

AN OVERVIEW OF THE ALGORITHMIC SESSION OF MCCMB'05

GREGORY KUCHEROV*

LIFL/CNRS

Bât. M3, 59655 Villeneuve d'Ascq, France

Gregory.Kucherov@lifl.fr

ANDREI A. MIRONOV

Department of Bioengineering and Bioinformatics

Moscow State University

Vorobiev Gory, 1-73, Moscow, 119992, Russia

Institute for Information Transmission Problems, RAS

Bolshoi Karetny pereulok 19, Moscow, 127994, Russia

State Scientific Center GosNIIGenetika

Pervy Dorozhny proezd 1, Moscow, 117545, Russia

mironov@bioinf.fbb.msu.ru

MIKHAIL A. ROYTBURG

Institute of Mathematical Problems in Biology

Pushchino, Moscow Region, 142292, Russia

roytberg@impb.psn.ru

MCCMB is a bi-annual international meeting in the general field of computational biology. In 2005, the conference program year included a one-day session devoted to the algorithmic aspects of genome analysis and corresponding software. This is a short overview of the talks of this session arranged according to the general topic.

Genome analysis. This is a most classical bioinformatics topic that includes problems of local and global alignment of genomic sequences and related issues. Grigory Kucherov (LORIA, France) gave a talk on a general method for computing the sensitivity of seeds (small patterns used to spot potential local similarities), and more specifically “subset seeds”, a new concept of seeds. Inna Dubchak (Lawrence Berkeley National Lab, USA) described the VISTA system which is a software system that features several useful programs of genome alignment and analysis.

Motifs in genomic sequences. Mireille Régner (INRIA, France) gave a presentation on how to estimate the statistical significance of motifs using Z-scores, p-values and other statistical estimators. The talk of Eugene Fratkin (Stanford University, USA) dealt with a difficult problem of finding motifs common to a sample of sequences, and proposed a new graph-theoretic approach for this problem.

Andrei Mironov (Moscow University) considered in his talk a new problem — computing significant positional clusters of sites of different score — and proposed an efficient algorithm for this problem. Irina Abnizova (Cambridge, UK) presented an approach to motif finding based on short-memory Markov models.

Systems biology. Oleg Rokhlenko (Technion, Haifa, Israel) presented MetaPathwayHunter, a tool for computing alignments between pathway graphs representing metabolic networks. Paul Horton (Computational Biology Research Center, Japan) introduced the S-systems as a tool for quantitative modelling of genetic networks and presented an algorithm for optimizing those systems.

Classification. Several talks were devoted to classification methods for different bioinformatics problems. Nikolai Zagoruiko (Institute of Mathematics, Novosibirsk) made a talk on extracting the most informative subset of disease-linked genes from expression array data. Alexander Favorov (GosNIIGenetika, Russia) presented APSampler, an algorithm for identification of allelic variants responsible for complex diseases based on the Markov chain Monte Carlo technique. Finally, Sang Jun Kim (Seoul University, Korea) presented a new haplotype reconstruction algorithm which is an improvement of the well-known Clark's algorithm.

RNA. Two talks dealt with RNA modelling. Mikhail Roytberg (Institute of Mathematical Problems in Biology, Russia) presented a new efficient algorithm for RNA folding based on the technique of sparse dynamic programming. Tariq Alsheddi (George Mason University, USA) in his talk studied the problem of designing possible small interference RNA (siRNA) having the property of being specific with respect to a set of target sequences under study.

The above overview shows that the algorithmic session featured interesting talks giving a representative image of the modern directions in algorithmic studies in computational biology. We consider it as a successful event and hope that this subarea will continue to be represented at future MCCMB conferences, possibly growing up into a satellite workshop.