UNIVERSITY OF NOTRE DAME
College of Science
DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY

Dr. Greg Warr

Cc: Dr. Parag Chitnis
Dr. Karen Cone

September 25, 2012

Dear Dr. Warr,

We are pleased to submit our report of the NSF-sponsored meeting on Evolutionary Cell Biology, supported by grant MCB 1228570, that took place from May 29-June 1, 2012 at the Airlie Conference Center. The meeting was organized to define a nascent field of Evolutionary Cell Biology, to identify its most important aims, and to develop a strategy for this endeavor. The participants, consisting of twenty-four scientists from North America, Europe, and Australia, included both cell and evolutionary biologists, but also mathematicians and physicists with interests in biological problems. This turned out to be a very stimulating mix, and a wide range of ideas was considered in spirited formal and informal discussions. The central recommendation, around which the group coalesced, was that the field, and indeed all areas of biology, would benefit enormously from a well-coordinated and systematic sampling of true organismal diversity at the levels of cellular architecture and behavior in addition to sequence. Among other things, this effort would greatly enhance our ability to understand the structure and function of cellular machinery in currently well-studied systems. Moreover, in order to understand how such structures evolved, it will be imperative to use phylogenetic considerations to develop a collection of experimentally-tractable organisms that represent key lineages. The proposed initiative is called the Atlas for the Biology of Cells (ABC). We hope that you will find it to be as exciting as we do.

Sincerely,

Holly Goodson, Michael Lynch, and Aaron Turkewitz

Holly Goodson
Dept. Chemistry and Biochemistry
University of Notre Dame
hgoodson@nd.edu

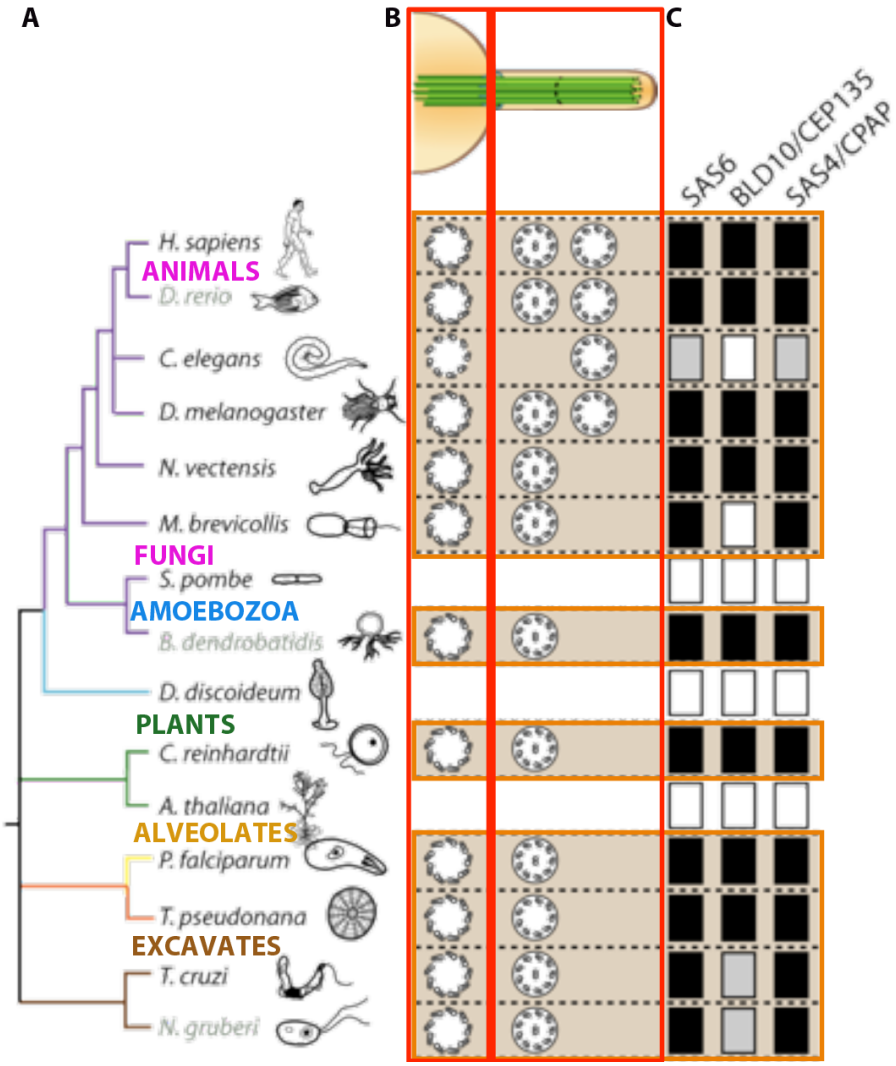
Michael Lynch
Department of Biology
Indiana University, Bloomington
milynch@indiana.edu

Aaron Turkewitz
Dept Molecular Genetics and Cell Biology
The University of Chicago
apturkew@uchicago.edu

251 NIEUWLAND SCIENCE HALL
NOTRE DAME, INDIANA 46556 -5670 USA
TELEPHONE (574) 631-7058 – FACSIMILE (574) 631-6652
EMAIL CHEMISTRY@ND.EDU – WEBSITE WWW.CHEMISTRY.ND.EDU

Evolutionary Cell Biology

A report of the findings of the NSF-sponsored Workshop on Evolutionary Cell Biology held May 29-June 1, 2012 at Airlie Conference Center, Warrenton, VA



This document is an outgrowth of discussions between participants who were physically at the workshop, and also those off-site, as contributed through the Evolutionary Cell Biology website (www.evolutionarycellbiology.org) and through a survey of more than 60 scientists conducted by the workshop organizers. These efforts were supported by grant MCB 1228570 from the National Science Foundation to the University of Notre Dame. Any opinions, findings, conclusions, or recommendations expressed in this document are those of the participants, and do not necessarily represent the official view, opinions, or policy of the National Science Foundation.

Cover art: Structure and distribution of centrioles, cilia, and the molecules involved in their assembly in eukaryotes. During evolution, centriole, cilia and associated genes have been lost in several eukaryotic lineages, such as yeasts, amoebas and higher plants. Correlating this pattern of gene loss with evolutionary relationships between organisms and structural differences between organelles provides testable hypotheses about the identity of genes involved in particular structures. **a)** Simplified taxonomic tree representing selected eukaryotic groups in different colors. **b**) Comparative centriole and cilia structures. **c**) Genes coding for centriole-assembly proteins are only present in the genomes of species that have those structures. White (gene absent); Black (gene present); grey (similar gene present). (modified from Carvalho-Santos et al. 2010).

Document number BIO 12-09

EXECUTIVE SUMMARY

Because cells are the fundamental units of life, two of the most significant questions in the biological sciences are "How do cells work?" and "How did the diversity of cellular structures and pathways arise?" The first question is the provenance of the field of cell biology, while the second, in principle, should be a central focus for evolutionary biology. Historically, cell and evolutionary biology have developed rather independently, with little exchange between the two. However, astonishing technological advances over the past decade raise the strong possibility that bringing these fields together will yield dramatic increases in our understanding of cell biological structures, functions, and processes, while also providing deep insights into the mechanisms of evolutionary change. The Workshop on Evolutionary Cell Biology (held May 29-June 1, 2012 at Airlie Conference Center outside Warrenton, VA) focused on the potential and requirements for advancing a field of Evolutionary Cell Biology. Twenty-four participants, bringing expertise in both cell and evolutionary biology, considered the key outstanding questions in Evolutionary Cell Biology (See Appendix) and envisioned an ambitious initiative anchored on an Atlas of the Biology of Cells (ABC).

The ABC would begin with a broad consideration of existing organisms based on phylogenetic considerations, enabling a rigorous sampling of cellular diversity. Comprehensive genomic, proteomic, and transcriptomic analysis of a large set of informative organisms would be complemented by ultrastructural analysis, so the combined data could be used to trace the gain and loss of major features, and to identify candidates for the molecular components thereof, in the full range of lineages. A critical aspect of the ABC would be the development of a subset of species as experimentally-tractable organisms (ETOs) amenable to critical approaches in cell biology, such as gene knockdown or knockout. The panel of ETOs would then allow testing of hypotheses about the assembly and function of specific structures in systems that are ideally suited for specific questions. Another transformative feature of the ABC will involve the collection of data on mutational and recombinational features as well as on parameters associated with the power of random genetic drift. An exciting anticipated outcome of this interdisciplinary venture is the potential for modeling the evolutionary forces underlying cellular features, which would then be tested in bench-top evolution scenarios.

The development of a field of Evolutionary Cell Biology will also involve initiatives aimed at fostering cross-disciplinary education, and the establishment and curatorship of novel databases. These long-term goals will require a combination of community initiatives as well as coordinated work by many individual labs. The promotion of many of these activities will require novel funding mechanisms.

WORKSHOP PARTICIPANTS

Monica Bettencourt-Dias	D. Allan Drummond	Eugene Koonin
Magdalena Bezanilla	Nels Elde	Michael Lynch
Dan Bolon	Cassandra Extavour	Harmit Malik
Zac Cande	Mark Field	Joanna Masel
Joel Dacks	Dan Fletcher	Dyche Mullins
Scott Dawson	Ursula Goodenough	Jose Pereira-Leal
Michael Desai	Holly Goodson	David Roos
Damien Devos	Greg Huber	Aaron Turkewitz

These participants were selected from list of >85 community members who expressed interest in attending, with an attempt to represent diversity in terms of area of research, organism of research, academic rank, size and geographic location of institution (including institutions outside the US), gender, and ethnicity.

INTRODUCTION

Remarkably, although molecular evolution, genome evolution, and even developmental evolution are now mature fields, there is no recognized field of Evolutionary Cell Biology. This glaring gap, which might be regarded as the last frontier of evolutionary biology, offers a signal opportunity for cell biologists and evolutionary biologists alike. A combined knowledge of cell biological diversity and functions and the evolutionary mechanisms that gave rise to them should help clarify the fundamental principles governing cell biological systems. And because all organisms are built from cells, the development of a field of Evolutionary Cell Biology (hereafter, ECB) should have a major impact on every dimension of the biological sciences.

Among many cell biologists, there is growing appreciation that evolutionary tools and perspectives have great utility in enhancing our understanding of cell biological processes. For example, phylogenetic analysis of protein families provides a powerful platform for rationally naming proteins, determining orthlog-paralog relationships that are essential for use of model systems in life-sciences research, and predicting functions of uncharacterized proteins. Moreover, the vast history of "nature's mutagenesis" stored in protein-sequence databases is a powerful but largely untapped resource for elucidating protein structure-function relationships (i.e., determining how proteins work). The combined knowledge of cell biological diversity and the evolutionary mechanisms and relationships that gave rise to this diversity should help clarify the fundamental principles that govern cell biological systems.

The potential payoff of developing a field of ECB is at least as significant for evolutionary biologists as it is for cell biologists. Among some evolutionary biologists, there is an emerging consensus that enhanced understanding of cell biology is required if evolutionary biology is to develop to its full potential – the establishment of a general theoretical framework based on principles that transcend species boundaries. Achieving this goal requires an understanding of how variations in cell structures and functions have arisen, and describing these in terms of the same variables that act at other levels of biological organization, i.e., population size, mutation rates, and recombination rates. We will also need to understand the relative contributions of adaptive vs. nonadaptive processes to cellular innovation. Moreover, the vast time span underlying the diversification of cellular features presents unique opportunities for identifying the fundamental physical and chemical principles and processes that have constrained and shaped the evolution of living systems.

By providing scientists with a rational basis for altering cellular features or even inventing new functions, the development of ECB will influence a diverse set of applied fields, including protein engineering, synthetic biology, and artificial life. Many of the goals of these emerging fields, which involve informed manipulation and design of cellular systems, are of major global significance, e.g., the development of efficient mechanisms for producing food, fiber, and biofuels. The establishment of bioremediation strategies for environmental restoration requires a deep understanding of variation in metabolic diversity among microbes. The evolutionary interplay between host and disease-causing organisms is a critical driver of developing strategies for fighting AIDS, influenza, malaria, and other emerging pathogens. Finally, cancer progression is an unfortunately spectacular example of cellular evolution in action, with a detailed understanding of somatic mutation being central to minimizing the consequences of this complex disorder.

To address the void in our evolutionary understanding of cell biological systems, NSF provided funding to bring together a small group of interactive scientists from a range of backgrounds to consider several fundamental questions. First, supposing that the time is ripe for the development of a field of ECB, what key questions need to be answered, and how should these be prioritized? Second, which areas of inquiry are accessible with current technologies, and what new technologies need to be developed to facilitate progress and remove barriers to

advancement? Finally, from the standpoints of methodology, community, and educational development, what programs and/or resources are required to move this emerging area of research forward in a timely, effective, and constructive manner?

A simple example of how NSF advanced a field of biology is given by the integration of evolutionary biology and developmental biology. Twenty years ago, work in both areas proceeded on largely independent pathways, with most references to evolution by developmental biologists being disconnected from known evolutionary mechanisms, and most references to development by evolutionary biologists ignoring the underlying details of cellular interactions. Now, thanks to a focused area of funding, an increasingly synthetic and demystified field of "evo-devo" has emerged as one of the most active areas of research on multicellular species. Given that cells are the building blocks of life, the argument can be made that this renaissance should have been preceded by the emergence of the field of Evolutionary Cell Biology.

WHAT IS EVOLUTIONARY CELL BIOLOGY?

Evolutionary cell biology is the study of patterns of variation in cellular features within and between species and of the mechanisms (molecular building blocks and population-genetic underpinnings) responsible for their establishment and maintenance. One aspect of ECB is the application of evolutionary perspectives and methodology to aid in elucidating the structure, function, and mechanisms of cellular processes. A second aspect is the study of cell biological diversity to gain insight into the mechanisms of evolution and the history of life on earth. For the field of ECB to have its full impact, both facets must be developed in tandem. Moreover, because cellular processes span the range from individual molecules to whole cells, ECB requires expertise from fields as diverse as biophysics and biochemistry to cytology and cell physiology.

While there is as yet no widely recognized field of Evolutionary Cell Biology (the phrase is found in fewer than ten abstracts in PubMed), scientists have applied evolutionary perspectives to cell biological problems for a long time. For example, in 1973 Pollard and Korn identified the muscle protein myosin in the single-celled *Acanthamoeba*, demonstrating that myosin is not unique to animal muscle cells. While myosin was certainly not the first protein found in both animals and protists, this paper was important in that it exploited a comparative study to simultaneously provide insight into the function of a protein (showing that myosin could operate outside the context of a muscle) and a cellular process (suggesting that amoeboid motility might involve myosin). Similarly, the early realization that humans and bacteria share biochemical pathways helped elucidate the biochemistry of both organisms and provided a unifying perspective for understanding metabolism.

Likewise, evolutionary biologists have long studied unicellular organisms in attempts to gain insight into the history of life and its diversification. For example, in 1980, Ford Doolittle used the term evolutionary cell biology to describe studies of the origin of eukaryotic cells, and a vigorous research community is actively focused on this problem. However, for the most part, molecular evolutionary studies simply evaluate the comparative features of linear protein sequences from diverse organisms. Such a strategy is a necessary first step to understanding the molecular basis of evolutionary change, but without a full understanding of the cellular contexts in which proteins evolve, confident interpretations of the mechanisms of sequence evolution and their phenotypic consequences are impossible. And without an understanding of how the population-genetic environment (the unique mix of forces of random genetic drift, recombination, and mutation in different lineages) defines the pathways that are open vs. closed to evolutionary exploitation, it is easy to become side-tracked by unrealistic evolutionary hypotheses. As in all aspects of evolution, selection operates on standing variation within

populations, whereas for the most part, cell biologists have embraced a small number of model systems explicitly designed to lack variation (other than manufactured mutants). An understanding of the genetic basis of within-species variation of cellular features is essential if ECB is to ever be fully embedded in the broader context of evolutionary biology.

In summary, a field of Evolutionary Cell Biology must encompass the application of comparative evolutionary perspectives and tools to the study of *diverse lineages of organisms* to tell us what evolution has produced. But ECB must also harness information on cell biological variation *within species* to tell us how evolution happens. Varying substantially in temporal scale, these two aspects of ECB are nevertheless united in their common utilization of cell biological diversity and in the synergism from coordinated consideration of how cell biological systems work and how such systems came to be.

SCIENTIFIC OBJECTIVES OF EVOLUTIONARY CELL BIOLOGY

Through a survey of more than 60 scientists (including workshop participants) prior to the workshop, a list of Key Questions in Evolutionary Cell Biology was generated and formed the basis for many of the workshop discussions (See Appendix). A general consensus was that a central goal of ECB is the generation of an organized, comparative perspective on the diversity of cellular features and processes within and among species of eubacteria, archaea, and eukaryotes. However, the ultimate desire is not simply a catalog of cellular diversity. Of equal importance is the establishment of common principles governing how cells and subcellular structures work, and development of an understanding of the evolutionary mechanisms responsible for the origin and maintenance of diversity at the morphological and metabolic levels. As discussed below, the meeting participants identified multiple ways to go about achieving these goals, concluding that the most broad-reaching, efficient, and significant avenue would be to undertake a large-scale and coordinated initiative that we have christened the Atlas of the Biology of Cells (ABC).

How Evolutionary Cell Biology can transform Cell Biology

The major goals of cell biology include: 1) the identification and characterization of subcellular structures and functions; 2) description of their underlying molecular components and mechanisms of assembly; and 3) elucidation of the complex processes by which these structures interact to generate cellular phenotypes, which in turn determine organismal responses to the environment. Recognizing the complexity of these issues, cell biologists have exploited a wide range of approaches including microscopy, molecular biology, biochemistry, genetics, biophysics and computational modeling.

However, because most cell biologists lack much rigorous education in evolutionary biology, a perspective common to most other areas of biology has been largely missing from the field. Moreover, most cell biological research has focused on a small group of model organisms, selected primarily for their accessibility to experimental interrogation or explicit similarity to mammalian cells rather than for their intrinsic interest in terms of biodiversity. Consequently, although many structures have been studied in great detail in mammalian cells and in model organisms such as the budding yeast *Saccharomyces cerevisiae*, any shared similarities between these taxa may have limited applicability to other eukaryotes because animals and fungi are members of just one of the major eukaryotic lineages. In other words, to the extent that they are exploited at all, the evolutionary perspectives that are possible with many of today's model species have limited power in the context of the broader Tree of Life.

It is time to change this situation. Spectacular advances in genome sequencing of the past decade have highlighted the vast diversity of the living world. If cell biologists can embrace and

exploit this diversity, and use the tools and perspectives of evolution as part of their arsenal of approaches, the pace of discovery should be greatly accelerated at all scales of inquiry:

- **What proteins do, and how they work**.

  At the molecular level, one of the most basic goals of cell biology and biochemistry is to determine the functions of proteins. Another central goal is to determine how proteins work. Classically, protein structure/function relationships are established through arduous experiments in which a range of systematically generated mutants are used to identify the parts of proteins that are most important. However, by utilizing the record of experiments that nature has already performed, i.e., by examining patterns of conservation in protein-sequence alignments, one can quickly identify the most invariant amino acids and obtain strong hypotheses about the location of the functionally significant regions without any bench-work. Indeed, by mapping patterns of conservation onto the surfaces of homology-modeled protein structures, one can often identify likely binding sites in proteins that have never themselves been biochemically characterized, even when the ligands (binding partners) themselves are unknown.

- **The functional integration of individual genetically encoded components in cells**.

  Insight into these issues derive from examination of variation at the level of structures, organelles, and processes. More specifically, by mapping complex features across the broad diversity of cellular life and determining patterns of gene and feature acquisition and loss, comparative biologists can identify the proteins and protein systems involved in particular features, structures, and processes. A striking example is provided by the study of cilia and centrioles across a range of organisms, as shown on the cover of this report and discussed more below (see the section "vignettes" for this and other examples).

- **The chemical and physical constraints on biological diversity**.

  Identification of differences between organisms provides information about which aspects of cell biology are malleable; and identification of unifying features of cellular life will help elucidate the chemical and physical constraints (such as the well-recognized surface area: volume ratios) that govern living systems.

A key requirement of many investigations into cellular structures and processes is data at the level of light and electron microscopy, behavior (e.g., mode of motility), and/or features such as metabolic activities. This means that realization of the full potential of ECB will require not just sequencing of an appropriately diverse set of organisms, but also a basic cell biological characterization of many of these same organisms. As with sequence data, these cell biological data will need to be accessible and comparable, which as described more below, will require extensive database development. Other important considerations in the development of ECB as noted by the workshop participants include:

- **Development of appropriately diverse "experimentally tractable organisms" (ETOs)**.

  Chosen from across the breadth of tree of life, the minimal requirement for an ETO might be as simple as culturability, but ideal ETOs would have features such as amenability to manipulation by tools such as transformation and RNA interference. Efforts to identify and develop an appropriate range of ETOs will need to coordinate and synergize with existing efforts in protistology and environmental microbiology.

• **Experimental testing of hypotheses generated from comparative analyses**.

Basic characterization of selected ETOs could be combined with approaches such as RNA-knockdown and GFP localizations to provide a first level of experimental testing of the hypotheses generated by the kinds of correlative analyses discussed above. These initial characterizations will provide the foundation for more in-depth analysis of specific structures, processes, and pathways, including guidance of researchers to organisms most appropriate for specific problems. Questions that could be addressed through such approaches include:

- How do cellular features (e.g., basal bodies, clathrin-coated vesicles, microtubules, nuclear pores, aspects of metabolism) vary within and among species, and what are the genetic origins of this variation?
- What is the minimal set of proteins required to generate, maintain, and regulate each of the specific structures or pathways in cells?
- Are there universal principles underlying common processes such as formation of the mitotic spindle, and if so, what are they? Or have different organisms solved the same problem in different ways, i.e., to what extent do shared features reflect common ancestry as opposed to parallel adaptation or physical constraints?
- What does the observed range of processes and mechanisms tell us about the origins of eukaryotic life or cellular life itself?

As noted above, though progress into these and similar questions of Evolutionary Cell Biology has been and will continue to be made by single-individual driven researcher, the meeting participants felt strongly that the most efficient progress and greatest advances would result from coordination of efforts in an ambitious initiative termed the Atlas of the Biology of Cells (ABC), discussed more below.

How Evolutionary Cell Biology can transform Evolutionary Biology

Today's field of evolutionary biology is well-grounded in a formal theoretical framework of population genetics, which provides a basis for understanding what is evolutionarily possible (and not possible) in lineages experiencing various strengths of mutation, recombination, random genetic drift, and selection. The central importance of this framework cannot be overstated, as any credible hypothesis about evolutionary change must be based on realistic genetic mechanisms. Evolution is a population-level process, with new adaptations arising from mutations that must navigate a long-history of stochastic processes before becoming established throughout a population. Numerous examples exist in which cellular improvements are inaccessible owing to internal population-level constraints (e.g., heterozygote inferiority in diploid species).

However, whereas the framework of population genetics provides strong guidance for the rapidity by which evolution can proceed in various contexts, the contexts themselves are defined by features of the biological world. Here we refer not only to the mutational and recombinational landscape in various taxa, but also to the specific molecular building blocks from which cellular features are built over evolutionary time by descent with modification. Much of the theory of population genetics is couched in very general terms, which is part of the power of this framework. However, the details of gene structure, protein architecture, and cell-biological contexts can no longer be ignored if we are to arrive at a fully mechanistic theory of evolution.

For example, most of the complex features of cells are defined by the properties of their constituent proteins and lipids. Thus, if we are to understand how cells evolve, we must understand how protein complexes emerge, e.g., how two previously noninteracting proteins come to participate in coevolving higher-order structures with novel functions. And if we are to understand this issue, we must start with a comparative analysis of the features of organisms with close enough relationships that the likely steps of divergence can be deciphered and, ideally, reconstructed and studied in an experimental laboratory setting. Only with closely related taxa is it possible to order the history of single-step mutations and their cumulative consequences for complex traits. Unfortunately, not only is most of today's cell biology restricted to just a tiny fraction of cellular diversity, but the few model species that have been adopted are so distantly diverged that there is no hope of confidently reconstructing shared ancestral states. Numerous examples of cell biological features are known where moderately related lineages deploy nonorthologous proteins for the same function, e.g., amino-acid synthesis (Hébert et al. 2011) and licensing of DNA replication origins (Drury and Diffley 2009). Thus, ECB would profit enormously from the development of a broad set of examples of cellular diversification at the level of reasonably closely related species, this being the reigning paradigm for studies of evolution at the level of external phenotypes and behavior.

Cell biology occupies a location in the hierarchy of life that is pivotal to our understanding of the mechanisms of evolution. Although many biologists view essentially every aspect of biodiversity as a product of natural selection, evolutionary geneticists have long known paths of evolutionary change that proceed with little involvement of Darwinian processes. Indeed, the further a biological feature is from the target of selection (the phenotype of the individual), the more likely it is to be influenced by nonadaptive mechanisms of evolution such as drift. For example, the diversification of a wide variety of genomic features (including introns and intergenic spacer DNA) among lineages appears to have arisen by differential forces of mutation and random genetic drift, which can sometimes completely overwhelm the power of selection (Lynch 2007a). Plausible arguments have been made that various aspects of cellular infrastructure may have also originated by effectively neutral mechanisms (Lynch 2007b; Lukeš et al. 2011). In addition, although it is easy to marvel at the bewildering array of features devoted to surveillance of internal cellular problems and their contribution to the robustness of organisms (e.g., DNA-replication proof-reading, decay of erroneous mRNAs, and chaperone guidance of protein folding), the case has been made that the establishment of layers of complexity need not have any long-term benefit (Frank 2007). One obvious disadvantage of a complex feature is that it is a larger target for mutational inactivation relative to a simpler feature carrying out the same task (Lynch 2007a).

Thus, a major challenge for evolutionary biology is to determine the extent to which the very infrastructure upon which organisms are built is driven by adaptive vs. nonadaptive processes, or combinations thereof. Resolution of these issues, which will require a synthesis of detailed comparative cell biology and evolutionary theory, will play a central role in the field of ECB if for no other reason than the fact that confidence in any adaptive arguments for the evolution of cellular features must remain suspect unless the hypothesis of effectively neutral evolution can be ruled out. This is not a trivial problem, as our understanding of comparative cell biology is so rudimentary that it is not even clear how a neutral theory of cellular evolution might be constructed. Nonetheless, the evidence for nonadaptive evolution is compelling. For example, numerous studies have demonstrated that the diversification of duplicate-gene function is often not due to the origin of new cellular functions but the simple partitioning of ancestral functions (Prince and Pickett 2002).

Finally, a key unresolved question at the heart of many ECB issues concerns the mechanisms that impose evolutionary limits on the levels of molecular perfection that can be achieved by natural selection. Although the argument is often made that selection is capable of pushing the refinements of molecular attributes until they meet the constraints imposed by

principles of physics and/or chemistry (e.g., Albery and Knowles 1976), it is also known that once an adaptation approaches a high level of refinement, further improvements can be blocked by the power of random genetic drift, with the level of imperfection being defined by the inverse of the effective population size (Hartl et al. 1985; Lynch 2011). Determining the situations in which a population-genetic process (drift) as opposed to a mechanical process (biophysics) constrains the adaptive evolution of cellular features will require the refined measures that cell biologists are often capable of achieving as well as estimates of the relative power of evolutionary forces that can be obtained through population-genetic study. Answers to questions such as these have obvious practical implications for applied studies that seek to improve cellular performance.

EXAMPLES OF THE POWER OF EVOLUTIONARY CELL BIOLOGY

Before proceeding with a description of the needs of ECB, we conclude this section with three brief vignettes of recent advances in cell biology (presented by workshop participants) that only became possible after the incorporation of an evolutionary perspective.

Discovery of a new adaptor protein. The modern eukaryotic cell is divided into distinct compartments by a complex organization of internal membranes consisting of lipid bilayers. This elaborate membrane system includes the Golgi apparatus, the endoplasmic reticulum (ER), and its continuity, the outer membrane of the nuclear envelope (NE). Movement of proteins and lipids between these compartments is critical for normal cellular functions, and such transport is mediated in part by a set of protein complexes called "adaptins." It had been accepted for over a decade that only four adaptin complexes exist in eukaryotic cells, and that these complexes are extremely ancient, arising in eukaryotes over a billion years ago. Although an additional adaptin-like protein had been suggested in humans, this had been dismissed and never characterized. However, after versions of the disputed protein were found (by sequence analysis) to be conserved in organisms across the span of eukaryotes (e.g., plants and amoebae), it was deduced that a fifth adaptor protein with key functions must exist (Hirst et al. 2011). Subsequent characterization of the protein in human cells identified a set of interacting partners, a cellular location, and a function. Thus, by adopting an evolutionary cell biological approach, a fifth adaptin complex was discovered, altering our basic understanding of how eukaryotic cells function, and in turn leading to a better understanding of how the transport system in cells has evolved over the past two billion years.

An evolutionary link between coated vesicles and the nuclear pore complex. Traffic between the ER, Golgi, and the cell membrane is carried out by three kinds of cargo-carrying, membrane-bound vesicles, each of which is surrounded by a different set of coat proteins: 1) clathrin/adaptin complexes are responsible for endocytosis (from the plasma membrane), and vesicular trafficking between the Golgi, lysosomes, and endosomes; 2) COPI complexes mediate intra-Golgi and Golgi-to-ER trafficking; and 3) COPII complexes support vesicle movement from the ER to the Golgi. In addition, the NE is perforated by nuclear pore complexes (NPCs), which form channels between the cytoplasm and nucleoplasm and stabilize the highly curved membranes that surround them. The three coated-vesicle complexes and the NPC are composed of multi-protein complexes, a central role of which is to stabilize curved membranes. Moreover, comparative structural work suggests that, despite their diverse roles in the cell and absence of sequence similarity, all four structures share a common molecular architecture, likely reflecting an ancient common origin (Devos et al. 2004).

This observation has had several important implications for ECB. First, comparative studies led to structural and functional information that was critical to developing a mechanistic

understanding of the nuclear pore complex (Devos et al. 2006). Second, it can now be concluded that the last eukaryotic common ancestor (LECA) possessed ancestral versions of membrane-curving complexes that duplicated and diverged to produce the diverse systems in modern eukaryotes. These inferences lead, in turn, to a novel evolutionary proposal – the "protocoatomer" hypothesis, which postulates that the NPC and vesicle-coating complexes arose by descent with modification (Devos et al. 2004). Third, by extension, these observations suggest the testable hypothesis that assembly/disassembly mechanisms might also be conserved between complexes with divergent functions. Because such mechanisms are better understood for coated vesicles like clathrin than for the NPC, it is tantalizing to speculate that the knowledge obtained from coated vesicles can be transferred to the NPC. Thus, this series of studies clearly shows how an evolutionary perspective can lead not only to important gain of knowledge in an area of broad significance for cell biology, but also to the design of a rational research agenda for resolving key unanswered questions.

The diversification of centrioles and cilia. Centrioles are microtubule-based cylinders that reside within centrosomes and can give rise to cilia. Both centrioles and cilia are involved in diverse functions ranging from cell motility to cell division. Centrosome defects are seen in many cancers, and abnormalities in cilia lead to numerous diseases including polycystic kidneys and infertility. Centrioles are found in most eukaryotic groups, implying their presence in the LECA, with secondary loss occurring in specific branches such as yeasts and higher plants. The distinctive features of these structures suggest unique assembly machinery that should only be present in the genomes of species that assemble cilia.

This hypothesis was successfully tested by several groups, and lead to the identification of novel components of centrioles and cilia, many of which were found subsequently to be involved in human disease. Using comparative genomics, an evolutionarily inferred, ancestral molecular module associated with centrioles and cilia has been proposed (Carvalho-Santos et al. 2011; Avidor-Reiss et al. 2004; Li et al. 2004; Carvalho-Santos et al. 2010).

Additional studies of the phylogenetic profiles of other players in centriole formation suggested that the coordination of centriole and cilia biogenesis and function in different cellular contexts is achieved by tissue-specific molecular innovations, gained through duplication and divergence of an ancestral gene set (Carvalho-Santos et al. 2010; see also Cover Art for this Report). Work is now being carried out to extend these approaches to the diversity of structures encountered within centrioles and cilia and correlating them with molecular components. Such work has required the development of new tools such as a web-based platform with a novel controlled vocabulary that integrates molecular and morphological data (including many decades of previously published electron microscopy studies) in an evolutionary context.

A RESEARCH AGENDA FOR EVOLUTIONARY CELL BIOLOGY

The central goals of ECB are to establish a deep understanding of:

- The cellular features of all of the major lineages across the Tree of Life;
- The range of structures, processes, and mechanisms that allow these varied cells to function, including the common principles that apply in diverse organisms, and conversely, the varied ways in which different organisms solve the same problem;
- The evolutionary (population-genetic) mechanisms that give rise to this variation;
- The elemental (molecular) building blocks from which evolution proceeds.

The workshop participants felt that some important questions can be pursued now with existing tools and even existing data sets (the vignettes give some examples), but a longer-term vision is essential if the field is to move forward in an efficient and effective manner. The consensus was that the greatest impact, by far, would come initially from a coordinated effort to catalog the full range of cell biological diversity across the Tree of Life. Such cataloging would include features shared by all lineages as well as those specific to individual lineages, and almost certainly would lead to the identification of new cellular features. We will start with an overview of the heart of the long-term research agenda proposed for this area, the Atlas of the Biology of Cells (ABC) plan, and then outline several key developments that will need to be pursued en route to the overall goal. Although the ABC may seem overly ambitious, the same was said about the Human Genome Project when first proposed over a decade ago, and yet today we are far beyond anything imagined at the time of the initial proposal, with all areas of biology profiting from the resultant knowledge base and technology development.

The Atlas of the Biology of Cells (ABC):

Achieving the goals of Evolutionary Cell Biology as outlined above will require a coordinated effort to sample and characterize cell biological diversity at all scales, meaning that it will be necessary to both determine the range of cell biological diversity as it occurs across the full tree of life and the detailed structure of the variation observed within particular branches. In other words, we must characterize both selected members of all major groups (both eukaryotic and prokaryotic) and specified sets of closely related species. We will call the resulting compendium of knowledge the Atlas of the Biology of Cells (ABC).

Considerations in setting up the ABC project:

- **Critical to the establishment of ECB is the availability of an accurate description of the Tree of Life.** Thanks to substantial work by the systematics community, data for the development of this framework are already abundant, but much of eukaryotic and prokaryotic diversity remains unsampled. Therefore, the ABC initiative will need to be carefully coordinated with existing projects such as the Tree of Life (ToL) so that the genome sequences and evolutionary relationships of informative organisms can be determined. It is important to clarify that the ABC and ToL efforts are complementary rather than redundant: while ToL focuses largely on sequence-based information and analysis, a major goal the ABC effort will be to connect this sequence information to data on cellular structures and processes.

- **Achieving the ultimate goal of a full description of cellular features will require the identification and development of a phylogenetically diverse set of experimentally tractable organisms (ETOs).** Minimally, ETOs will need to be amenable to laboratory culture and accessible to genetic manipulations such as siRNA and transformation. Although numerous existing model organisms already meet these criteria, a large fraction of the Tree of Life is strongly or completely under-represented in current laboratory research, with many phylogenetic lineages being represented only by pathogens with extraordinarily derived genomes and cellular attributes. The investment of time and resources in developing even a single new model organism is enormous, and it is acknowledged that there may (and should) be considerable debate among cell biologists as to which branches constitute the major lineages of life and which species within each lineage provide the greatest potential for the development of novel methodologies. As examples of such community efforts, after more than two decades by

more than 30 research groups, two protists in the Alveolate lineage, *Toxoplasma gondii* and *Tetrahymena thermophila*, have become powerful model systems for cell biological research.

- **ABC taxa will need to be culturable in multiple environments.** Because organisms have evolved in temporally and spatially varying environments, which are generally substantially different than optimal laboratory growth conditions, a full understanding of the functions of various cellular features will need to be explored in the context of a range of realistic conditions. Among other things, such investigation will also enable investigators to define the morphospace of species, i.e., the ranges of phenotypes that can be expressed without a change in genotype.

- **The ABC will not only require sampling of species across the whole Tree of Life, but also deeper sampling on selected branches.** To the extent possible, reliance on single "reference" species per major lineage should be avoided. Even closely related taxa can differ dramatically in genome content, and evolutionary analysis relies on observations about standing variation within and among species. Indeed, while many cell biological questions may be best-addressed by comparing a diverse range of major lineages, because evolution is a population-level process, an understanding of the likely evolutionary mechanisms driving such diversity will generally require studies on variation among closely related organisms.

- **The ABC will necessitate descriptions of the full range of cellular features for each candidate lineage**. The initial objectives will be to collect information at the genomic, transcriptomic, and proteomic levels, complemented by detailed ultrastructural (electron and light microscope) analysis. When relevant, separate analyses should be conducted for all accessible developmental stages. For selected ETOs with well-developed toolkits, these approaches will be complemented by more in-depth analysis such as siRNA and GFP-fusion libraries as has been done for model organisms such as *Arabidopsis* and budding yeast. Experience with current model systems can serve as a filter for the types of surveys/methodologies that are most informative.

- **A community-level effort will be required to formulate detailed plans.** Although the guidelines provided above represent the broad and preliminary outline arrived at by the meeting participants, more detailed plans for ABC and prioritization of ETO-focused projects will have to be established through future discussions and meetings to foster communication between all interested parties. As discussed further below, one mechanism to support such phase would be the Research Communication Network program already in place at NSF.

Expected outcomes of the ABC project:

- The trove of data that is expected to result from ABC, itself of immediate value to cell and evolutionary biologists, is an essential starting point for: 1) understanding the function of each of the parts and their involvement in functional modules; 2) obtaining descriptions of the physical structure of each protein subunit and assembly mechanisms of higher-order complexes; and 3) procuring a dynamic description of temporal patterns of gene expression, subcellular localization, and compartmentalization.

- Given the resultant organized comparative data bases of the ABC, the connections that will have been drawn between proteins and cell parts, and a well-described structure for

the Tree of Life, it will be possible to determine the phylogenetic positions of lineage-specific gains and losses of key metabolic and structural innovations of cells, as well as to ascertain the temporal positions of gene-duplication/loss events and points of accelerated evolution of the relevant cellular components. For example, the ABC database will allow one to ascertain the fates of gene duplicates in terms of subfunctionalization and neofunctionalization, as well as to infer other key events such as horizontal gene transfer. It will also be possible to determine whether common cellular attributes in distant lineages have arisen by shared descent or by parallel gains of their underlying constituents, and if so, whether the same molecular mechanisms have been involved.

- Although the preceding analyses will provide an atlas of cellular features and their structural underpinnings, thereby telling us what has evolved, comparative biology is rather silent on the underlying evolutionary mechanisms promoting change. Understanding at this level requires quantification of population-genetic features (the power of mutation, recombination, and random genetic drift), which together constrain the paths that are open to evolutionary exploitation on various lineages. For the molecular spectra of mutation and recombination, which are known to vary by orders of magnitude among species, direct estimates are achievable via mutation-accumulation and, for sexual organisms, meiotic-crossing experiments provide information on recombination rates. Indirect estimates of the power of drift are also obtainable with existing methodologies using population-genomic data. Although this kind of work can only establish the features of modern-day species, elucidating generalities (or lack thereof) is critical to determining how the "ground rules" of evolution are manifested across the Tree of Life. For example, recent theoretical work demonstrates that the paths by which complex traits (involving multiple mutations to achieve a final end state) become established (or are prevented from establishing) is very much dependent on the population-genetic environment (Lynch 2007b; Weissman et al. 2010). Thus, a broad and quantitative understanding of how the power of drift, mutation, and recombination varies across the Tree of Life has the potential to move ideas about the origin of cellular features beyond the realm of pure speculation.

- Studies using the ABC database will lead to new hypotheses about the evolution of cellular features. One way to test those hypotheses will be to reconstruct ancestral states, using molecular-genetic approaches in favorable model systems. A second useful approach may be the application of selection on specific cellular features in organisms with short enough generation times to make it possible to detect evolution on reasonable time scales. With the short generation times of most unicellular species and emerging opportunities in nanotechnology and microfluidics, enormous potential exists for experimental approaches to cellular evolution, e.g., targeted selection applied to cellular features in populations of cells (analogous to selection for phenotypic extremes in economically important species of plants and animals).Given a large population of cells, for example, if prolonged selection were to be imposed for particular aspects of cellular infrastructure (metabolic or structural), what would the response be, and it would vary among replicate populations?

Other ECB Research Initiatives:

While the workshop participants felt that the greatest progress by far would come through establishment of the ABC, which will require considerable resources, many aspects of evolutionary cell biology can be pursued with existing sequence databases and experimental

systems. The preceding vignettes provide snapshots of how the evolutionary perspective has been used to gain insight into cell biological problems. Other examples include the use of phylogenetic information to identify previously uncharacterized actin-related proteins (ARPs) as likely chromatin remodelers (Blessing et al. 2006), the use of conservation mapping to predict ligand binding sites by mapping of sequence constraint onto protein surfaces (Ashkenazy et al. 2010), or even (more broadly) the identification of simple physically-dictated mechanisms as a way to explain both the mechanism and origin of complex cell biological processes (e.g., Meyers et al. 2006; James and Vale 2012). Issues of the population-genetic mechanisms that might promote the evolution of oligomeric structures of proteins, network topologies, and complex cellular features are now being explored (Weissman et al. 2010; Lynch 2007b, 2012), although these need to be tied to specific biological examples.

MAJOR SCIENTICIFIC AREAS THAT WILL BE IMPACTED BY AN ECB INITIATIVE

Developed through close coordination with ongoing efforts such as the Tree of Life projects, by fostering a quantum leap in our understanding of the structure and function of cellular machineries, the ABC research agenda is expected to transform every field of the life sciences. Because most of life's diversity has a cellular basis, the ABC will amount not only to the establishment of an atlas for all of biology, but also to a detailed understanding of where the atlas came from in an evolutionary sense. The resultant knowledge base will further provide the basis for a fully integrative field of evolutionary biology, firmly informed by both population-genetic and cell-biological mechanisms. And finally, the ABC will constitute a permanent set of resources for a wide array of research communities. A few examples of the anticipated breadth of influence of ABC follow:

- **The relationship between protein structure and function**, initially informed by approaches such as mapping patterns of conservation onto 3D surfaces, followed up by empirical analysis. Work of this nature is fundamental to the fields of protein engineering and biotechnology, whose goals include harnessing the observed range of biochemical diversity to enhance the design of biomaterials and biofuels.

- **The identification of proteins associated with particular structural modules and processes**, ascertained in part by correlating losses and gains of structures and processes with losses and gains of particular genes and whole modules, capitalizing on the spectacular advances in sequencing to connect genes to structures, phenotypes, and behaviors. While enhancing our understanding of well-known processes, such efforts are also likely to reveal a range of new enzymes and new biological processes.

- **The elucidation of the chemical and physical constraints that govern living systems**, accomplished in part by determining which aspects of cellular life are shared by diverse organisms, but also guided by principles of biochemistry and biophysics. Applications of such understanding will extend to the field of exobiology, providing insights into how physical and chemical constraints may have influenced the emergence of life elsewhere in the universe.

- **The history of life on earth**, not simply as discerned through the genealogical relationships of organisms, but with the additional integration of information on the diversification of cell biological features and the underlying proteins, structures, and processes.

- **The integration of evolutionary genetic theory with specific molecular-level changes**. Medicine and agriculture are among the fields that will profit from such understanding, in areas as diverse as the control of pathogens to the development of improved systems for food and fiber production.

- **The establishment of a rational field of synthetic biology**, informed by an understanding of the identity and evolution of naturally occurring cellular processes in the pursuit of the design and manipulation of artificial cellular systems and/or drug design.

FUNDING MECHANISMS

Assembling and exploiting the ABC will necessarily be a composite effort of many individual laboratories, and will therefore require novel coordination, collaboration, and establishment of core facilities and databases. Here, we provide an outline of how the individual initiatives required for the development of ECB might be organized on the basis of funding that will be required. A subset of the initiatives can, in principle, begin immediately, requiring minimum to modest additional resources for their early stages. Other initiatives might largely be supported through existing funding mechanisms, primarily individual or collaborative awards, while a third important group will clearly require the development of novel funding mechanisms distinct from individual awards. We also indicate the approximate time scales for beginning each of these initiatives. The accomplishment of some goals will require further development of key methodologies and model systems. In addition, since many of these efforts will require the coordination of research groups in multiple countries, it is important that international participation be accommodated or even encouraged through joint programs or other mechanisms.

Exploiting Existing Resources

Education/community building. To publicize and promote this general ECB initiative, attendees will be writing brief meeting reports to cell and evolutionary biology journals, and will also organize cross-disciplinary sessions at prominent cell and evolutionary biology meetings. For example, a Member-Organized Special Interest Subgroup session on ECB will be held at the annual meeting of the American Society for Cell Biology in December 2012.

Coordination with the Tree of Life project. Efforts to expand and refine the phylogenetic assignments of the taxa in the Tree of Life are currently underway, e.g., the NSF-sponsored initiative on Assembling, Visualizing, and Analyzing the Tree of Life (AVAToL). One of the main foci of ECB will be eukaryotic microbes, a polyphyletic group constituting most of eukaryotic evolutionary diversity that has been only sparsely sampled by genome sequencing projects. As a consequence, the detailed phylogenetic assignments of many eukaryotic microbes are not yet secure. Coordination of the phylogenetic aspects of the ABC initiative with the ToL is essential to ensure that the tools and data developed by each group are complementary and shared resources rather than redundant efforts. For example, as the ABC develops experimental tools for a wide set of single-celled organisms, the relevant catalog of tools for each species and the data resulting from their application should be incorporated into an annotated ToL, enhancing the overall utility of this resource to experimental cell and evolutionary biologists. This process should be initiated in the near future to summarize information on existing tools for the complete range of experimental organisms.

Encouraging the further development of ECB. Organizing meeting symposia, publishing short synopses of such meetings, and the like, are examples of cost-free mechanisms for advancing ECB that can encourage the development of a large community of researchers with expertise on widely divergent organisms. One mechanism for encouraging communication and participation is EVOLUTIONARYCELLBIOLOGY.ORG website that was established for the NSF-sponsored workshop. Prior to the workshop, short questionnaires were sent to a group of approximately 85 cell and evolutionary biologists to learn their opinions on the most important questions in ECB and to determine their interest in participating in the workshop. The number of positive responses far exceeded the number of people that could be accommodated at the workshop, stimulating the establishment of this interactive web site for "virtual participants," in an effort to provide a forum for communication within the larger community.

Pursuing Individual or Collaborative Grants

In order for existing funding mechanisms (e.g., grants to individuals or small consortia) to contribute to the development of ECB, funding agencies will need to aggressively work to inform the community of their interest in supporting ECB-related work, by website postings and more direct outreach to the ECB community. The main issue here is that ECB crosses traditional programmatic boundaries within most life-sciences funding agencies. Thus, a slight broadening of conventional funding mechanisms could have an important influence on ECB. For example, the estimation of the strengths of drift, mutation, and recombination in phylogenetically diverse organisms would shed light on how the population-genetic environment, within which cellular evolution occurs, varies across the Tree of Life.

One key to developing a broad phylogenetic basis for understanding ECB will be the establishment of a diverse collection of experimentally tractable organisms (ETOs), well beyond the currently narrow range of model organisms upon which most molecular and cellular biologists are now focused. To be successful, this kind of pursuit may require a special funding mechanism for small groups of interacting investigators. In any event, it is clear that much can be done to advance ECB, before a fully organized ABC initiative is established, by simply changing the culture among investigators attempting to understand the cellular basis of evolution and the evolutionary basis of cellular diversification.

Comprehensive databases already exist for model organisms such as yeast, fly, and Arabidopsis; these repositories include DNA sequence data along with a large array of other information including EST libraries, protein-protein interaction data and protein localization data curated from the literature. These could serve as models for establishing databases for newly characterized ETOs. There are also numerous organism-specific databases that contain primarily nucleic acid sequences and are, for the most part, extensions of individual or small collective projects that are typically narrow in scope. Individual researchers funded by conventional grants will keep developing these tools in response to their own needs. To serve the needs of ECB, the content of these existing databases would have to be greatly expanded to include information relevant to cell biology. One major challenge for the broader scientific community and ECB in particular will be the establishment of mechanisms to make all of the information in these resources comparable, stable, reliable and easily accessible. However, such such enterprises will only be successful if coupled with rigorous evaluation of quality, impact, and cost-efficiency.

Many organism-specific databases already exist for model organisms such as yeast, fly, and nematode. These academic efforts are, for the most part, extensions of individual or small collective projects and typically narrow in scope. Individual researchers funded by conventional grants will keep developing these tools in response to their own needs. However, a major challenge for ECB will be the establishment of mechanisms for expanding these classical databases to incorporate a much larger set of ETOs. Such resources will need to evolve into

stable and reliable information repositories, but such enterprises will only be successful if coupled with rigorous evaluation of quality, impact, and cost-efficiency.

The ABC initiative represents a clear example of a general phenomenon in today's biological research, namely the enormous increase in data generation resulting from emerging technologies. One of the major challenges of this 'omics' paradigm is the establishment of data-analysis methods capable of matching the scale of data acquisition. From the standpoint of ECB, an early step that could be taken in this direction would be the pursuit of NSF funded Research Coordination Networks (RCNs) to organize groups for developing and maintaining databases relevant to the creation of synergies within ECB, and for establishing standards, evaluation metrics, etc.

A Need for New Funding Mechanisms

Although substantial progress in ECB might be made through traditional funding pathways and by dovetailing with other large-scale initiatives, the field is sufficiently novel that few current investigators have the expertise to integrate ideas from cell biology and evolutionary biology. Thus, if the field is to move forward in a significant way, mechanisms for encouraging the merging of these technically demanding fields will need to be promoted.

- A large part of the initial experimental work essential to the ABC will require specialized funding mechanisms, in large part because such research does not neatly correspond to "hypothesis-driven research". Examples include testing a wide variety of organisms, judged to occupy highly informative positions, for their ability to be cultured in the laboratory. It must be understood at the outset that many of these attempts are likely to fail, so steady initial funding will be essential to find the subset of useful species. For such purposes, the set of organisms currently available via the American Type Culture Collection (ATCC) is an obvious short-term starting point, as culture methods (although perhaps not optimal) are already established. Similarly, developing a set of core functional tools and approaches for a large set of ETOs will require a sustained effort supported by dedicated funding. Both types of research will almost certainly require funding to communities of interacting investigators. For example, in order to develop an understanding of the evolutionary features of a particular pathway or structure, a mechanism that facilitates collaborations between researchers working on the same structure or pathway, but in different model systems/ETOs, would be imperative.

- Given the interdisciplinary nature of ECB, and the paucity of established investigators crossing the disciplines of evolution and cell biology, one of the most critical needs of the field exists at the educational level. Funds should be available for cross-disciplinary training at all levels (e.g., graduate student training grants, postdoctoral fellowships, and sabbatical support for faculty). Such funding would be most likely to bear fruit when allocated to situations in which the participant already has established expertise in one of the subdisciplines (e.g., a postdoctoral fellow with a cell biology background moving into a lab with expertise in evolutionary theory). Because individuals receiving such training will be poised to make the most seminal, founding contributions to this nascent field, the optimal situation will be the allocation of resources at early points in awardees' careers, as this will help ensure the further training of the next cohort of students at little extra cost.

- Database curation, ideally in a framework that transcends species boundaries, will be critical for ECB, distinct from maintaining the research-driven databases discussed above. The large databases envisioned would be analogous to the enormous

contribution made by the NCBI repositories for genetics, genomics, and many other areas of biology (e.g., Pubchem). The resources for establishing and curating large ECB databases will need to be provided outside of the limitations of investigator-driven proposals (just as the NCBI repositories are open to the general public in a standardized manner).

- Similarly, the establishment and maintenance of collections of mutants, morpholino/RNAi collections, antibodies, vectors, etc., will be central to the development of each ETO. The engagement of companies to collaborate in the development and distribution of these reagents must be sought, but the small scale of the communities focused on each ETO may preclude this option in many cases. Thus, a mechanism that ensures the sustainable maintenance and accessibility of ETO-specific reagents must be developed.

Research Funding Timescales

Immediate Research Coordination Network **(**RCN) grants would enable the ECB community to meet and set priorities, lay groundwork for coordination with the Tree of Life project, and begin identifying candidates for ETOs. Early funding should then be made available to begin basic analysis of potential ETOs (genomic, proteomic, and/or transcriptomic). Ultimately, the ABC will need to be funded by a large project-type grant or series of such grants.

LITERATURE CITED

Albery, W. J., and J. R. Knowles. 1976. Evolution of enzyme function and the development of catalytic efficiency. Biochemistry 15: 5631-5640.

Avidor-Reiss, T., A. M. Maer, E. Koundakjian, A. Polyanovsky, T. Keil, S. Subramaniam, and C. S. Zuker. 2004. Decoding cilia function: defining specialized genes required for compartmentalized cilia biogenesis. Cell 117: 527-539.

Blessing, C. A., G. T. Ugrinova, and H. V. Goodson. 2004. Actin and ARPs: action in the nucleus. Trends Cell Biol. 14: 435-442.

Carvalho-Santos, Z., J. Azimzadeh, J. B. Pereira-Leal, and M. Bettencourt-Dias. 2011. Evolution: tracing the origins of centrioles, cilia, and flagella. J. Cell Biol. 2011. 194: 165-175.

Carvalho-Santos, Z., P. Machado, P. Branco, F. Tavares-Cadete, A. Rodrigues-Martins, J. B. Pereira-Leal, and M. Bettencourt-Dias. 2010. Stepwise evolution of the centriole-assembly pathway. J. Cell Sci. 123: 1414-1426.

Devos, D., S. Dokudovskaya, F. Alber, R. Williams, B. T. Chait, A. Sali, and M. P. Rout. 2004. Components of coated vesicles and nuclear pore complexes share a common molecular architecture. PLoS Biol. 2: e380.

Devos, D., S. Dokudovskaya, R. Williams, F. Alber, N. Eswar, B. T. Chait, M. P. Rout, and A. Sali. 2006. Simple fold composition and modular architecture of the nuclear pore complex. Proc. Natl. Acad. Sci. USA 103: 2172-2177.

Doolittle, W. F. 1980. Revolutionary concepts in evolutionary cell biology. Trends Biohem. Sci. 5: 146-148.

Drury, L. S., and L. F. Diffley. 2009. Factors affecting the diversity of DNA replication licensing control in eukaryotes. Curr. Biol. 19: 530-535.

Frank, S. A. 2007. Maladaptation and the paradox of robustness in evolution. PLoS One 2: e1021.

Hartl, D. L., D. E. Dykhuizen, and A. M. Dean. 1985. Limits of adaptation: the evolution of selective neutrality. Genetics 111: 655-674.

Hébert, A., S. Casaregola, and J. M. Beckerich. 2011. Biodiversity in sulfur metabolism in hemiascomycetous yeasts. FEMS Yeast Res. 11: 366-378.

Hirst, J., L. D. Barlow, G. C. Francisco, D. A. Sahlender, M. N. Seaman, J. B. Dacks, and M. S. Robinson. 2011. The fifth adaptor protein complex. PLoS Biol. 9: e1001170.

James ,J. R., and R. D. Vale. 2012. Biophysical mechanism of T-cell receptor triggering in a reconstituted system. Nature 487: 64-69.

Li, J. B., J. M. Gerdes, C. J. Haycraft, Y. Fan, T. M. Teslovich,. et al. 2004. Comparative genomics identifies a flagellar and basal body proteome that includes the BBS5 human disease gene. Cell 117: 541-552.

Lukeš, J., J. M. Archibald, P. J. Keeling, W. F. Doolittle, and M. W. Gray. 2011. How a neutral evolutionary ratchet can build cellular complexity. IUBMB Life 63: 528-537.

Lynch, M. 2007a. The Origins of Genome Architecture. Sinauer Assocs., Inc. Sunderland, MA.

Lynch, M. 2007b. The evolution of genetic networks by non-adaptive processes. Nat. Rev. Genet. 8: 803-813.

Lynch, M. 2011. The lower bound to the evolution of mutation rates. Genome Biol. Evol. 3: 1107-1118.

Lynch, M. 2012. The evolution of multimeric protein assemblages. Mol. Biol. Evol. 29: 1353-1366.

Meyers, J., J. Craig, and D. J. Odde. 2006. Potential for control of signaling pathways via cell size and shape. Curr. Biol. 16: 1685-1693.

Pollard, T. D., and E. D. Korn. 1973. *Acanthamoeba* myosin. I. Isolation from *Acanthamoeba castellanii* of an enzyme similar to muscle myosin. J. Biol. Chem. 248: 4682-4690.

Prince, V. E., and F. B. Pickett. 2002. Splitting pairs: the diverging fates of duplicated genes. Nat. Rev. Genet. 3: 827-837.

Weissman, D. B., M. W. Feldman, and D. S. Fisher. 2010. The rate of fitness-valley crossing in sexual populations. Genetics 186: 1389-1410.

Appendix:

Final Meeting Agenda

Workshop on Evolutionary Cell Biology
May 29-June 1, 2012
Airlie Conference Center, Warrenton VA

Evolutionary Cell Biology
Workshop May 29-June 1 2012

Workshop sponsored by the National Science Foundation (Award MCB - 1228570 to explore issues of common interest to the cell biology and evolutionary biology communities, identify and consider the key questions and challenges and the approaches necessary to move the field forward. Several NSF program officers will attend the workshop as observers, to learn about emerging areas and to answer questions.

The questions below were selected from participant suggestions, and the foci of several sessions will be determined during discussions at the meeting.

Arrival:

- Arrive afternoon of Tuesday May 29 into Dulles (preferable) or National (alternative, with longer travel to the meeting)
- Attendees can either take a taxi or a shuttle. Participants should send arrival and contact information so that we can coordinate ground transport to the meetin

Schedule:
Meeting will start with Tuesday evening dinner 6:00-7:00, and end Friday mornin with breakfast

Day 1 (May 29): Laying the foundations

Dinner 6:00-7:00

Session 0: "Finding Common Ground" 7:15 - 10:00pm

7:15-7:30 Welcome and Introduction: Holly Goodson and Aaron Turkewitz
7:30- 8:15 The relevance of evolutionary theory to cell biology: Mike Lynch
8:15-8:45 Discussion
8:45-9:30 Cell biological diversity: Zac Cande
9:30-10:00 Discussion

Reception 10:00pm

• Continue discussion over drinks/snacks (cash bar until midnight

Day 2 (May 30): Identifying Key Questions in Evolutionary Cell Biology

Breakfast 7:30-8:30

9:00-10:20 Session 1: How do physical constraints influence the evolution of cells?

- What physical constraints are relevant to cell biological evolution?
- How have they limited cell biological structures and functions?
- What are the roles of physical constraints in convergent evolution?
- How can these issues be addressed experimentally?

Introduction and Discussion Leader: Holly Goodson
9:00-9:15: Introduction
9:15-10:00 Small group discussions
10:00-10:20 Reassemble: small groups report back
10:20-10:40 Discussion as group

10:40-12:20 Session 2: What is the significance of adaptive vs. nonadaptive mechanisms for cellular evolution?

- Are the limits of molecular perfection dictated by the power of genetic drift in different lineages, or by physical / chemical barriers?
- Is the mutational cost of increased complexity an issue in the evolution of cellular features?
- How can these issues be addressed experimentally?

Introduction and Discussion Leader: Mike Lynch
10:40-10:55 Introduction
10:55-11:40 Small group discussions
11:40-12:00 Reassemble: small groups report back
12:00-12.20 Discussion as group

Lunch: 12:20-1.30

1:30-2:50 Session 3: What are the molecular mechanisms leading to the origins of cell biological features?

- How do novel protein activities arise?
- How do novel cell biological structures arise?
- What determines whether a new mutation goes to fixation?
- How can these issues be addressed experimentally

Introduction and Discussion Leader: Michael Desai
1:30-1:45: Introduction
1:45-2:30 Small group discussions
2:30-2:50 Reassemble: small groups report back
2:50-3:10 Discussion as group

3:15-4:20 Session 4: How can studying the evolution of proteins and cell biological systems provide insight into their present-day function and mechanism?

- How can the array of information in the sequence databases be put togethe with knowledge of evolutionary relationships to gain insight into protein structure, function, and dynamics?
- How can comparative cell biology be used to identify fundamental cell biological mechanisms?

Introduction and discussion leader: TBA
3:15-3:30 Introduction
3:30-3:50 Small group discussions
3:50-4:10 Reassemble: small groups report back
4:10-4:30 Discussion as a group

4:30-5:00 Writing break

- Moderators finish up and post discussion summaries
- Participants post comments, questions, and ideas.

5:00-5:45: Discussion of topics to be addressed during Day 3 – additional "big questions" of evolutionary cell biology?

Dinner 6:30-7:30

Informal discussions at the Pub

Day 3 (May 31):
1) More Key Questions
2) What is needed to move the field forward?

Breakfast 7:30 -8:30

9:00-10:25 Session 5: Parallel question discussions –(Moderator chosen at meeting)
Questions that individual groups can choose to discuss include:

- What would it take to create a hypothesis-driven evolutionary cell biology?
- What have been the major challenges to cells, and how does this impac evolutionary cell biology?
- Is there any predictability to cell biological evolution?

9:00-9:10 Introduction to topics in session
9:10-9:55 Small group discussions
9:55-10:25 Reassemble: small groups report back

10:25-11:45 Session 6: Focus TBA. *Either one question to be considered by all groups or a series to choose from, as seems appropriate from the progress of the meetin*
10:30-11:00 Small group discussions
11:00-11:30 Reassemble: small groups report back
11:30-12:00 Group Discussion

Lunch: 12:00-1:00

1:15-2:30 Session 7: Moderator: TBA
Choose one of these two questions:

- What new tools/technologies/infrastructure are needed to advance Evolutionary Cell Biology?
- What changes to graduate, undergraduate, and post-graduate education are needed to push Evolutionary Cell Biology forward?

1:15-2:00 Small group discussions
2:00-2:30 Reassemble: small groups report back

2:30-3:50 Session 8: Moderator: TBA
Choose one of these two questions:

- What types of funding mechanisms would be most useful for advancing the field of Evolutionary Cell Biology?
- What types of communication/community infrastructure/meetings would advance the field of Evolutionary Cell Biology?

2:30-2:45 Introduction to topics
2:45-3:30 Small group discussions
3:30-3:50 Reassemble: small groups report back

4:00-5:45 Session 9: to be chosen based on topics that arise during meeting
4:00-4:45 Small group discussions
4:45-5:15 Reassemble, small groups report back
5:15-5:45 Group discussion

5:45-6:15 Writing Break

- Moderators finish up and post discussion summaries
- Participants post comments, questions, ideas

Dinner 6:30-7:30

7:45-8.30 Workshop Wrap-up
Summarize:

- Key Questions
- What is needed to move the field forwards

Reception 9:00pm -11:00pm
• Continue discussion over drinks/snacks (cash bar at reception until 11:00 pm those who desire can continue at the pub)

Day 4: Departure

Breakfast 7:00 -8:30