

Archive for Project for CS-LS for year 2004

Isaiah (Shy) Arkin

Department of Biological Chemistry

Tel/Fax: +972-(0)2-6584329

Email: arkin@cc.huji.ac.il

Computational modeling of membrane proteins based on silent amino-acid substitution data.

We have decided to employ mutagenesis and evolutionary conservation data in a novel way to simulate membrane proteins. Rather than looking at those residues which destroy the complex (or highly conserved during evolution) as we have done previously, we are now concentrating on the effect of non-disturbing residues in global search schemes. The reason being that what these non-disturbing are doing is to obstruct non-native structures (out-liars) but have little effect on the native structure, thereby enabling one to determine what is the correct structure, without the need for any prior assumptions. So far we have shown that this method predicts the correct structure for human glycoporphin A dimersing transmembrane alpha-helices based on silent mutations and evolutionary non-conserved residues. The wealth of readily available evolutionary conservation data provides potential for the widespread application of this method, not only to polytopic membrane proteins using global searching molecular dynamics, but in concert with any technique that generates multiple candidate protein structures. One area in which this method will be readily applicable would be low-resolution EM structures in which the positions of the helices are known (in terms of translation and tilt), but the their relative rotation with respect to each other is unknown. Multiple homologs can then be used in simulations, in the hopes of finding a single structure that persists in all instances.

Sigal Ben-Yehuda

Department of Molecular Biology

Phone number: 6758600

e-mail: sigalbe@md.huji.ac.il

Identification and comparison of RacA binding sites among bacilli

Triggered by nutrient limitation, *Bacillus subtilis* enters a pathway of differentiation that culminates in the formation of a dormant stress-resistant cell-type called a spore. Entry into sporulation involves the remodelling of the sister chromosomes into an axial filament structure and is mediated by RacA, a novel DNA binding protein that we identified. Biochemical, microarray and bioinformatics approaches revealed a potential RacA binding sequence which is present 20 times in the *B. subtilis* chromosome in a centromere-like region. We would like to find out if these binding sequences and their adjacent regions are conserved among other bacilli (that their genome sequence is known) and whether RacA homologues in these species have similar recognition sequences and binding properties.

Yehudit Bergman

Department of Experimental Medicine and Cancer Research

Phone number: 6758362

e-mail: yberg@md.huji.ac.il

Mapping the murine immunoglobulin variable region genes.

Opening and closing of gene expression is a very basic phenomenon. We would like to assemble the 100 immunoglobulin variable region genes together (about 2 mega base), and to find unique sequences that we will use to analyze their chromatin structure. This will help us in analyzing the kinetics of the opening up of particular variable genes, as a function of their relative location in the locus during B cell differentiation. This will help to understand chromatin changes that occur in very complex loci.

Guy Bloch

Evolution, Systematics, and Ecology (ESE)

Phone: 6584320

bloch@vms.huji.ac.il

Develop methods to compare the amplitude and significance of circadian oscillations in gene expression data from microarrays.

Honey bees show socially-mediated chronobiological plasticity in their circadian clock. Bees that forage for flowers (iforagersⁱ) outside the hive have strong circadian rhythms that are used for sun-compass navigation and to time visits to flowers. By contrast, nest bees that care for brood (inursesⁱ) are active around the clock with no circadian rhythms. The nursesⁱ arrhythmicity may be functionally significant because honey bee larvae need continuous care. We use microarray technologies to explore the molecular underpinning underlying this ecologically relevant chronobiological plasticity. The aim of the proposed project is to develop computational tools necessary to compare oscillations in gene expression between nurses (behaviorally arrhythmic) and foragers (behaviorally rhythmic). Genes that show strong circadian oscillation in foragers but not in nurses may be involved in chronobiological plasticity.

Dan Cohen

Dept. of Evolution, Systematics and Ecology. The Silberman Institute of Life Sciences.

Phone: 972-2-6584640.

E-Mail: dancohen@vms.huji.ac.il

1. Modeling biologically plausible algorithmic mechanisms of genetic evolution
2. Individual learning of multi-level and multi-step optimal learning rules and preferences.

Optimal evolution and learning of complex tasks by living organisms necessarily include multi-level sampling and learning of many intermediate steps and many levels of performance. Important principles for the achievement of such optimal algorithms have been recently developed as Reinforcement Learning (RL) algorithms. I suggest that RL principles can be effectively applied for developing optimal sampling and learning rules of multi-step and multi-level tasks.

I can suggest several possible interesting biological phenomena and processes that can be investigated using the RL methodology, and many more are also possible:

1. Optimal sampling and learning of complex spatial movement patterns, e.g. foraging, escaping predators.
2. Optimal sampling and learning of complex multi-step tasks, e.g. foraging and hunting
3. Optimal sampling and learning of a complex hierarchy of multi-level tasks. e.g. optimal allocation of sampling and learning effort between multi-level skills and preferences.

Amatzia Genin

Dept. Evolution, Systematics & Ecology. Inter-university Institute for Marine Sciences in Eilat

Phone number: (08) 636-0124

e-mail: genin@huji.ac.il

Acoustic tracking and behavior of microscopic plankton in the ocean

Zooplankton are minute animals suspended in the water column from the surface to the abyss. They form a major link in the food web of all oceans and lakes. Their importance to the general ocean ecology is immense and understanding their behavior is a fundamental objective in biological oceanography. Yet, due to their small size and transparency almost nothing is known about their behavior. Together with a group of marine engineers from San Diego (USA), we have recently developed a multi-beam, high-frequency sonar that allows tracking of individual zooplankton in the sea. This is a one-of-a-kind instrument, available only at my lab. So far we have developed only basic algorithms to acquire data and process it for individual tracks in 3-D. Our objective is to develop new algorithms to:

- (a) automatically characterize the swimming behavior of zooplankton (e.g. swimming directionality, curvature, active swim vs. passive swept with the currents).
- (b) visualize 3-D tracks using advance computing tools.

Joseph Hirschberg

Department of Genetics

E-mail: hirschu@vms.huji.ac.il

Phone:(972)2-6585662

Gene discovery using cDNA microarray analysis.

The purpose is to discover new plant genes with potential importance for biotechnology. Work will focus on scanning specific arrays and analysis (computational mainly) of the results.

Alex Levine

Plant Sciences

Phone 658 6543.

Email: AlexLevine@huji.ac.il

Title: Reconstruction of Stress Signal Transduction Networks in Plants

Elucidation of the signal transduction pathways regulating stress and disease resistance in plants is presently one of the hottest fields in Plant Biology. The completion of the Arabidopsis genome sequencing and almost complete genome data already available from rice and maize plus partial genomic information on numerous other plants make genomic approaches feasible for the study of plant defense mechanisms, and identification of the regulatory genes.

The proposed project will be based on analyzing Arabidopsis Microarray gene expression data collected during various stress conditions from public (Stanford, European consortium, etc.) sources, and criss-crossing the transcription data across stresses and across species to select genes induced early-on by very specific stresses. Since plant cells lack specialization/differentiation at the level of animal cells, they must rapidly adjust their transcription to given stress conditions, making plant-derived expression data a better source for gene hunting by genomic methods than in animal systems. The selected genes will then be analysed for already present functional genomics data and also compared to the functional information of the gene homologs in yeast. Yeast homologs will also be used to identify the interacting partners of the selected regulatory genes for further elucidation of the signaling networks that operate during stress.

The tools that will be used throughout the project will be based on some of the already present Bioinformatics programs, but will necessitate development of combining various programs that will enable us to analyse genes on a global genomic level rather than selecting individual genes or gene families.

Michal Linial*

Dept of Biological Chemistry

Institute of Life Sciences

mail: michall@cc.huji.ac.il <http://www.ls.huji.ac.il/~michall/>

*on sabbatical, working with a graduate student

1. Informative features in protein families

Many protein families can be described by a small set of features. Those features may provide rich functional information. For example, a protein having 12 transmembrane domains is probably a transporter while 7 transmembrane domains is a strong predictor for a membrane receptor. Similarly, short proteins with abundant cysteines may belong to growth factors or to snake toxins. How can a family be described in condensed, reduced representation that captures functional relevance. Protein families participate in endocytosis and in cytoskeletal dynamics will be used to formulate the problem.

2. Representation of the protein space via annotation and experimental data.

Collecting raw data from large-scale experiments is a must for all high throughput technologies. This trend asks for developing new ways for presenting data and more importantly, for analysing such data. We suggest developing tools based on quantitative and qualitative information. Data of gene expression levels, protein-protein interactions, protein levels, localization data etc can be all combined to present groups of proteins based on their associated parameters and annotation.

3. Data-mining and statistical value of large-scale proteomics and Gene Expression data.

4. Target selection for Structural Genomics – Evaluation of the current state using computational and statistical tools.

Ed Mitrani

mitrani@vms.huji.ac.il

1. Tissue engineering from embryonic stem cells.
2. Replacing by-passes with normal blood vessels.

Etana Padan

Molecular Microbial Ecology, Life Sciences,
Tel: 65 85094
Etana Padan <etana@vms.huji.ac.il>

Genomics of Na⁺/H⁺ antiporters.

Na⁺/H⁺ antiporters are ubiquitous polytopic membrane proteins composed of 12 trans membrane domains. They exist in the cytoplasmic membrane of all cells and membranes of many organelles and are responsible for H⁺ and Na⁺ homeostasis. We cloned the first genes encoding antiporters, *nhaA* and *nhaB* of *E. coli* and now the genome project provided a burst of antiporter genes that are ready to be classified with a genomic and proteomic approaches.

Ophri Pines

Department of Molecular Biology
Phone: 6757203
e-mail: ophry@md2.huji.ac.il

Signal peptides and exon boundaries

Signal peptides are located in the N terminus of proteins and target them to various cellular compartments. Is there a correspondence between the signal peptide/protein boundaries and exon/intron/exon boundaries in the mRNA?

Anton Post

Dept of Microbial and Molecular Ecology, Givat Ram, Marine Biology Lab, Eilat
Phone number: 85090 or 08-6360122
<anton.post@huji.ac.il>

Cyanobacterial Genome Analyses for Abundance of Transport Proteins and their Transcriptional Regulation. (continuation)

Cyanobacteria are important components of aquatic plankton communities. They contribute about 50% of global yearly photosynthesis and are especially abundant in the ocean. Their success in natural ecosystems is largely determined by their capacity to interact with the environment. This is most explicit in their capacity to take up essential nutrients and other solutes via transport protein complexes which are located in the cytoplasmic membrane. To date 8 cyanobacterial genomes have been analyzed, among which are 4 marine and 4 freshwater strains. Three more genomes are in progress. A visual inspection of these genomes suggests that marine strains rely on permeases as transport systems, whereas freshwater cyanobacteria employ ABC-type transporters. Across the board analyses of these genomes for the identification of genes encoding membrane proteins has not been performed so far. Moreover, these genes form excellent targets for identification of regulatory circuits through the inspection of their upstream sequences which contain signatures characteristic for e.g. two component regulatory systems or those of CRP type transcriptional activators. Such systems are known to operate in cyanobacteria. The project will require development and/or adaptation of algorithms and search strings for the specific genome analyses. Outcome will be interpreted against an ecological/physiological background and so give an extra dimension to the project.

Tsvi Sachs

Dept Plant and Environmental Sciences
Phone: 6585253

Random elements in leaf vein patterns

Contrary to common text book statements, ontogeny need not follow a strict developmental program. Unpredictable, possibly random details pervade biology, as seen, for example, in chromosome movements, the course of early embryonic divisions of mammal embryos, and the detailed course of blood vessels. Leaf vein patterns are a convenient system for the study of such unpredictable elements, which has been largely neglected. Digital pictures of all the veins of leaves [including the ones not seen without some preparation] of a single and of genetically identical plants can be readily obtained. Their functional roles are known and an hypothesis about the determination of their patterns is currently receiving molecular support. Quantitative work could define the parameters of vein distribution that are least variable. Though the possibility has not been studied, let alone measured, it is possible that the

presence of a vein is influenced by vein density in some unknown area and thus by the occurrence of veins that are not its immediate neighbors. Any results should thus be significant to discussions of both developmental mechanisms and vein function. Variable structure is presumably not the most efficient possibility, and a reasonable hypothesis that requires testing is that robust, plastic development compensates for possible functional prices. The work could be continued to include quantitative effects of stress conditions, a comparison of leaves of different types, and other biological patterns.

Shimon Schuldiner

Silberman Bldg. 3-345

Phone: 972-2-6585992

Email: Shimon.Schuldiner@huji.ac.il

Membrane proteins: what to do while waiting for high resolution structures?

Of the genomes sequenced so far, over 20% of the gene products are known or predicted to be membrane proteins. These proteins catalyse a multitude of essential functions, one of which is the transport of molecules into and out of cells or intracellular organelles, and across epithelia. However, membrane proteins in general — and membrane transport in particular — are notoriously resistant to the determination of high-resolution structure.

Although advances in molecular biology have led to rapid progress in understanding structure–function relationships for some membrane proteins, structures have been obtained at atomic resolution in only a handful of instances. So, the level of understanding of membrane proteins is almost inversely proportional to their roles in living systems.

Others and we are now developing and using a battery of non-traditional approaches to structure. These include a number of site directed biochemical and biophysical approaches. In addition, the wealth of available novel information provided by the genome sequence projects offers an opportunity to study the differences and similarities “in silico” and express and clone relatively rapidly proteins with properties different from the ones we characterized until recently.

Several projects can be performed in our lab that is now using the above-mentioned approaches to study EmrE, a bacterial multidrug transporter. EmrE is one of the many proteins that have been associated with resistance to the effects of multiple drugs, antibiotics, and antineoplastic agents, a phenomenon that poses a serious problem in treatment of infectious diseases and resistant tumors.

Nadav Shashar

The Interuniversity Inst. of Eilat

Phone: (972)-7-6360123 (direct) (972)-7-6360114 (lab) (972)-7-6360111 (main office)

E-mail: nadavs@cc.huji.ac.il

How does a fish look to another fish?

Generating a computerized simulation for the reception of a fish body pattern at the eye of another fish.

Coral reefs are considered among the most colorful environments on earth. Much of this spectral variability arises from the body coloration of fishes living in and around reefs. Indeed coral reef fishes are exceptional between animal groups in the wide distribution complex body patterns- both texture and spectra- otherwise known as colorfulness, which appear in a large portion of reef species. In general body coloration has two visual functions a) camouflage b) signaling. The colorfulness of coral reef fishes gives the impression that many of these fishes abandon the camouflage function all together. However, this description is often bias since it is mostly based on the human perception of the fishes.

The proposed project will attempt to create a computerized simulation of the way a fish may be perceived at the level of the eye of another fish, either from the same species or from another species. The simulation will take into account physical properties such as optical properties of the seawater, spectral and texture properties of body coloration of the observed fish, and the distances of visual interactions; along with neuro-anatomical information such as the visual acuity and spectral sensitivity of the observing fishes.

The project will take place at the Inter Univ. Inst. for Marine Sciences in Eilat and at the computer facilities of the Hebrew Univ. Accommodation in Eilat will be provided. It will include outlining and breakdown of the question and project in Eilat, participation in data collection in Eilat, and generation of the simulation in Jerusalem and in Eilat. In general this is an image-orientated project with significant parts of mathematical calculations. The simulation is expected to be developed gradually and in a modular fashion, with each section examining the impact of a specific factor in the visual pathway.

Itamar Simon

Department of molecular biology
Phone number 6758544
Email - itamars@md.huji.ac.il

What is the role of the yeast Histone H1?

Histone H1 in mammals has an important role in chromatin condensation. In yeast there is also an histone H1 protein but its role is not clear. The gene is not essential and the mutant has no phenotype. We have used location analysis to profile the genome wide interactions between the histone h1 protein and the yeast genome. This data is available for analysis in light of other published microarray datasets.

Is stress response cell cycle dependent?

Yeast responds to stress (for example exposure to H₂O₂) by changing the expression of numerous genes. However it is not clear whether this response is cell cycle dependent. In order to examine this question we have performed expression profile experiments on yeast arrested in distinct cell cycle stages that we exposed to stress. The analysis of this data may reveal novel connections between stress response and cell cycle.

Hermona Soreq,

Department of Biological Chemistry
Phone: 02-658-5109

Title of the project: Alternate promoter usage and alternative splicing in the mammalian transcriptome

Both alternate usage of multiple promoters and alternative splicing of pre-mRNA transcripts contribute significantly to the proteomic and functional complexity of the mammalian transcriptome. However, the rule(s) controlling the suspected linkage between these two processes have not yet been determined. Several principal approaches may be utilized to approach this intriguing subject, including EST database analysis, search for associations between consensus promoter sequences and specific splice variants or datamining of relevant research articles.

Ruth Sperling

Department of Genetics
Phone: 972 2 658 5162
E-mail: sperling@vms.huji.ac.il

1. A novel nuclear proofreading mechanism as part of the splicing machine.

This project is based on computer data base analysis, and experimental results.

2. Three dimensional image reconstruction from electron micrographs of the native splicing machine.