**PREDICTING RNA FOLDING BY GENETIC ALGORITHM WITH LOCAL EXHAUSTION.**

TITOV I.I.+, IVANISENKO V.A., KOLCHANOV N.A.

Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, 10 Lavrentiev Ave., Novosibirsk, 630090, Russia;
e-mail: titov@bionet.nsc.ru

+Coresponding author

Keywords: antisense oligonucleotide, RNA secondary structure, fitness landscape, dynamic phase transition

The additivity of energetic rules for RNA secondary structure is a good basis for validity (important for GA performance) of building blocks hypothesis; in addition, a number of examples demonstrates successful GA application for secondary structure energy minimization. At least, not last, GA could take into account tertiary interactions (e.g., pseudoknots, A.P. Gultyaev et al., 1995)

We present the algorithm for RNA secondary structure prediction that adheres to the following concepts:

1. The greatest possible diversity of initial random set.
2. Extensive search by obtaining the recombination offspring that is maximally distinct from the parents.
3. Deep and fast local optimization by exhaustion in a given neighborhood.

We show that the described algorithm possesses a higher performance compared to the GA with conventional rules. On IBM-486, it takes our algorithm less than a minute to find global optimum state of RNA in a space of 1015 states. The algorithm was tested on the leader region of threonyl-tRNA synthetase mRNA of*E.Coli*and its mutants (H.Moine et al., 1990) and 89.1 % average prediction accuracy was found.

The problem of GA convergence is analyzed using extensive computer simulations and a simple model of GA kinetics. The dynamic phase transition of GA is observed. Combined with mean-field type theory of RNA secondary structure (I.I. Titov, to be published), these results allow us to suggest a criterion for GA convergence in terms of statistical properties of fitness landscape in the vicinity of global optimum.

The "refoldability" of RNA is investigated by computer simulation of refolding of RNA random sequences under the influence of antisense oligonucleotide. The applications of the analysis to molecular processes where RNA refolding under antisense interactions is believed to take place (splicing, etc.) are discussed.

Finally we address the issue whether is it possible to associate GA operators with the processes that govern RNA folding*in vivo*.

The work was supported by the grants INTAS-RFBR 95-0653 and RFBR 97-04-49740 and the Integration Program of the Siberian Branch of the Russian Academy of Sciences.