STANFORD UNIVERSITY
CENTER FOR COMPUTATIONAL GENETICS AND BIOLOGICAL MODELING

Biocomplexity Faculty Search Committee,
c/o Prof. Rob de Ruyter van Steveninck,
Biocomplexity Institute,
Indiana University,
Swain Hall West 117,
Bloomington IN, 47405-7105.

Dear Prof. Rob de Ruyter van Steveninck,

I write in order to apply for a Systems / Computational Biology faculty position in the Biocomplexity initiative The College of Arts and Sciences at Indiana University Bloomington. I have been a Senior Scientist and co-director (and founder) of The Center for Computational Genetics and Biological Modeling at Stanford University for the last six years.

My research agenda involves multidisciplinary investigation into evolutionary and developmental biology that can be approached using a combination of mathematical, computational and experimental tools. Though the focus of my program is mainly on quantitative aspects, the relationship between these subjects and experimental molecular genetic studies of evolution and is an integral part of my efforts.

My current research was posed initially by Waddington some 60 years ago, that is, why is phenotypic expression robust, or canalized, to mutations and environmental variations, while occasionally, genetic mutations or environmental assaults disrupt the system sufficiently to cause phenotypic disorder? This is the context in which I study an important property of complex genetic systems, namely, their apparent robustness in the face of mutations and environmental variation. I model gene networks in order to generate hypotheses and predictions that are then tested with biological data. These tests are made possible by systematic genomics efforts undertaken in recent years. One example is provided by the project to delete each gene in the yeast *Saccharomyces cerevisiae* and to assay the effects of each knockout on the expression of other genes. Analyses of microarray and other data in the context of canalization will shed light on the evolution of robustness and other properties of complex genetic systems. My approach also helps us to understand the contribution to proper development of specific control mechanisms, such as Hsp90 and other chaperones.

To further our understanding of the phenomenon, in a collaborative experimental investigation, I study variation in segmental patterning in response to a null mutation of a gene involved in the *Drosophila* segment-polarity network. This network has been chosen for two reasons: it is known to be highly conserved across all insect species; and there exists a detailed knowledge of its components and their interactions. It has been recently been debated whether the observed conservation is due to robustness of the network or to strong stabilizing selection on pleiotropic effects of mutations in the network. Our study addresses the present controversy, furthermore, correlating phenotypic response to null mutations, with the genotypic variation, using microarray, makes it possible to identify genes as member of the segment polarity network that have not been identified otherwise. In turn, more realistic model of the segment-polarity network will be built leading to new hypothesis regarding phenotypic response to variation in the network's components and environmental heterogeneity.

In the long run these studies will enhance our understanding of, the relationship between genetic network architecture and phenotypic fidelity; gene-environment (e.g. gene-nutrient) interactions in complex disease, such as diabetes, cancer, and heart diseases; the response of different genetic background, such as single nucleotide polymorphisms, to environmental variation; as well as, the relationship between micro- and macroevolution.

Enclosed you will find the relevant materials to support my candidacy. I would be happy to provide any additional materials that you might find useful.

Yours sincerely,

Aviv Bergman, Ph.D.

385 SERRA MALL (HERRIN LABS 428) • STANFORD, CALIFORNIA 94305-5020
TEL: (650) 725-7727 • FAX: (650) 725-7745 • compgen@compgen.stanford.edu

# *Synopsis of current and future research directions.*

## *Goals:*

***Understanding the Evolutionary Causes and Consequences of Robustness in Development***

## *Approach*

Canalization, a concept which relies heavily on the notion of epigenetic landscape, was introduced by C.H. Waddington to explain the robustness of developmental processes in the face of environmental and genetic variation, remains one of biology's most confusing and contentious subjects. A fundamental question concerning robustness is whether it constrains or promotes evolutionary divergence. At first glance, it may seem natural to suppose that robustness only constrains divergence, because it prevents mutations from having phenotypic consequences. However, by providing for the buildup of stores of genetic variation, robustness may ultimately promote divergence.

The prospect of bringing experimental and computational approaches together to a greater degree to address such question in the context of canalization is particularly exciting. Whereas the modeling approach incorporates a rather abstract representation of biological systems, the growing body of knowledge of actual developmental pathways, such as regulatory gene-networks, makes it possible both to construct more realistic models and to test the predictions of the models empirically. For example, to enhance the biological relevancy of an abstract computational model, it should be possible to incorporate an explicitly model of Hsp90 and its interaction with specific signaling pathways, so as to compare the relative effects on robustness and 'evolutionary capacitance' of the buffering afforded by Hsp90 and the buffering intrinsic to complex networks. Such a model might build on simulation study of known signaling and metabolic pathway, such as the mitogen-activated protein kinase (MAPK) cascade. On the experimental side, it is now feasible to perform a Waddingtonian artificial selection experiment with a microorganism such as yeast, and to trace in the selected lineages not only the phenotypic changes that occur, but also the expression changes in every gene, some of which must underlie the phenotypic evolution. Such an experiment would no doubt yield valuable insights into the regulatory and evolutionary properties of the network involved, and would suggest new directions for theoretical work.

In the long run these studies will enhance our understanding of, the relationship between genetic network architecture and phenotypic fidelity; gene-environment (e.g. gene-nutrient) interactions in health and complex disease, such as diabetes, cancer and heart diseases; the response of different genetic background such as single nucleotide polymorphisms, to environmental variation; as well as, the relationship between micro- and macroevolution.

Bellow I describe my theoretical and experimental studies that follow this approach as it applies to 'robustness'.

## *Evolutionary Causes of Robustness*

Ultimately, according to the neo-Darwinian paradigm, divergence between species is the product of micro-evolutionary processes taking place within species. Thus, if one wishes not just to answer how developmental systems evolve, but why they evolve the way they do, one needs to understand the forces shaping within-species genetic variation, thus, to understand the developmental-genetic influences on the expression of this variation. Since

Waddington first laid out his theory of developmental canalization there has been much theorizing on the subject, yet very little consensus, let alone direct experimental data. Recently, however, I conducted an analytical and numerical study to address this issue which led us to propose a biologically plausible genetic mechanism that buffer genetic and environmental variation.

Most species maintain abundant genetic variation and experience a range of environmental conditions, yet phenotypic variation is low. That is, development is robust to changes in genotype and environment. It has been claimed that this robustness, termed canalization, evolves due to long-term natural selection for optimal phenotypes. Understanding this phenomenon as it pertains to traits such as cancer, heart disease and other complex traits requires the understanding of the developmental process at organization levels from gene activation through network of interacting genes to population dynamics. To understand the evolutionary properties of robustness and is breakdown a detailed study of networks properties and architecture is to be understood. I started this investigation using computational approaches to the study of the evolutionary causes and consequences of robustness. In this study I have shown that the developmental process, modeled as a network of interacting transcriptional regulators, constrains the genetic system to produce canalization, even without selection toward an optimum phenotype. The extent of canalization, measured as the insensitivity to mutation of a network's equilibrium state, have shown to be dependent on the complexity of the network, such that more highly connected networks evolve to be more canalized. Thus, I argue that canalization may be an inevitable consequence of complex developmental-genetic, epigenetic, processes, and thus requires no explanation in terms of evolution to suppress phenotypic variation.

***Evolutionary Consequences of Robustness***

The harboring of genetic variation lead to the potential to increase the rate of evolution if revealed when needed. The investigate the evolution of mechanism revealing variation as well as their consequences, I begun my study by revisiting a recent study by Lindquist and colleagues on the buffering properties of the chaperone protein Hsp90. The molecular chaperone Hsp90 has been termed by, the Lindquist and her colleagues an 'evolutionary capacitor' because it has been shown, in *Drosophila* and *Arabidopsis*, to: 1) suppress phenotypic variation under normal conditions and release this variation when functionally compromised, 2) have its function overwhelmed by environmental stress, and 3) exert pleiotropic effects on key developmental processes. There still is a considerable debate whether these properties necessarily make Hsp90 a significant and unique facilitator of adaptation. Using numerical simulations of complex gene networks, as well as genome-scale expression data from yeast single-gene deletion strains, I presented a novel mechanism that extends the scope of evolutionary capacitance well beyond the action of Hsp90 alone. We illustrated that most, perhaps all, genes reveal phenotypic variation when functionally compromised. It was also shown, in the first demonstration of the plausibility of evolutionary capacitance, that the availability of loss-of-function mutations speeds adaptation to a new optimum phenotype. However, this effect does not require the mutations to be conditional on the environment. Thus, I maintain, there may exist a large class of evolutionary capacitors whose effects on phenotypic variation complement the systemic, environment-induced effects of Hsp90.

***Robustness: Experimental Study*** **(current and future study)**

To further our understanding of the phenomenon, in collaboration with Mark L. Siegal, I study variation in segmental patterning in response to a null mutation of a gene involved in the *Drosophila* segment-polarity network. This network has been chosen for two reasons: it is known to be highly conserved across all insect species; and there exists a detailed knowledge of its components and their interactions. Recent computational study, using differential-equation model of the *Drosophila* segment-polarity network, showed that the equilibrium expression state, which defines segmental pattern in the developing embryo, is robust, in that the correct output is insensitive to a wide range of parameter values for the kinetic constants in the model. From this observation, it was concluded that the conservation of the segmented germband stage in insects results from robustness of the segment-polarity network. It has been recently been debated whether the observed conservation is due to robustness of the network or to strong stabilizing selection on pleiotropic effects of mutations in the network. Clearly the computational model must be extended to incorporate all known inputs to the segment-polarity network, as well as the numerous patterning events that the network organizes, so that the effects of pleiotropy can be more fully appreciated. The computational approach should also include a population-genetic process, for as we have seen above, robustness does not necessarily imply evolutionary stasis. We use the transcription factor *cubitus interruptus (ci)*, a component of the segment-polarity network, as the gene whose functionality is compromised by a null mutation, $ci^{94}$ , which causes segmentation defects evident on the ventral surface of the larva. We use 13 inbred (isogenic) wildtype *Drosophila* strains of diverse geographic origin as a source of variation in genetic background representative of wildtype populations. The null mutation is introgressed into the 13 strains to study the pattern of variation between and within both mutant and wildtype strains. This study addresses the present controversy, furthermore, correlating phenotypic response to mutations in *ci*, between the 13 different strains with their genotypic variation, using microarray, makes it possible to identify genes as member of the segment polarity network that have not been identified otherwise. In turn, more realistic model of the segment-polarity network could be built leading to new hypothesis regarding phenotypic response to variation in the network's components.

***The Evolution of Evolvability***

In a closely related theoretical study we investigated the evolution and evolvability properties of the yeast, *Saccharomyces cerevistae,* prion [*PSI*$^+$]. Here evolvability is defined as the rate of appearance of heritable and potentially adaptive phenotypic variants. The increase in phenotypic variants occurs when the appearance of prion causes read-through translation and reveals hidden variation in otherwise untranslated regions. Eventually the portion of adaptive phenotypic variation loses its dependence of this molecular mechanism. This molecular mechanism, just as the mechanism of the chaperone Hsp90, can be seen as analogues to revealing hidden variation through the break of buffering. In the model, the mechanism is reversible, so that restoration of normal termination conceals the revealed deleterious variation, restoring the organism's fidelity. It is known that the ability to from [*PSI*$^+$] is known to be conserved in yeast, using constructed a mathematical model to calculate whether this ability is more likely to have become fixed due to chance alone or due to is adaptive evolvability characteristics. We have shown that is environmental stress induces partial read-through at a frequency as low as once every million years. Thus, we concluded that the observed conservation of the yeas prion is more likely explained by the adaptive advantage of its evolvability properties.

***Role of Network Topology*** **(current and future study)**

It has been suggested that functional modules and their interconnections occupy a fundamental level of biological organization at which to understand such properties as robustness and evolvability. Indeed, the very fact that certain types of modules, such as signal-transduction cascades, are few in number, deeply conserved, and integral to generating the vast diversity of animal form, lends credence to this claim. Simulations of the dynamics of such modules, such as the mitogen-activated protein kinase (MAPK) cascade, clearly add to our understanding of their evolutionary properties. Still there is much to be gained from the detailed investigation of pathways that do not constitute conserved modules, such as the *Drosophila* sex-determination pathway. There is also much to be gained by maintaining a healthy skepticism toward claims that higher-order network organization determines network behavior to a large degree. To that end, using numerical simulation, I study the properties of a class of networks obeying power law distributions in their connectivity. It has been conjectured that most biological networks fall within this class. Preliminary results makes it unclear whether the top-down inferences about network properties drawn from these observations offer much predictive power for individual genes. Put another way, I question whether knowing how many *cis*-regulatory inputs a certain transcription factor has, or how many direct targets it has, necessarily tells us anything useful about its buffering effects on interacting genes or about its propensity to diverge between species. My preliminary results suggests that unique details of particular network configurations are far more determinative in this respect, furthermore, they may reveal correlation between the position of a gene in a network and its pattern of divergence.

***Mode and Tempo of Evolutionary Changes*** **(current and future study)**

To relate this study to the broader issue of the relationship between micro and macro evolution, I plan to continue to study questions pertaining the time scale in which we are likely to observe a pattern of punctuation as opposed to gradual evolution. Some of the remaining issues include: the effect of selection, stabilizing and directional, on the sojourn time. Results based on diffusion theory have been obtained for the expected sojourn time for a finite population under drift, but very little has been done to estimate this time under selection. My findings for a finite population under drift are in accord with the analytical results, but we have no such expression for comparison when selection is present. I will explore properties of the sojourn time under various selection regimes and will attempt to draw, if not quantitative results, qualitative results pertaining to the sojourn time under selection.
In addition I am now in a position to forge a link between the conserved functionality due to developmental constraints and the stasis periods. That is, is the observed long periods of phenotypic stasis are the result of constraints imposed by a developmental process or are they the result of interaction between drift and mutation, as been proposed earlier? To address this issue, I plan to study the time course of the evolutionary trajectories, with and without developmental stages.

**Name** Aviv Bergman

## Teaching interests

It has long been recognized that scientific breakthroughs and groundbreaking research in the coming century require multidisciplinary approaches to many areas of research. These include crucial problems in the biological sciences, such as understanding and interpreting vast quantities of genetic information (at the sequence as well as the functional level), structural biology, gene and protein network analysis, and evolution. These types of research require the involvement of biological, mathematical, computational, and engineering sciences. Today, most biologists are already engaged in such multidisciplinary research in their daily practice. However, our education curricula have not been adapted to these changing needs. Teaching for scope and depth at the same time, however difficult, is an urgent matter in training our future scientists.

One of my basic commitments as a scholar and practicing scientist is to the development of new integrative curricula on both the graduate and under-graduate levels. I have experience in (and enjoy) teaching traditional courses in population dynamics, biostatistics, mathematical modeling, and simulation techniques, as well as courses on general evolution. However, an example of the type of integrative course that I am interested in developing would concentrate on the interface between development and evolution. It would first introduce classical approaches to the study of development and evolutionary biology. It would turn then to methods of analysis of molecular information generated by modern technology such as microarrays. Tools like these provide the bases from which to infer molecular mechanisms. The study of how specific systems evolve, e.g. *Drosophila melanogaster* sex determination hierarchy, can be used pedagogically in valuable ways. Through case studies of these systems, mechanistic models can be constructed which illustrate fundamental evolutionary questions. The course outline will look like this:

**The Evolution of Development**

**Text Books:** Required: *"The Evolution of Developmental Pathways"*-Wilkins; *"Statistical Methods in Bioinformatics"*- Grant & Ewens; *"The Nature of Mathematical Modeling"*-Gershenfeld; *"Genomic Regulatory Systems"*-Davidson. Optional: *"Mathematical and Statistical Methods in Genetic Analysis"*-Lange; *"Principle of Development"*-Wolpert; *"Biometry"*-Sokal & Rohlf.

- Introduction to Developmental Biology
  - History and basic concepts
  - Genotype vs. Phenotype
  - Structural vs. Functional
  - Model organisms.
- Molecular Evolution
  - History and basic concepts
  - Population dynamics
  - Recombination mapping and linkage
  - Mutant phenotypes and the physical mapping of genes
- Mathematical Modeling
  - Analytical Models
    - Ordinary differential and difference equations
    - Introductory partial differential equations
    - Random systems
  - Numerical Models
  - Observational Models
  - Graphical and mathematical software
- Statistical Hypothesis Testing
  - Parametric statistics and Non-parametric statistics
  - Sample analysis of microarray data (e.g. yeast knockout)
- Genetic Networks in Development
  - Segmentation
    - Activation of pair-rule genes
    - Segment polarity genes
    - Selector and Homeotic genes
  - Determination of the sexual phenotype
- Case Study in Developmental Processes Evolution
  - Metabolic Pathways
  - Sex determination hierarchy and its representation
  - Modification of developmental processes in evolution