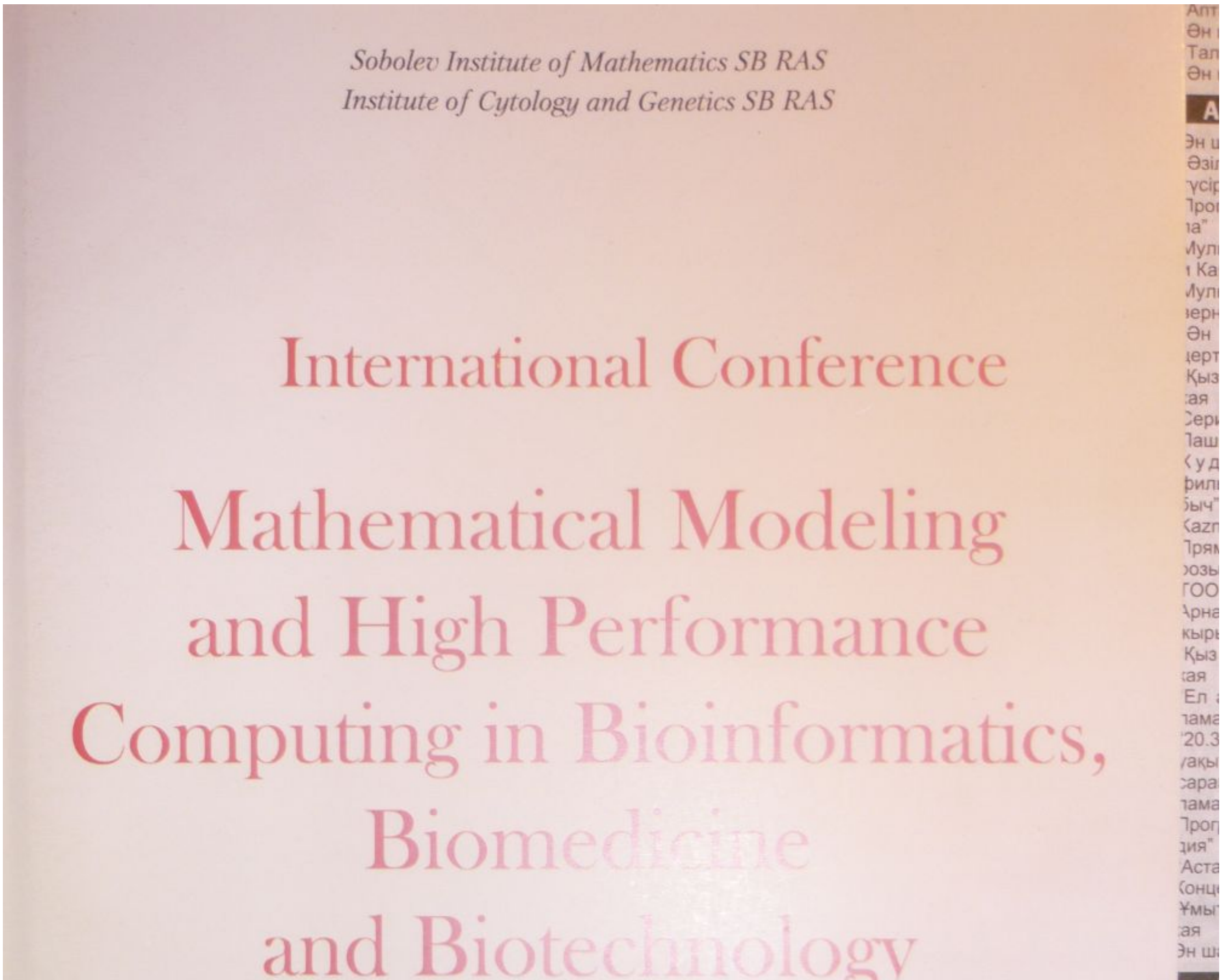

Sobolev Institute of Mathematics SB RAS
Institute of Cytology and Genetics SB RAS

International Conference

Mathematical Modeling and High Performance Computing in Bioinformatics, Biomedicine and Biotechnology

ABSTRACTS

MM-HPC-BBB-2014
NOVOSIBIRSK, RUSSIA
JUNE 24-27, 2014

APPLICATION OF
FOR INVESTIGAT
ON MACROHECT
IN THE LAKE BAI

Afanasyev I.V.
Institute of Computational Mat
e-mail: ivafanas@gmail.com

Key words: cellular automata
Baikal

Motivation and Aim: I
biochemical processes is
ence is usually local, mo
distribution of parameter
leadership in the biomass
own young. The aim of
Comephorus baikalensis
into account spatial indivi

Methods and Algorith
is eight groups of organi
defined between groups c
the model. The model pro
each of them being entitle
account seasonality, orgar

Results: Model verific
rameters differ from asses
developed for computatio
lelized for systems with s
oscillations and non-unifo
seasonality. Non-uniform

Conclusion: Simulatio
polluted territory. Assessn
death of individuals, and t

Availability: Program c

References:
1. V.I. Zorkalcev, I.V. Mokry,
организмов озера Байкал,
2. O.L. Bandman (2010) Cell
Simulating Complex Syster

MODELLING OF THE PROBLEM OF MULTIPLE ALIGNMENT OF THE NUCLEOTIDE SEQUENCES AND DENDROGRAM CONSTRUCTION

Pyrkova A.Yu.*, Ivashchenko A.T., Berillo O.A.
Al-Farabi Kazakh National University, Almaty, Kazakhstan
e-mail: Anna.Pyrkova@kaznu.kz
**Corresponding author*

Key words: algorithm, alignment, dendrogram, Java MPI, mathematical model, miRNA

Motivation and Aim: The detection of the quantitative characteristic of similarity of nucleotide and aminoacid sequences is often problem in genomics and proteomics [1]. The difficulties of the solution of this problem is a great number of RNA sequences and their length. One of the disadvantages of the dynamic programming algorithms is their enough low speed of alignment. The alignment of sequences more than one million symbols or performance of hundreds thousands comparisons of sequences consisted of some thousands symbols is impossible to realize for acceptable time even on the modern computers. Aim of this research is optimize decision of the problem of multiple alignment of the nucleotide sequences and dendrogram construction using the computational cluster.

Methods and Algorithms: For achievement more exact result of multiple alignment and optimization of the time demanded for data processing at multiple alignment, the set of breaks of sequences on the M independent groups processed by M parallel processes. Each of them carries out the alignment of its group independently of others.

Results: The mathematical model of multiple alignment of the nucleotide sequences was constructed. The algorithm of multiple alignment constructed on the basis of the analyzed Needleman-Wunsch algorithm which was modified for processing of files with big data via parallelization using MPJ. The parallelized algorithm of dendrogram construction was realized on the cluster computing platform. The results of work of the program were tested on the nucleotide sequences of microRNAs (miRNAs). Several hundreds and thousands the short miRNA nucleotide sequences with length from 17 to 27 nucleotides was compared. They consist of only four different symbols. Their specific length and content represent a complex problem. The adequate solution of this issue is necessary for identification miRNA binding sites with mRNAs. In this case the received characteristics helped to establish belonging miRNA to one or different families. Interpretation of biological function miRNAs can be depended on the solution of this question. The membership of some miRNAs of Arabidopsis to one family was established by created program. For example, miR156 and miR157 are members of one family as well as ath-miR171 and ath-miR171a belong to another family. The developed program will be applied to distribution 2500 human miRNAs to different families.

Conclusion: Different types the programs can be demanded in process of a sequencing of genomes of animals and plants. The developed algorithm used for analysis of miRNAs of different organisms.

Availability: The program works on the computing cluster of al-Farabi Kazakh National University (http://ursa.kaznu.kz/).

References:
1. N.C. Jones, P.A. Pevzner. (2004) An Introduction to Bioinformatics Algorithms. Massachusetts Institute of Technology Press, Massachusetts.

PARALL
OF miRN
ON THE

Pyrkova A.Y
Al-Farabi Kaz
e-mail: Anna.P
*Correspondi

Key words: p

Motivat
(miRNAs)
tant role of
free energy
are a wide
data. It is
tween nuc
of RNAs.
many of th
that bindi
cation of
tions. Ma
not allow
is creatio
reliability
Metho
nucleotid
sequence
miRNA.
sequence
energy fc
Resul
gap in m
dence of
The algo
miRNA
the nucl
the sites
computa
Conc
program
Avail
tional Ur

Reference
K.M. Guck
aromatic st