

## NEWS AND VIEWS

# International alliances for quantitative modeling in systems biology

Hiroaki Kitano

Sony Computer Science Laboratories Inc, Shinagawa, Tokyo, Japan

*Molecular Systems Biology* 29 March 2005; doi:10.1038/msb4100011

Creating comprehensive models that can predict cellular behaviors is one of the major goals of systems biology. This requires the integration of experimental, computational, and theoretical approaches. Molecular interactions have to be precisely described in mathematical formula that reflects the right level of abstraction suitable for specific biology, and the necessary parameters, such as initial concentration of each component and kinetic constants, have to be estimated from a set of experiments. Although abundant experimental data and published articles are available, creating a quality model by assembling these resources is undesirable as each data set is based on an experiment with a different setup. There is serious need to obtain consistent and comprehensive data measuring different aspects of the focused system so that it can be a basis for quantitative modeling. The problem is that it requires a broad range of expertise and resources, often beyond the capability of a single research group, or even beyond the institution. For example, a particular research group may be able to carry out expression profiles; however, the group may not have the expertise to run image-based time-lapse quantitative localization assays. An experimental group does not necessarily have the computational resources required for modeling. This has led to the establishment of a number of interdisciplinary research groups with extensive networks of collaboration. Collaborations among groups have been common practice in the research community. The major difference in recent collaborations is that of scale leading to self-organization of global alliances to tackle biological complexity. The recent success of global alliance in biology is signified by the human genome project. The emerging alliance requires quite a different approach. In the human genome project, the main issue was how to efficiently and accurately sequence the genome, so the challenge has been quite technical in nature. However, alliances for quantitative modeling are exploratory in nature as what needs to be done has to be investigated as the project move on.

The need for large-scale integrated data sets and the expertise required to comprehend these data is often overwhelming for any single group, and has led to collaboration on a global scale. Several alliances have formed such as the International *Escherichia coli* Alliance (IECA), the Yeast Systems Biology Network (YSBN: <http://www.ysbn.org/>), and the Receptor Tyrosine Kinase network (<http://receptorkinase.gsc.riken.jp>), and these will undoubtedly form a blueprint for numerous alliances in the pipeline. These projects commonly set a goal to build large-scale comprehensive

models that fit currently available data and have predictive power.

It is important to remember that these new alliances are being created in a bottom-up manner, almost solely driven by research needs. They have also come into existence even before any funds have been secured. This position is not through choice but because there are no practical mechanisms for global collaborations, leaving participants to obtain funding from their respective government agencies. The YSBN and RTKsys consortia have written up white papers defining their aims and international scope to provide a springboard for participating researchers to raise funds consistent with the global initiative in each of their funding agencies. Already YSBN is exploring potential funding opportunities within the European Commission's framework program and the SysMO (Systems Biology of MicroOrganisms) initiative, which may lead to pressure to establish liaisons within funding agencies to co-ordinate such alliance funding. Also the soon to be formed International Society for Systems Biology (<http://www.issb.org/>), could help them coordinate and obtain funding. This situation contrasts with other efforts such as the Alliance for Cellular Signaling (AfCS: <http://www.afcs.org/>) and the Systems Biology Mark-up Language (SBML: <http://www.sbml.org/>) initiatives. AfCS is a major US initiative lead by Nobel-prize winner Alfred Gilman and funded by NIH, so that participants are primarily US researchers, and major funding was secured upon the launch of the alliance. The SBML initiative is a successful global initiative, but its initial funding was secured through the Japanese ERATO Kitano Project to maintain a core steering team and central development efforts. Additionally, the SBML initiative targets an issue of practical concern, that is, the standardization of model representations, rather than being focused on modeling itself.

Funding is not the only issue to be resolved. Many scientific hurdles also need to be passed. To obtain consistent experimental data, which is a necessity for successful modeling, it is ideal that a single strain or cell line is used under precisely defined and controlled culture conditions throughout the project, unless comparative studies of multiple strains/cell lines are an explicit topic. However, historically, individual experimental groups have already committed to experimental setup and switching may undermine the value of accumulated data and established expertise. While principles and ideals are well understood, it is hard to convince participating experimental groups to switch to a newly defined setup without

substantial practical benefits. In AfCS, a cell line and culture conditions have been defined by the steering committee, and provided by a group in charge of cell line maintenance. This is possible because it is funded as one national project by a single agency. In the emerging alliances, there is no such top-down scheme available. A possible remedy was suggested recently in a YSBN meeting, which is to recommend a few culture protocols and strains, and once researchers realize the benefit of fast accumulation of controlled data sets and integration into models, the recommended protocols will be adopted. If this approach is adopted, it may also be beneficial to the research community outside any alliance, as experimental results, obtained using different protocols and strains, could be compared with the data produced from the alliance. Therefore, the alliance would be a 'hub' mediating the comparison of a variety of experimental results. A major motivation for researchers within any alliance is the generation of reference mathematical models that have wide visibility and impact in the community. In fact, in yeast this process is already well underway with global deletion strains available in one genetic background.

It is essential that experimental efforts be orchestrated to fill the needs of model development. This means that any project needs to have strong modeling efforts from the beginning in trying to identify the needs for measurements and determine the experimental design. It is reasonable that alliances have been launched in Yeast and RTKsys pathways because they represent two of the best studied model systems for computational modeling, which is sharp contrast to AfCS that started the project without computational methods. Nevertheless, the methods to estimate useful parameters from data are not fully established, and it is unclear at this time which experimental data and specific protocols will prove to be most useful in model development. Hence, large-scale systems biology alliances such as YSBN and RTKsys also require significant components of development of technologies of quantitative biology.

Another issue is whether an alliance should build one single model, or whether there should be multiple models. It is vaguely agreed that there should be one 'reference model' that covers an entire system. However, it may not be suitable for the

needs of those who are interested in specific aspects of the cellular system. While creating predictive models of limited scope could be a practical target for the next five years, a predictive model of an entire cell is not within our foreseeable reach at this moment. A practical approach may be the creation of a large-scale 'wiring diagram', including the map of interactions, and dynamic models for specific sub-networks or phenomena, compatible with other focused models. If integrity is maintained, all dynamic models could eventually be merged into a large-scale whole cell model.

Integration of the experimental, computational, and theoretical approach at this magnitude has not been tried before. It is interesting to note that these alliances recognize the challenge to mid- to long-term enterprise and emphasize systematic education program for next generation systems biologists. The alliances attract young researchers or researchers from different fields, direct their potential to current problems and promising approaches, may provide access to appropriate interdisciplinary training, and sharpen standards in scientific research and communication to carry these challenges to conclusion.

Finally, and perhaps most importantly, the goal of an alliance must be defined explicitly with milestones to evaluate its progress. This is critical in sustaining fund raising and accountability to taxpayers. It is a feature of successful challenging projects such as the Apollo mission to send man to the moon, computer chess to beat human chess champions, RoboCup to beat a human World Cup soccer team by 2050, and the human genome project to complete sequencing of the entire human genome. What should be the goal of these new alliances? To create a whole-cell or pathway-specific comprehensive model could be an appealing goal. Then, we must answer a series of questions: What is meant by a whole-cell model? What is the level of accuracy, predictability, and explanatory power aimed for? What are the scientific and practical advances gained by accomplishing the goal? By when should it be accomplished? In fact, these very questions are major topics in the alliances cited here. Whatever the road map created, alliances are now embarking on challenges where none has gone before.