Nova Acta Leopoldina NF 96, Nr. 357, XXX–XXX (2009)

Integrative Computational Approaches to Complex Ecophysiological Systems

Andreas Bohn (Oeiras, Portugal)

With 6 Figures

Abstract

The present work highlights the application of integrative computational tools in ecophysiological studies. With the example of circadian rhythms in Crassulacean acid metabolism, modeling approaches for the integration of diverse levels of biological organization, as well as different time- and space scales are assessed. The integration of heterogeneous data sources is discussed with the case of a web-based computational infrastructure in a multinational, multidisciplinary project on phototrophic biofilms. Both examples underline the importance of aligning the respective scientific cultures and communication forms of the project partners for the successful application of computational techniques in the research on complex biological systems.

Zusammenfassung

Die vorliegende Arbeit behandelt die Anwendung integrativer Computerwerkzeuge in ökophysiologischen Studien. Am Beispiel circadianer Rhythmen im Crassulaceen-Säurestoffwechsel werden Modellierungsansätze zur Integration verschiedener Ebenen biologischer Organisation sowie verschiedener Zeit- und Längenskalen erörtert. Die Integration heterogener Datenquellen wird anhand der Fallstudie einer internetbasierten Dateninfrastruktur im Rahmen eines multinationalen, multidisziplinären Projekts über phototrophe Biofilme diskutiert. Beide Beispiele unterstreichen die Bedeutung der wechselseitigen Abstimmung der wissenschaftlichen Kulturen und Kommunikationsformen der Projektpartner für die erfolgreiche Anwendung von computerbasierten Techniken in der Erforschung komplexer biologischer Systeme.

1. Introduction

In the second half of the 20th century, technological breakthroughs in biochemistry, imaging and information processing have triggered a tremendous increase in the generation of information about the constituent parts of living organisms. Until the full sequencing of the human genome, the predominant approach to understanding the complex nature of biological systems followed a reductionist paradigm, relating organismal functionality and dynamics to the activity of individual molecules (Keller 2005, Hütt and Lüttge 2005). The evidence that the functionality or disease of entire biological systems could not be explained by deciphering solely the letters of the 'Book of Life' (Noble 2003), lead to a massive paradigmatic change in life sciences and the emergence of a large number of novel research approaches, attempting to unravel how whole-organismic function surges from the interactions between the parts of the system (Busch and Eils 2005). One of the most prominent approaches taking a systemwide perspective on life has been entitled systems biology (Ideker et al. 2001, Kitano

2001). Since its beginnings, this field has attracted scientists of many different disciplines, from experimental biology over engineering to computer sciences and physics (Keller 2005). Despite the existing lack of a clear definition of what exactly constitutes systems biology, common elements to most suggestions for a definition include the quantitative modeling of biological systems across different levels of organization, the integration of heterogeneous data sets, and the interdisciplinary networking of experimental biologists and quantitatively trained scientists (Morris et al. 2005).

The inherently quantitative character of systems biology, together with the traditionally strong connection between molecular high-throughput studies and bioinformatics, has founded an implicit tendency to understand systems biology as the genome-wide, or generally 'ome-wide', study of cellular and organismal function emerging from the interactions of its molecular parts (Westerhoff and Palsson 2004). Yet, it has been suggested that in studies of multicellular organisms and their environmental interactions, e.g., in crop development, a dialectic between bottom-up and top-down approaches provides a more efficient approach to biocomplexity and biotechnological developments, than hierarchical, unidirectional advances, working from the molecular level up to higher levels of organization (Hammer et al. 2004). These arguments gain even more weight if one understands systems biology as a truly holistic attempt to integrate all space-scales of the biosphere from molecules to ecosystems.

Bridging the entire spectrum of scales will also increase the spectrum of computational tools to be applied. While molecular studies are mainly challenged by the large volumes of information to be processed, quantitative ecological studies also face a bewildering variety of data types, sources and logical structures which need to be integrated (Jones et al. 2006). Situated between the molecular and the ecological scale, computational ecophysiology is about to become an interesting meeting point of bottom-up and top-down approaches in full-scale systems biology.

By means of two case studies, the present work discusses the implementation of the basic elements of systems biology in ecophysiological studies. Section 2 discusses multilevel modeling with the example of circadian rhythms of whole-leaf gas exchange in a Crassulacean acid metabolism (CAM) plant. It is demonstrated how the development of logically connected models with differing degrees of abstraction can elucidate the connection of dynamical processes at the cellular and organismal level, and enhance related experimental studies. Section 2.3 draws some general conclusions on multilevel modeling approaches in environmental physiology. The second example, presented in Section 3, highlights the integration and analysis of heterogeneous data sources in ecophysiology, generated in the context of a multinational, multidisciplinary project on phototrophic biofilms. It stresses the importance of metadata management to balance flexibility with consistence in data integration, and the importance of matching the applied analysis tools with the size and structure of the given data pool. Section 3.4 discusses general aspects in the creation of effective scientific data workflows and underlines the importance of the interdisciplinary communication between the involved partners. The conclusions on both examples are integrated in Section 4, discussing the necessity to complement new technologies for knowledge discovery and hypothesis testing with cultural and sociological advances in networked team science, to promote successful interdisciplinary research in biocomplexity.

2. Circadian Rhythms in Crassulacean Acid Metabolism: Integrating Data and Hypotheses on Different Levels of Organization

2.1 Chronobiology of CAM

Cyclic, oscillatory dynamics are deeply entrenched into the temporal organization of living organisms. An elusive example is the adaptation to geophysical cycles, in particular the 24 h-cycle of day and night (Pittendrigh 1993). These co-called circadian rhythms are ubiquitously observed in plants, animals and microorganisms, and there is increasing evidence that the coordinated timing driven by endogenous, circadian clocks enhances organismal fitness (Paranjpe and Sharma 2005, Dodd et al. 2005). A prominent example for how the temporal organization of metabolic processes by a circadian system can provide ecological advantages to plants is Crassulacean acid metabolism (CAM), an adaptation of plants to drought stress (Black and Osmond 2003, Lüttge 2004, and references therein): Governed by an endogenous circadian system, the uptake of CO_2 from the environment is shifted to occur predominantly at nighttime. The temporal separation of CO_2 uptake from its fixation and storage as starch via the light-dependent C_3 pathway, allows the use of the internal CO_2 store accumulated during the night. Diurnal photosynthesis can then take place behind closed stomata during the hottest and driest phase of the day, yielding an overall improvement of water-use efficiency.

Over the last two decades, two principal hypotheses about the origin of the endogenous oscillations in the carbon metabolism of CAM plants have been proposed. One is based on a molecular feedback system which hierarchically drives metabolic rhythmicity by modulating the activity of key enzymes in CAM carbon metabolism (Hartwell 2005, and references therein). The second approach features a biophysical pacemaker localized at the the vacuolar membrane, the tonoplast. The principal feedback mechanism is based on the nonlinear interdependence of the efflux rate of vacuolar malic acid, the principal store for nocturnally acquired CO_2 and the order of the vacuolar membrane (Lüttge 2000). While the former hypothesis to date has not been modeled in a quantitative fashion, the latter mechanism was subject to an ongoing iteration of experimental studies and quantitative multi-level modeling.

2.2 Modeling CAM Rhythmicity: From Single- to Multi-Oscillator Systems and Back

The first quantitative model of CAM, based on the experimental knowledge available at that time was presented by Nungesser et al. (1984). By interdisciplinary collaboration between engineers and botanists, a computational model was developed, featuring 6 ordinary differential equations (ODEs), representing 6 metabolic pools, interacting by first order reaction and regulation terms. Already in that model, the principal point of impact for environmental parameters like light intensity was the transport of malic acid at the tonoplast. As described in detail by Lüttge (2000), this model evolved in several steps in alignment with surging experimental evidences. The hitherto final point in the evolution of cellular CAM models was reached with the model by Blasius et al. (1999), quantifying the mentioned nonlinear interdependency of the efflux of vacuolar malic acid and its level of accumulation.

A principal merit of this model is the representation of the conditionality of the CAM cycle in continuous light: here, the circadian cycle is arrested in steady states with a filled vacuole at low temperatures, and an empty vacuole when the plant is exposed to high temperatures (Grams et al. 1997). Starting to lower the temperature from the latter arrested state, the model predicts the onset of circadian oscillations once the temperature crosses the bifurcation threshold, independent on the rate of temperature change. This prediction was contradicted by experiments by Rascher et al. (1998): rhythm re-initiation could only be observed experimentally in response to fast temperature changes, while a slow transition between the two temperature regimes maintained the gas-exchange cycle arrested in the arrhythmic state.

This finding induced a fundamental change in the modeling approach to CAM rhythms. The experimental results by Rascher et al. (1998) became interpretable by considering populations of several copies of the CAM model (Blasius et al. 1999) with an additional noise term (Beck et al. 2001). Taking into account the multi-cellular nature of the measured whole-leaf gas exchange, and the stochastic dynamics of the oscillations emerging from omnipresent noise in real systems, rhythm re-initiation after a fast temperature transition could be understood as the synchronization of a population of noisy oscillators by a quick common transition of all oscillators from a fixed-point to a limit-cycle regime. Slow temperature transitions would yield an onset of oscillations in each individual oscillator, however, phase desynchronization of the population would maintain the arrhythmic global signal. In addition to that, this multi-level approach gave rise to interpret macroscopic arrhythmicity and rhythm damping as a noise-induced loss of phase coherence among the microscopic oscillating elements of the system. The success of this multi-oscillator model, introducing the “clockshop” hypothesis (Winfree 1975) to CAM rhythms, lead to a new experimental approach to whole-leaf rhythmicity. By implementing a chlorophyll fluorescence imaging facility, image sequences of photosynthetic efficiency could be recorded to assess the spatio-temporal metabolic dynamics in CAM leaves in day-night cycles and in continuous light (Rascher et al. 2001). This new technique unraveled a significant amount of spatio-temporal heterogeneity in the CAM leaves in day-night cycles and continuous light conditions, which was a new and unexpected result for a physiologically and anatomically homogeneous leaf.

By taking a pixelwise time-series approach, Bohn (2003) analyzed the spatio-temporal data quantitatively and compared them with numerical simulations of populations of uncoupled oscillators. As the hypotheses about whole-leaf rhythm generation to be tested were concerned with purely dynamical processes, not involving biophysical details of rhythm generation, the CAM model by Blasius could be further reduced in complexity and degrees of freedom. For the given purposes it can be substituted by the FitzHugh-Nagumo (FHN) model, a generic, phenomenological model for excitatory behavior and limit cycle oscillations, consisting of two differential equations (Keener and Sneyd 1998), whose dynamical properties fully represent the CAM model. Arrested circadian cycles in continuous conditions had been modeled earlier for rhythms observed in insects (Peterson 1980) and fungi (Gooch et al. 1994), using the so-called displaced limit-cycle (DLC) model. As this model also incorporates rhythm damping at the level of a single oscillator, it was suited to create an alternative hypothesis to the suggested loss of phase coherence in the population exhibited in the FHN model, whose single elements show either oscillations with unattenuated amplitude, or no oscillation at all. Figure 1 exhibits simulations of the spatiotemporal dynamics of uncoupled ensembles of both FHN (left panels) and DLC (right panels) systems. Underneath the space-time plots in the top row, the temporal development of the spatial average of each population and a representative time-series of a single oscillator is given. The middle panels of Figure 1 show the corresponding time-series from a data set derived from the chlorophyll image sequences by Rascher et al. (2001). From visual inspection, one may conclude that

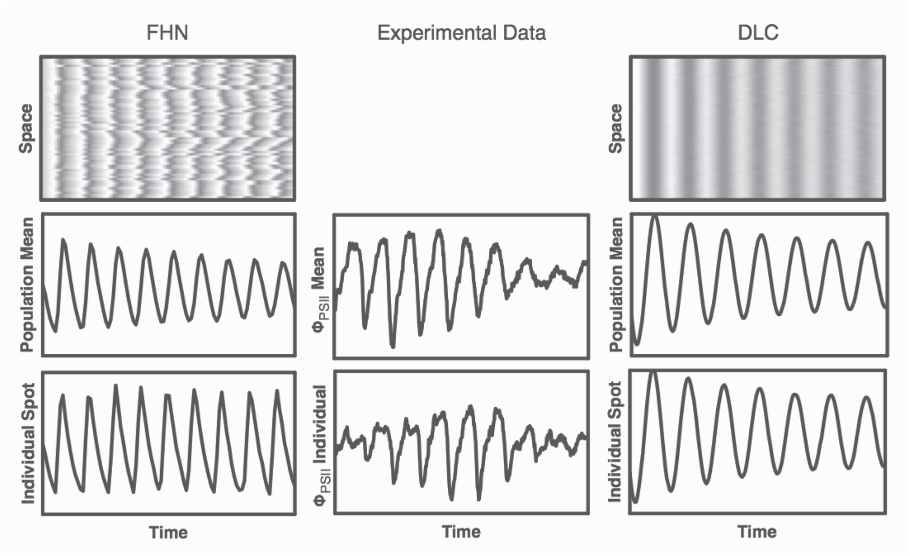

Fig. 1 Comparison of two hypotheses on the amplitude attenuation of overt circadian rhythmicity in the CAM plant *Kalanchoe daigremontiana*. *Top row*: Spatiotemporal dynamics of 100 uncoupled oscillators. *Middle row*: Temporal dynamics of the spatial average of the population. Bottom row: typical temporal dynamics of a single oscillator. *Left column*: FitzHugh-Nagumo (FHN) model (Keener and Sneyd 1998). *Middle column*: Quantum efficiency of photosystem II, data from Rascher et al. (2001). *Right column*: Simulations of the Displaced Limit-Cycle (DLC) model by (Peterson 1980)

the global dynamics of leaf metabolism is correlated to the local dynamics, which favors the hypothesis represented by the DLC model. This has been further substantiated by quantitative statistical analyses (Bohn 2003), suggesting that the origins of amplitude attenuation and arrhythmicity should emerge from the interactions of the biophysical and chemical entities at the cellular level.

The pixelwise time-series analysis applied to the mentioned image sequence of photosynthetic efficiency furthermore revealed synchronized patches in the leaf, which could be related to a slight 24h-modulation of light intensity (Bohn 2003). This observation pointed to a dynamical phenomenon which had not been studied in depth in fundamental nonlinear dynamics: the spatio-temporal dynamics of a population of spatially arranged oscillators under the influence of an external periodic driver, which acts with a spatially heterogeneous amplitude on the oscillators (Bohn and García-Ojalvo 2008). The central question of biological interest in this system is if and which type of inter-oscillator coupling could bring the entire array into synchronization with itself and the environmental driver, in spite of the heterogeneous impact of the latter on the array. As synchronization of non-chaotic oscillators is characterized by phase differences between oscillators, rather than amplitude deviations (Pikovsky et al. 2001), the additional complexity introduced by coupling was balanced by considering phase oscillators, a minimal model for oscillatory processes, which consists of one single differential equation describing the evolution of the phase of the system (Acebrón et al. 2005). The spatio-temporal plots in Figure 2 show the principal effects of two different types of coupling on a heterogeneously driven array. The left panel shows the dynamics of an uncoupled array under a driver which has its peak amplitude in the center of the array: this leads to the synchronization of the oscillators in a center strip, while at the margin, where the external driver strength drops below the critical entrainment strength, oscillators show independent free-running dynamics, resulting from their individual period mismatch with the external force (see Bohn and García-Ojalvo 2008 for details). Applying global coupling, i.e. each oscillator adjusting its own phase in dependence on its difference to the average of all phases in the array (Fig. 2, top right panel), the entire array becomes synchronized both internally, as well as to the external driver. Applying local, next-neighbor coupling (Fig. 2, lower right panel), global synchronization is replaced by running phase waves in the non-synchronized zone in the uncoupled state. Hence, under heterogeneous external driving, local coupling provides order in the array, however it is less efficient in comparison to global coupling. This is in general compliance with the theory of coupled phase oscillators (Acebrón et al. 2005).

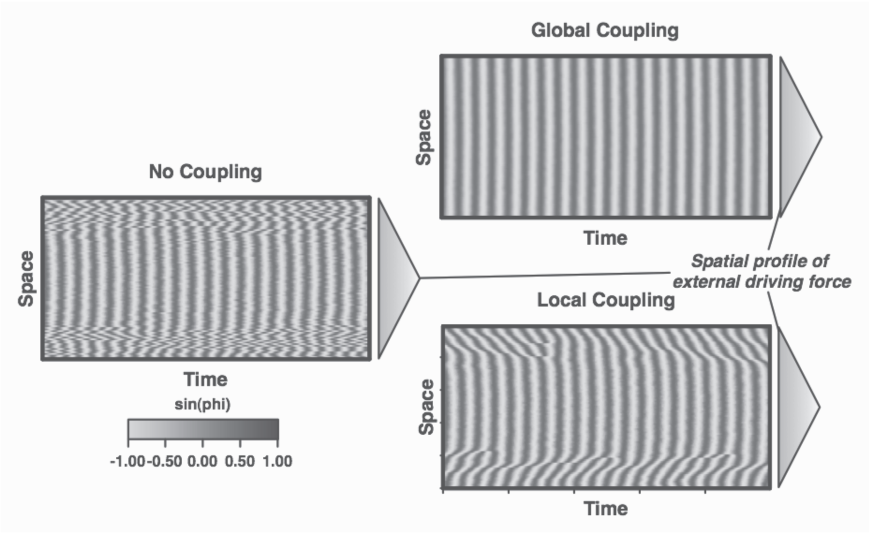

Fig. 2 The effect of different coupling types on the synchronization of 100 phase oscillators (Bohn and García-Ojalvo 2008). *Left*: spatiotemporal dynamics of 100 uncoupled oscillators. *Top right*: 100 globally coupled oscillators. *Bottom right*: 100 locally coupled oscillators.

An advantage of the generic nature of the used model is the ease of comparison of rhythmic phenomena in various organisms. Besides the present case of spatio-temporal CAM rhythmicity, global coupling by the medium surrounding the cell population has also been indicated as the main synchronization mechanism between circadian rhythms in populations of unicellular algae (Broda et al. 1985) and neurons in the suprachiasmatic nucleus, the central mammalian pacemakes (Gonze et al. 2005). This suggests that if circadian oscillators are indeed coupled across cell boundaries, the coupling agent must be quickly diffusing throughout the entire system on time-scale much faster than the 24h-period of the oscillation. Taking this hypothesis into account in future research on intercellular coupling agents of circadian rhythms could catalyze the screening of possible molecular candidates and hence guide the design of future experiments.

In the concrete case of circadian CAM oscillations, this hypothesis might inspire the design of prospective experiments. Assuming CO_2 to be the intercellular coupling agent as proposed by Duarte et al. (2005), suggest that fully synchronized leaves in the presence of heterogeneous environmental signals could be easier achieved in plants with low internal CO_2 diffusion resistance. The succulence of CAM leaves and the resulting low internal CO_2 conductivity could thus be an important factor for the emergence of running phase waves, as phase adjustments by CO_2 signaling might be limited to cells in close vicinity, corresponding to the numerical scenario of local next-neighbor coupling (Fig. 2). This hypothesis could be tested by exposing CAM plants with different internal CO_2 conductivities to a light source with controlled temporal modulation of its intensity and a controlled spatial geometry of light incidence on the leaf.

Returning from the organismal to the cellular scale, the results of the presented numerical studies using phenomenological models could also trigger novel modeling approaches to the biochemical bases of CAM rhythmicity. The present mechanistic model should be extended in order to include rhythm attenuation in continuous conditions at the cellular level, to provide a mechanistic background to the phenomenological DLC model. As the comparison of simulations exposing the model by Blasius et al. (1999) to CO_2-free epochs with corresponding experiments suggest, a second oscillator robust to severe metabolic perturbations must be involved in the generation of overt CAM rhythm generation (Wyka et al. 2004). Hence, a new model at the cellular level is likely to be a multi-level model also, integrating the metabolic dynamics with circadian gene expression cycles. Given the similarity of the molecular CAM clock with the molecular *Arabidopsis* oscillator (Hartwell 2005), the gene expression cycles might be represented by a model similar to the one developed by Locke et al. (2006) for *Arabidopsis.* As this class of model gives a very detailed description of the involved molecular feedback loops, it might contain an excess of complexity to the study of the interaction between gene expression and metabolism, which should be a bidirectional connection, as in some cases, metabolic signals have been shown to override periodic gene expression cycles in CAM plants (Borland et al. 1999). It is thus suggested that future numerical studies of the multi-level nature of cellular CAM rhythms might – at least initially – rely on models with a reduced number of variables. The complexities at both levels could be reduced in analogy to Forger and Kronauer (2002), who explicitly showed the mathematical equivalence of the two-dimensional van-der-Pol oscillator, a phenomenological model for nonlinear oscillations, with the five dimensional model of circadian gene expression by Goldbeter (1995).

2.3 Modeling in Systems Biology: Integration across Organisms and Levels of Organization

Integrated understanding of the behavior of living organisms in their natural environments requires the integration of diverse levels of organization. Successful system models of entire organisms, e.g. the human heart, suggest that the most efficient way to achieve multilevel understanding is not by a hierarchical, unidirectional modeling attempt, be it bottom-up or top-down, but by starting at those intermediate levels where sufficient data are available (middle-out approach, Noble 2003). Hybrid multilevel models, relying on the flexible handling of the model granularity in function of the data availability, might reconcile inductive and deductive concepts in systems biology (Coveney and Fowler 2005). In practice, it is both the failure as well as the success of a given model that can advance the knowledge on a given system, as long as modeling activity remains logically connected to experimental work (Noble 2002). It is in this sense, that the modeling of CAM rhythms, using an entire spectrum of models with different levels of detailedness has guided and inspired corresponding experimental endeavors for more than two decades.

Perceiving quantitative models as navigation tools in tackling the multilevel complexity of living organisms (Hammer et al. 2006), and as being "no more, but no less, than a way of thinking clearly" (May 2004), could be a promising stance to foster the acceptance of a heterogeneous landscape of coexisting models of a given organism. In the same fashion that complex functions of living organism emerge from the interaction of many diverse parts on different levels, the knowledge of this complexity might rather emerge from a heterogeneous network of models, than form one single, optimally designed model. The success of such a "modelomics" approach requires new technologies to facilitate the connection of diverse models. Modeling meta-languages such as the Systems Biology Markup Language SBML (http://www.sbml.org), which serve as common descriptor for a large spectrum of models, are one example. However, as will be further discussed in the concluding Section 4 of this work, to be fully efficient, these technological advances need to be accompanied by an increasing awareness of the socio-psychological challenges to interdisciplinary team research in life sciences.

3. Phototrophic Biofilms: Integrating Data Sources, Analysis Tools and Scientific Activity

3.1 The PHOBIA Project: Diverse Perspectives on Phototrophic Biofilms

Most hard substrates in nature are covered with biofilms, aggregations of microorganisms encapsulated in a protective and adhesive matrix. They are increasingly recognized as the preferred mode of growth of microbes in a wide range of habitats (Costerton et al. 1987, Stoodley et al. 2002). Research on biofilms has increased to a great amount over the last decades. They play a central role in many pathogenic processes in biomedicine (Parsek and Singh 2003), and cause significant damage to technical processes and transportation through biocorrosion and biofouling (Coetser and Cloete 2005). On the other hand, biofilms bear a large potential for the development of novel biotechnologies, e.g. for wastewater treatment (Rittmann 2006).

Phototrophic biofilms, mixed cultures of hetero- and autotrophic organisms (Fig. 3), are crucial elements of aquatic ecosystems and are prospective points of departure for novel environmental biotechnologies (e.g. Sansone et al. 1998, van Dam et al. 2002, Nagarkar et al. 2004). Here, phototrophic organisms, e.g. microalgae and cyanobacteria, fuel heterotrophic bacteria and fungi, which in return provide nutrients for the phototrophs. Both groups exude extracellular polymeric substances (EPS), which provide surface adhesion of the biofilm and contribute to its protection.

The project PHOBIA was a first large-scale integrative approach to phototrophic biofilms, joining 6 laboratories in 5 European countries (Fig. 4, and http:// www.photobiofilms.org). All partners used identical freshwater and marine inocula gained from the same sampling

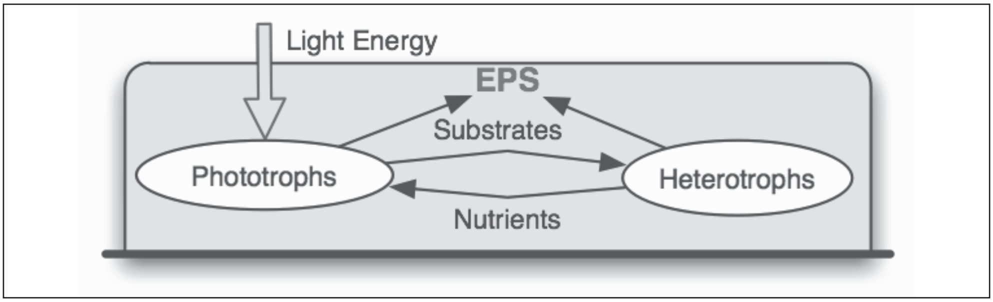

Fig. 3 Scheme of interactions between autotrophic and heterotropic organisms in phototrophic biofilms. EPS stands for “Extracellular Polymeric Substances” (see text)

sites, as well as the same incubator and protocols to assess the effects of temperature, light intensity and water flow velocity (Zippel et al. 2007). Each partner then analyzed the biofilms in different developmental stages with the specific tools of their expertise (Fig. 4). As is depicted in the following, the quantitative approach to integrate these heterogeneous and geographically dispersed data sources was based on a web-based computational infrastructure, WebPHOBIA, incorporating a data warehouse for the reposition of the data generated in the diverse laboratories, connected to a computational module based on Artificial Neural Networks for data modeling and variable selection.

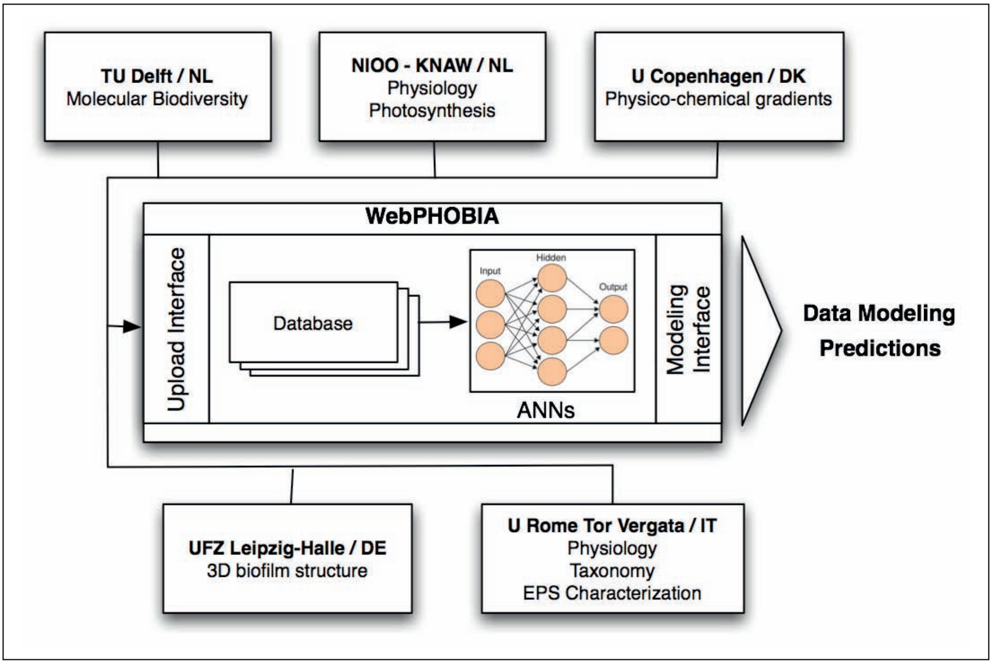

Fig. 4 Scheme of the data flux in the PHOBIA project and the composition of the WebPHOBIA data management facility.

3.2 Integration of Heterogeneous Data Sources by Data Warehousing

One approach to integrate heterogeneous data sources is data warehousing (Schönbach et al. 2000). It involves the translation of the data from the diverse sources to a central repository, on which the querying and data mining is performed. Like for other models as well, the data model on which a data warehouse is based can be developed in either a bottom-up or a top-down fashion: the former consists in integrating diverse local data repositories (data marts) into a single facility, i.e. queries submitted to the central data warehouse are translated and conveyed to the local repositories. In the top-down approach, the structure of the warehouse is designed beforehand, and all generated data in the laboratories are submitted directly to the warehouse, which also serves as the site to which further processing modules are deployed.

Inspired by an earlier web-based infrastructure developed for another European-wide research project in microbiology (Silva et al. 2003), the data warehouse in WebPHOBIA followed the top-down approach. Each PHOBIA partner communicated the structure of their data in the startup meeting of the project. Based on this knowledge, a database was designed and implemented using the relational database management system (RDBMS) PostgreSQL. A software engine was programmed in PHP to provide the user interface for data upload and querying, and to connect the database with the analysis module.

While the top-down approach based on a relational database provides a high level of consistency in the integration of the diverse sources, two of its disadvantages became prominent in the course of PHOBIA. First, changes to the data structure, due to additional experimental endeavors, cannot be accomplished without exporting the already submitted data and resubmitting them to the rebuild structure. The second drawback involves the dependence of a data warehouse relying on a complex relational database model on the knowledge of the domain expert, i.e. the database creator and programmer. A career change of the WebPHOBIA domain expert before the completion of data submission and the end of the project made a significant amount of the contextual information needed to understand and use the data repository unavailable, which significantly hampered the implementation of the necessary changes to the data structure and the completion of the database.

Both these issues relate to the importance of metadata, i.e. data about data, in integrative computational approaches in ecology and life sciences (Michener et al. 1997). Managing metadata is thus considered a key strategy in the development of data integration tools, which allow incremental modifications of the structure of the base, without compromising the already uploaded data, and grant transparency to its users without having to unravel complex data structures at the level of RDBMS. For the implementation of metadata management, and to create controlled vocabularies and ontologies for data integration in life sciences in the long run, the use of semantic web technologies has been proposed (Cheung et al. 2005, Jones et al. 2006).

A prototype software for data integration using semantic concepts is S3DB, the Simple Sloppy Semantic Database (Almeida et al. 2006). This framework refrains from pre-defining a fixed database structure, in which the data model is coded in the relationships between the tables of the database. Data are rather considered as resources with given properties, and are labeled with a Uniform Resource Identifier (URI). In analogy with the Resource Description Framework (RDF), the data are related and organized as triples [Resource][Property][Property Attribute]. The backbone of S3DB is a relational database, however, its tables contain the resources, the URIs, properties and values, and the data structure is not implicitly mapped in the entity relationships. The S3DB engine (built in PHP) checks for the consistency of the semantic data structure. The resulting data model, i.e. the network of relations between the different resources and their properties, are visualized by the S3DB interface, and can be easily modified (Fig. 5).

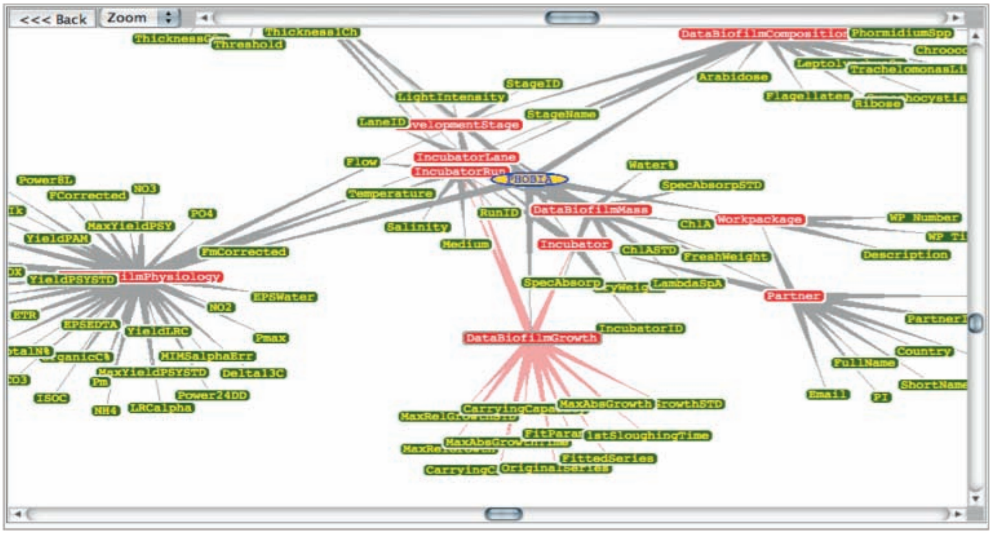

Fig. 5 Screenshot of the graphical interface to the S3DB database in the PHOBIA project, showing the semantic network.

Apart from the theoretical aspects of data integration using ontologies and controlled vocabularies, the practical advantages of applying the semantic approach mainly consist in the transparency of the metadata, and the fact that – once deployed to the webserver – no coding at the level of the RDBMS is needed to adjust the database to changes in the data model. Semantic approaches like S3DB represent thus a step towards providing data integrity across the sources to be integrated (like the top-down approach), yet maintaining some of the flexibility of the bottom-up approaches. The tradeoff for these advantages is the reduced computational efficiency and speed in query processing as compared to a taylor-made relational database.

3.3 Data Analysis and Modeling with Artificial Neural Networks

The second module of the web-based PHOBIA platform contained a modeling and analysis infrastructure relying on Artificial Neural Networks (Anns, Basheer and Hajmeer 2000, Fig. 4). The application of this non-linear data analysis tool was inspired by its prior successful and widespread application in environmental biotechnology and biochemical engineering (Almeida 2002). They represent a machine learning tool which can establish a functional relation between a set of independent (input) variables and a set of dependent (output) variables. In analogy to the neural learning process, consisting in adjusting synapsis strength in the network of neurons, the learning process determines the number of neurons in the hidden layer, and adjusts the weights of the connections between the neurons in the diverse layers, such that for empirically determined input values x, the prediction error of the ANN output with respect to the measured output variables is minimized. For the training process, the full data set is divided into a training part, used to adjust the weights, a testing part, from which the the prediction error is tested, and a validation part to avoid overfitting (Almeida 2002). This completely data-driven approach bears the advantage of not needing any beforehand mechanistic assumption for data modeling and prediction.

In WebPHOBIA, ANNs were deployed to perform two tasks. First, through a special interface for predictive modeling in the website, users could select any pairs of input and output variables among the quantities contained in the database. By computations based on algorithms programmed in Matlab, ANNs would then be trained with the data available in the base and then serve as predictor of the output values resulting from the combination of input values chosen by the user. The objective of the second task was to detect the most important relations between the variables in the different repositories, in order generate mechanistic knowledge on biofilm development and structure. Therefore the sensitivity of an given output variable to a certain input can be estimated via the relation

$$S_{ij} = \frac{\Delta Y_j}{Y_j} \Big/ \frac{\Delta X_i}{X_i} \qquad [1]$$

The resulting sensitivity patterns can then be integrated with mechanistic knowledge in order to create testable hypotheses on the relations between the observed biofilm variables. Given the above issues of the database, ANN analyses were carried out manually using data from spreadsheets as input form. An example spreadsheet containing a predictive model for biofilm growth can be downloaded at (http://phobia.itqb.unl.pt/ann.php).

Unfortunately the quality of the data used for the analysis indicated a loss of robustness of the obtained results. As the example of the dependence of a growth parameter (days needed to reach 50 % of the carrying capacity) on temperature shown in Fig. 6 indicates, the data used to train the ANN do not share equal variances and contain outliers. Furthermore, the division of the total data set into a training, test and validation subsets in ANN training, requires a large number of cases to produce reliable results. The number of available data in PHOBIA was close to the lower limit where the application of ANNs can be considered to be reliable. Hence, in analogy to the S3DB database, the ANN analysis was developed to the stage of a proof-of-concept only. Given the issues with the scarcity and the variation of the data, obtaining robust results suggests the application of classical statistical tools in parallel to the ANNs. Furthermore, the ANN approach needs to be preceded with additional tools for data normalization and outlier removal, in order to obtain more robust and reliable results (Basheer and Hajmeer 2000).

3.4 Integration and Analysis of Small-scale Data Sets in Ecophysiological Studies

The use of web-based computational tools for data reposition, sharing and analysis is likely to become more widespread in future international interdisciplinary research projects in ecophysiology. Compared to the currently common practice of data sharing, i.e. using e-mail to exchange spreadsheets based on diverse local data models, integrated data repositories with a clear metadata structure provide a more transparent and consistent platform for discussion, interpretation and further analysis among project partners and the rest of the scientific community. Also in terms of maintaining and curating the data beyond the end of a project, inte-

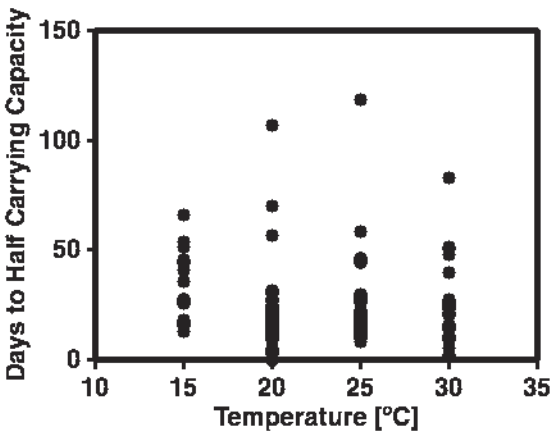

Fig. 6 Example of biofilm growth data from the PHOBIA project. The number of days to reach half the carrying capacity is plotted against the temperature of the corresponding incubator run.

grative platforms present a significant number of advantages. Finally, integrated data repositories are the basis for significant improvements of scientific data workflows by combined tools for data reposition and quantitative analysis (Jones et al. 2006).

The experience from the PHOBIA project highlights two technological challenges involved in the development of integrative data management tools. The first one consists in finding the optimal balance between flexibility and consistency in the definition of the data model: interdisciplinary, multiyear projects on complex biological systems are highly dynamical systems by themselves, hence the corresponding computational tools are required to allow modifications to the data model at a reasonable cost during the life time of the project and beyond. As was shown by the application of the prototype software S3DB, managing metadata using concepts from the semantic web are at the forefront of potential solutions to this challenge (Almeida et al. 2006). The second theme is the adaptation of the analysis tools to the data volume and the quality of the data sources. In general, the data-driven, exploratory, approach chosen here to detect patterns hidden in the data, and proceed with statistically testing concrete hypotheses only after this initial screening, is reasonable, given the situation of having many different variables from diverse sources in the repository. However, for data sets that correspond to a wide table, i.e. have a small number of cases and repetitions compared to a large number of trials, and in addition contain influential and outlying data points, machine learning tools like ANNs are likely to provide unreliable results. Tools using visual exploration (Cleveland 1993), could provide a more efficient first approach to the data structure. With high speed computers being commonly available today, the numerical exploration for the featured type of data set should give preference to modern methods of robust statistics, providing more statistical power (Wilcox 2005).

Just like there is no single “golden model” describing a biological system in a comprehensive manner across all levels of organization, there apparently is no “golden hammer” for data analysis. So far, there is no tool being optimally suited for any type of data set, ranging from high-throughput molecular data, where the sheer amount of data to be processed poses a challenge, to ecophysiogical studies, performed *in vivo* under controlled laboratory conditions or even in the original habitat, where the objective is to obtain the maximum amount of reliable quantitative information from a given set of (unrepeatable) measurements. In the latter scenario, comparing notes on data analysis with astrophysicists, psychologists or social scientists, who experience similar sets of challenges and inherent experimental constraints, might be more fruitful than importing quantitative solutions from computational biosciences, which are optimized for other levels of organization and data volumes.

While integrative computational tools on one hand are necessary to improve the communication and interdisciplinary interaction, their successful application depends on their integration and the overall quality of the communicational network between the members of the research team (Morris et al. 2005). As the PHOBIA example shows, the earlier in the project there is clarity about the data structure and quality, the closer one can get to obtain a seamless chain of knowledge generation, featuring computational tools that are optimally matched to the experimental design and the resulting data structure. These calls for a change of a widespread pattern of communication dynamics between biologists and quantitative scientists: hitherto there is usually a surge of interaction between experimental and computational scientists towards the end of the project, when the data are available. The presented example suggests that a dominant peak of interdisciplinary communication should occur in the very initial stage of each project, in the design phase of the experiments. By eliminating potential bottlenecks in the scientific information workflow at the beginning of a project, the return of knowledge on the investment of time and material resources could be optimized. As discussed in the following concluding discussion, this appeals to a larger awareness of the existing socio-psychological roadblocks and challenges involved in the development of interdisciplinary team science.

4. Summary and Conclusions

The present work features two cases of quantitative ecophysiological studies of phototrophic systems. The example of modeling circadian rhythms in Crassulacean acid metabolism at different levels of organization, from the cell to the whole leaf, demonstrates how quantitative models with diverse levels of abstraction can serve to conduct research on the complex nature of organismal function, and thus contribute to integrate similar phenomena across a variety of organisms. It suggests that a crucial asset of models proving to be successful in advancing knowledge about systems behavior is the potential to be connected with, and to influence other model activities and experimental endeavors. In this perspective, models take the role of navigation tools in complexity (Hammer et al. 2006), rather than the representation of a universal law, or an exact *in silico* map of experimentally observable items. Knowledge discovery itself might turn out be an emergent phenomenon of complex networks of logically connected multilevel models.

The second example, demonstrating the development of a web-based computational infrastructure to integrate heterogeneous datasets on phototrophic biofilms, demonstrates the potential benefits of this type of integrative tool to enhance scientific information- and workflows. It also demonstrates the necessity for intensive interdisciplinary communication between the team members from the very start of a project, as the efficiency of the applied computational tools is crucially depend on their adaptation to the specific character of the project. The central challenge in the development of integrative tools is to achieve an optimal balance between sufficient flexibility for local adaptation to the experimental realities and sufficient structure to provide global description patterns across different species and levels of organization.

Both cases underline the fact that integrative systems biology not only deals with networks of diverse entities in biological systems, but is itself increasingly comprized of dense networks of scientists and resources (Morris et al. 2005). From this point of view, the organization and traditions of conducting sciences have so far followed a reductionist approach: from ancient times until the 20th century, from Aristotle to Einstein, milestones of scientific progress were marked by outstanding individual accomplishments (Humphrey et al. 2005). While specialized expertise in a given experimental or computational method and in a determined area of life sciences continues to be a necessary condition to achieve a better understanding of the complexity of living organisms, it is increasingly clear that it is not sufficient, in analogy to individual molecules which are necessary but by themselves alone not sufficient to provide organismal function.

Significant breakthroughs in modern biology might most likely emerge from collaborations in which the involved partners are committed to the solution of a scientific problem of common interest, for which they hold themselves mutually accountable. This commitment generally involves changes in the attitudes and perspectives on both sides of the disciplinary spectrum (Humphrey et al. 2005). On the side of experimental biologists, it requires the commitment to produce quantitative data of the highest quality possible within a given technical frame, and the disposition to embrace quantitative tools as a vital part of their own scientific endeavor as of the beginning of the project. This requires extra effort to obtain the mathematical knowledge needed to communicate with quantitatively trained scientists, and to gain insight into the chances and limitations that arise from moving from qualitative to quantitative descriptions. This insight might then provide the willingness to put an extra amount of energy in the execution of the experiments, e.g., to perform apparently dull repetitions of measurements to obtain an optimum of statistical power to prove or disprove a given hypothesis. Computational scientists on the other side must increase their appreciation for the technical limitations and the temporal and material investments undertaken by their experimental partners, i.e. appreciate the effort behind each data point. This should be reflected in the optimization of the computational tools for the given data structure, even if that implies the application of already established algorithms, models and software, which do not warrant publication in biocomputational journals. The extra effort consists thus in the intent to satisfy and combine both the team requirements and the individual need to produce novel tools and models which make a difference in the theoretical or computational community. Designing projects that promise a win-win-situation among the project partners could be a strategy to provide a maximum return of quantitative, connectable systems knowledge on a given investment of time and resources.

The required extra investment of time and effort will make these behavioral changes unlikely, unless they are encouraged and facilitated by new modes of organization of the scientific research and its corresponding reward systems. While the importance of interdisciplinary research has become widely recognized by the scientific community, its recognition in scientific practice is still lagging behind (Payton and Zobak 2007). Practical improvements might consist, e.g., in composing committees for any type of scientific reviewing, such that the multidisciplinary spectrum of a given project is adequately mirrored. This could foster the judgment of the integrated contribution to the given biological problem, instead of isolated appreciations of the novelty brought to the diverse disciplines of origin of the partners. As institutional and cultural changes naturally occur on a slow time-scale, short-term measures to promote interdisciplinary research could consist in highlighting success cases in order to develop best-practice standard procedures, and in exposing students of both experimental and math-based disciplines to the world of their respective partners as early as possible in their careers (Morris et al. 2005, Humphrey et al. 2005).

Two of the most prominent examples of team sciences, the Manhattan project leading to the harnessing of nuclear energy, and the Apollo project, leading to man's landing on the moon, were motivated by perceived external threats during World War II and the Cold War. It might also be up to future challenges to mankind, arising from, e.g., the uncertainties of climate change, or the evolution of novel pathogens, to ultimately catalyze the establishment of integrative, team-based approaches as a widespread or even standard protocol in life and social sciences. Promoting interdisciplinary team research out of insight into the high productivity and intellectual inspiration it provides, rather than out of mere practical necessities caused by external pressures, would not only provide novel and powerful approaches to extend human perception and understanding of the complex organization of life, but could also catalyze the development of a general culture of collaboration between people with heterogeneous skills, perspectives and backgrounds.

Acknowledgement

The author acknowledges support by the *Fundação para a Ciência e a Tecnologia*, Portuguese Ministry for Science, Technology and Higher Education, through grant SFRH/BPD/25967/2005.

References

Acebrón, J. A., Bonilla, L. L., Pérez Vicente, C. J., Ritort, F., and Spigler, R.: The Kuramoto model: a simple paradigm for synchronization. Rev. Mod. Phys. *77*, 137–185 (2005)

Almeida, J. S.: Predictive non-linear modeling of complex data by artificial neural networks. Curr. Opin. Biotechnol. *13*, 72–76 (2002)

Almeida, J. S., Chen, C., Gorlitsky, R., Stanislaus, R., Aires de Sousa, M., Eleutério, P., Carriço, J., Maretzek, A., Bohn, A., Chang, A., Zhang, F., Mitra, R., Mills, G. B., Wang, X., and Deus, H. F.: Data integration gets sloppy. Nature Biotechnol. *24*, 1070–1071 (2006)

Basheer, I. A., and Hajmeer, M.: Artificial neural networks: fundamentals, computing, design, and application. J. Microbiol. Meth. *43*, 3–31 (2000)

Beck, F., Blasius, B., Lüttge, U., Neff, R., and Rascher, U.: Stochastic noise interferes coherently with a model biological clock and produces specific dynamic behaviour. Proc. Roy. Soc. Lond. B Biol. *268*, 1307–1313 (2001)

Black, C. C., and Osmond, C. B.: Crassulacean acid metabolism photosynthesis; "working the night shift". Photosynthesis Res. *76,* 329–341 (2003)

Blasius, B., Neff, R., Beck, F., and Lüttge, U.: Oscillatory model of Crassulacean acid metabolism with a dynamic hysteresis switch. Proc. Roy. Soc. Lond. B Biol. *266*, 93–101 (1999)

Bohn, A.: Analysis and simulation of multi-oscillator systems in a Crassulacean acid metabolism plant. Ph.D Thesis, Physics Dept., Darmstadt University of Technology, Germany (2003)

Bohn, A., and García-Ojalvo, J.: Synchronization of coupled biological oscillators under spatially heterogeneous environmental forcing. J. Theor. Biol. *250*, 37–47 (2008)

Borland, A. M., Hartwell, J., Jenkins, G. I., Wilkins, M. B., and Nimmo, H. G.: Metabolite control overrides circadian regulation of phosphoenolpyruvate carboxylase kinase and CO_2 fixation in Crassulacean acid metabolism. Plant Physiol. *121*, 889–896 (1999)

Broda, H., Brugge, D., Honma, K., and Hastings, J. W.: Cellular communication between unicells? Cell Biophys. *8*, 47–67 (1985)

Busch, H., and Eils, R.: Systems Biology. In: Meyers, R. A. (Ed.): Encyclopedia of Molecular Cell Biology and Molecular Medicine *14*; pp. 123–159. Berlin: Wiley-VCH 2005

Cheung, K. H., Smith, A. K., Yip, K. Y. L., Baker, C. J. O., and Gerstein, M. B.: Semantic web approach to database integration in the life sciences. In: Baker, C. J. O., and Cheung, K. H. (Eds.): Semanic Web. Revolutionizing Knowledge Discovery in the Life Sciences; pp. 11–30. New York: Springer 2007

Cleveland, W. S.: Visualizing data. Summit (NJ, USA): Hobart Press 1993

Coetser, S. E., and Cloete, T. E.: Biofouling and biocorrosion in industrial water systems. Crit. Rev. Microbiol. *31*, 213–232 (2005)

Costerton, J. W., Cheng, K. J., Geesey, G. G., Ladd, T. I., Nickel, J. C., Dasgupta, M., and Marrie, T. J.: Bacterial biofilms in nature and disease. Annu. Rev. Microbiol. *41*, 435–464 (1987)

Coveney, P. V., and Fowler, P. W.: Modelling biological complexity: a physical scientist's perspective. J. R. Soc. Interface *2*, 267–280 (2005)

Dodd, A. N., Salathia, N., Hall, A., Kével, E., Tóth, R., Nagy, F., Hibberd, J. M., Millar, A. J., and Webb. A. A. R.: Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. Science *309*, 603–633 (2005)

Duarte, H. M., Jakovljevic, I., Kaiser, F., and Lüttge, U.: Lateral diffusion of CO_2 in leaves of the Crassulacean acid metabolism plant *Kalanchoë daigremontiana* Hamet et Perrier. Planta *220*, 809–816 (2005)

Forger, D. B., and Kronauer, R. E.: Reconciling mathematical models of biological clocks by averaging on approximate manifolds. SIAM J. Appl. Math. *62*, 1281–1296 (2002)

Goldbeter, A.: A model for circadian oscillations in the Drosophila period protein (PER). Proc. Roy. Soc. London B *261*, 319–324 (1995)

Gonze, D., Bernard, S., Waltermann, C., Kramer, A., and Herzel, H.: Spontaneous synchronization of coupled circadian oscillators. Biophys. J. *89*, 120–129 (2005)

Gooch, V. D., Wehseler, R. A., and Gross, C. G.: Temperature effects on the resetting of the phase of the *Neurospora* circadian rhythm. J. Biol. Rhythms *9*, 83–94 (1994)

Grams, T. E. E., Borland, A. M., Roberts, A., Griffiths, H., Beck, F., and Lüttge, U.: On the mechanism of reinitiation of endogenous Crassulacean acid metabolism rhythm by temperature changes. Plant Physiol. *113*, 1309–1317 (1997)

Hammer, G. L., Sinclar, T. R., Chapman, S. C., and van Oosterom, E.: On systems thinking, systems biology, and the in silico plant. Plant Physiol. *134*, 909–911 (2004)

Hammer, G. L., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., Chapman, S. C., and Podlich, D.: Models for navigating biological complexity in breeding improved crop plants. Trends Plant Sci. *11*, 587–593 (2006)

Hartwell, J.: The co-ordination of central plant metabolism by the circadian clock. Biochem. Soc. Trans. *33*, 945–948 (2005)

Humphrey, J. D., Coté, G. L., Walton, J. R., Meininger, G. A., and Laine, G. A.: A new paradigm for graduate research and training in the biomedical sciences and engineering. Adv. Physiol. Educ. *29*, 98–102 (2005)

Hütt, M. T., and Lüttge, U.: Network dynamics in plant biology: current progress in historical perspective. Prog. Bot. *66*, 277–309 (2005)

Ideker, T., Galitski, T., and Hood, L.: A new approach to decoding life: systems biology. Annu. Rev. Genom. Hum. Genet. *2,* 343–372 (2001)

Jones, M. B., Schildhauer, M. P., Reichman, O. J., and Bowers, S.: The new bioinformatics: Integrating ecological data from the gene to the biosphere. Annu. Rev. Ecol. Evol. Syst. *37*, 519–544 (2006)

Keener, J., and Sneyd, J.: Mathematical Physiology. New York/NY, USA: Springer 1998

Keller, E. F.: The century beyond the gene. J. Biosci. *30*, 3–10 (2005)

Kitano, H. (Ed.): Foundations of Systems Biology. Cambridge (MA, USA): MIT Press 2001

Locke, J. C. W., Kozma-Bognar, L., Gould, P. D., Fehér, B., Kevei, E., Nagy, F., Turner, M. S., Hall, A., and Millar, A. J.: Experimental validation of a predicted feedback loop in the multi-oscillator clock of *Arabidopsis thaliana*. Mol. Syst. Biol. *2*, 53 (2006)

Lüttge, U.: The tonoplast functioning as the master switch for circadian regulation of crassulacean acid metabolism. Planta *211*, 761–769 (2000)

Lüttge, U.: Ecophysiology of Crassulacean acid metabolism (CAM). Ann. Bot. *93*, 629–652 (2004)

May, R. M.: Uses and abuses of mathematics in biology. Science *303*, 790–793 (2004)

Michener, W. K., Brunt, J. W., Helly, J. J., Kirchner, T. B., and Stafford, S. G.: Non-geospatial metadata for the ecological sciences. Ecol. Appl. *7*, 330–342 (1997)

Morris, R. W., Bean, C. A., Farber, G. K., Gallahan, D., Jakobsson, E., Liu, Y., Lyster, P. M., Peng, G. C. Y., Roberts, F. S., Twery, M., Whitmarsh, J., and Skinner, K.: Digital biology: an emerging and promising discipline. Trends Biotechnol. *23*, 113–117 (2005)

Nagarkar, S., Williams, G. A., Subramanian, G., and Saha, S. K.: Cyanobacteria-dominated biofilms: a high quality food source for intertidal grazers. Hydrobiologia *512*, 89–95 (2004)

Noble, D.: Modelling the heart: insights, failures and progress. BioEssays *24*, 1155–1163 (2002)

Noble, D.: The future: putting Humpty-Dumpty together again. Biochem. Soc. Trans. *31*, 156–158 (2003)

Nungesser, D., Kluge, M., Tolle, H., and Oppelt, W.: A dynamic computer model of the metabolic and regulatory processes in Crassulacean acid metabolism. Planta *162*, 204–214 (1984)

Paranjpe, D. A., and Sharma, V. K.: Evolution of temporal order in living organisms. J. Circad. Rhythms *3*, 7 (2005)

Parsek, M. R., and Singh, P. K.: Bacterial biofilms: an emerging link to disease pathogenesis. Annu. Rev. Microbiol. *57*, 677–701 (2003)

Payton, A., and Zoback, M. L.: Crossing boundaries, hitting barriers. Nature *445*, 950 (2007)

Peterson, E. L.: A limit cycle interpretation of a mosquito circadian oscillator. J. Theor. Biol. *84*, 281–310 (1980)

Pikovsky, A., Rosenblum, M., and Kurths, J.: Synchronization. A Universal Concept in Nonlinear Sciences. Cambridge (UK): Cambridge University Press 2001

Pittendrigh, C. S.: Temporal organization. Reflections of a Darwinian Clock-Watcher. Annu. Rev. Physiol. *55*, 17–54 (1993)

Rascher, U., Blasius, B., Beck, F., and Lüttge, U.: Temperature profiles for the expression of endogenous rhythmicity and arrhythmicity of CO_2 exchange in the CAM plant *Kalanchoë daigremontiana* can be shifted by slow temperature changes. Planta *207*, 76–82 (1998)

Rascher, U., Hütt, M. T., Siebke, K., Osmond, B., Beck, F., and Lüttge, U.: Spatiotemporal variation of metabolism in a plant circadian rhythm: the biological clock as an assembly of coupled individual oscillators. Proc. Natl. Acad. Sci. USA *98*, 11801–11805 (2001)

Rittmann, B. E.: Microbial ecology to manage processes in environmental biotechnology. Trends Biotechnol. *24*, 261–266 (2006)

Sansone, U., Belli, M., Riccardi, M., Alonzi, A., Jeran, Z., Radojko, J., Smodis, B., Montanari, M., and Cavolo, F.: Adhesion of water-borne particulates on freshwater biota. Sci. Total Environ. *219*, 21–28 (1998)

Schönbach, C., Kowalski-Saunders, P., and Brusic, V.: Data warehousing in molecular biology. Brief. Bioinform. *1*, 190–198 (2000)

Silva, S., Gouveia-Oliveira, R., Maretzek, A., Carriço, J., Gudnason, T., Kristinssion, K. G., Ekdahl, K., Brito-Avô, A., Tomasz, A., Santos Sanches, I., de Lencastre, H., and Almeida, J.: EURISWEB – Web-based epidemiological surveillance of antibiotic-resistant pneumococci in Day Care Centers. BMC Med. Inform. Dec. Mak. *3*, 9 (2003)

Stoodley, P., Sauer, K., Davies, D. G., and Costerton, J. W.: Biofilms as complex differentiated communities. Annu. Rev. Microbiol. *56*, 187–209 (2002)

van Dam, A. A., Beveridge, M. C. M., Ekram Azim, M., and Verdegem, M. C. J.: The potential of fish production based on periphyton. Rev. Fish Biol. Fish. *12*, 1–31 (2002)

Westerhoff, H. V., and Palsson, B. O.: The evolution of molecular biology into systems biology. Nature Biotechnol. *22*, 1249–1252 (2004)

Wilcox, R.: Introduction to robust estimation and hypothesis testing. San Diego (CA, USA): Academic Press 2005

Winfree, A. T.: Unclocklike behaviour of biological clocks. Nature *253*, 315–319 (1975)

Wyka, T. P., Bohn, A., Duarte, H. M., Kaiser, F., and Lüttge, U.: Pertubations of malate accumulation and the endogenous rhythms of gas exchange in the Crassulacean acid metabolism plant *Kalanchoë daigremontiana*: testing the tonoplast-as-oscillator model. Planta *219*, 705–713 (2004)

Zippel, B., Rijstenbil, J., and Neu, T. R.: A flow-lane incubator for studying freshwater and marine phototrophic biofilms. J. Microbiol. Meth. *70*, 336–345 (2007)

Dr. Andreas Bohn
Instituto de Tecnologia Química e Biológica
Universidade Nova de Lisboa
Avenida da República –
EAN, 2780-157 Oeiras
Portugal

Phone:
Fax:
E-Mail: abohn@itqb.unl.pt