal alignment and Smith-Waterman Algorithm for local alignment using the Octave, an open soucre program. The alignment program used for DNA sequences alignment from *Streptococcus pneumoniae* obtained from the GeneBank and the DNA sequence obtained from sequencing result and suspected as *Streptococcus pneumoniae*. From this alignment, we found that two DNA sequences have maximum similarity.