Fragmenting of algorithm for solving the problem of clustering miRNA families

Anna PYRKOVA [1], Anatoly IVASHCHENKO [2], Raigul NIYAZOVA [2]

[1] *Faculty of Mechanics and Mathematics, and* [2] *Faculty of Biology,*
Al-Farabi Kazakh National University, Almaty, Kazakhstan
E-mail: Anna.Pyrkova@kaznu.kz, a_ivashchenko@mail.ru, raiguln@mail.ru

Abstract: The division of miRNAs into families does not adequately reflect the degree of nucleotide sequence similarity, and thus the categorisation of miRNAs into families requires quantitative criteria defining the differences between miRNA families. The genomes of different organisms have orthologous miRNAs that should be distributed into families. Than it is necessary to establish the degree of similarity of orthologous miRNAs and their belonging to different families. The present research is aimed at solving the problem of clustering miRNAs based on their nucleotide sequence similarity using the fragmented programming method.

Using this developed method, the miRNA nucleotide sequences of Arabidopsis thaliana were clustered. The results identified a conserved polynucleotide, which serves as the Ath-miR156a-j binding site in paralogous SPL mRNAs. These nucleotides encode the conserved ALSLLS motif and the miR156 and miR157 subfamilies belong to the same family. The human miR-1273 family includes miRNAs with very different nucleotide sequences; therefore, based on our program they belong to different miRNAs families.

Keywords: miRNA family, algorithm, fragmented programming method.

References

[1] Eric Londina et al., Proc Natl Acad Sci USA. 2015 112: E1106 [PMID: 25713380]

[2] Wan L, Ding J, Jin T, Guan J, Zhou S., Automatically clustering large-scale miRNA sequences: methods and experiments. BMC Genomics. 2012;13 Suppl 8:S15. doi: 10.1186/1471-2164-13-S8-S15. Epub 2012 Dec 17. PubMed PMID: 23282099; PubMed Central PMCID: PMC3535721.

[3] Kaczkowski B., Torarinsson E., Reiche K., Havgaard JH., Stadler PF., Gorodkin J., Structural profiles of human miRNA families from pairwise clustering. Bioinformatics. 2009 Feb 1;25(3):291-4. doi: 10.1093/bioinformatics/btn628. Epub 2008 Dec 4. PubMed PMID: 19059941.