Editorial Open Acces

With the development and completion of genome projects, the genome data not only are available on the search of primary databases but also the structural and functional prediction powers are increased on the growth of secondary databases. Although, data retrieval tools suggested by the database updating centers are relatively simple for use of cellular and molecular investigators, but many prediction tools are based on understanding algorithms. is is the evidence, that the data management is related to mathematics and computer sciences in bioinformatics. ere is a gap on the interest of biologists and biochemists to the use of prediction tools, since a little information is observed on designing algorithms. Computational biology is a scope more considered by investigators; however, the mathematics schemes are also used in their owcharts.

It is obvious that the progressions in bioinformatics are rate-limited; it depends on the application of new matrices in algorithms and elevation of mathematical instruments. e basic question is that, “what should molecular and cellular investigators know about algorithms and their updated versions?”

On the excellent idea to know the details of bioinformatics algorithms, biologists and biochemists should spend the long and not-gracious educational terms on mathematical aspects of computer and information sciences. In addition, they should refresh your information on update of tools.

I thought these are the most important causes of investigators’ unwilling to apply bioinformatics tools in their studies. My opinion is that a biologist or a biochemist should know only common descriptions and updates about algorithms designed for tools, in di erent branches of bioinformatics including sequence homology, genomics, transcriptomics and structural modeling.

Algorithm

Bioinformatics algorithms are step-by-step procedures for predicting an event. ere are a set of mathematical and biological rules that support a prediction on calculating, processing and reasoning. e principles of each algorithm are commonly based on the two models; homology and ab initio. e models are updated on the development of mathematical schemes, whose details are not essential for a biologist to know. Although, the precision of updates can be considered on checking the outputs with quality control factors, but the link of tools within layers of a neural network point out that there are needs to know the descriptive characteristics of updated versions. e most important changes are primarily considered in the scoring matrices of structural and functional units (amino acid, base, rotamer), and then, the programs that try to relate each unit to its anking sections.

Matrices

e primary scoring matrices are obtained on the molecular distance/similarity and conservation. Other so matrices such as Position-Speci c Scoring Matrix (PSSM) are also created and corrected on the primary data and score weights [1].

Distance/Similarity

DNA matrices were not complicated on the base substitutions; however, there were constitutional tendencies on the transition and transversion changes. PAM (Point Accepted Mutation) amino acid matrix [2], developed by Margaret Dayho was based on the evolutionary model, while BLOSUM matrix (BLOcks of Amino Acid SUbstitution Matrix), developed by Heniko and Heniko [3] was created on the BLOCK identity.

Sequence conservation

e scores to predict the secondary structures were created by Chou-Fasman, namely the Chou-Fasman method [4], and were developed by Gor method [5]. Furthermore, some scores were suggested on

GHSHQGHQW URWDPHU SUHIHUHQFHV RI SURWHLQ VLGHFKDLQV 1DW 6WUXFW %LRO ... DSSURDFK WR ELRLQIRUPDWLFV %LRLQIRUPDWLFV

6KHQ +% &KRX .& 3UHGLFWLQJ SURWHLQ IROG SDWWHUQ ZLWK IXQFWLRQDO GRPDLQ ... SUREDELOLW\ DQG VHTXHQWLDO HYROXWLRQ LQIRUPDWLRQ - 7KHRU %LRO ... VFLHQFH 0HWKRGV 0RO %LRO