

MGAlign, A TOOL FOR ALIGNING mRNA SEQUENCES TO GENOMIC SEQUENCES

T. K. B. Lee¹, T. W. Tan¹, H. P. Too¹, S. Ranganathan^{1,2}

^{1/} Department of Biochemistry, National University of Singapore, Singapore

^{2/} Department of Biological Sciences, National University of Singapore, Singapore

In the study of splicing, there is a need for algorithms to align mRNAs and ESTs (Expressed Sequence Tags) to the genomic sequences in order to reveal the exon arrangements. Furthermore with the explosive increase in genomic, EST and mRNA sequences, there is a greater need than ever for accurate and efficient algorithms. Current algorithms search the genomic sequences for matches using portions of the mRNA/EST. These matches are then evaluated and used for the subsequent alignment. We have devised a new algorithm that searches for unique matches on the genomic sequence in order to reduce the number of matches used in the subsequent alignment of the mRNA/EST to the genomic sequence. The portion of mRNA/EST used for the search is selected such that the resulting matches would result in a reduction in the sequence space required for the subsequent alignment. This algorithm is implemented in the program MGAlign, which is available at <http://origin.bic.nus.edu.sg/mgalign/index.html>. The program is suitable for studying splice variants as well as construction of exon databases.