Please cite this article in press as: Gomez-Cabrero et al., Guidelines for Developing Successful Short Advanced Courses in Systems Medicine and Systems Biology, Cell Systems (2017), http://dx.doi.org/10.1016/j.cels.2017.05.013

Cell Systems
Commentary

# Guidelines for Developing Successful Short Advanced Courses in Systems Medicine and Systems Biology

David Gomez-Cabrero,[1,2,3,4,5,*] Francesco Marabita,[1,2,3,4] Sonia Tarazona,[6,10] Isaac Cano,[7,8] Josep Roca,[7,8] Ana Conesa,[6,11] Philippe Sabatier,[9] and Jesper Tegnér[1,2,3,4,12,13,*]
[1]Unit of Computational Medicine, Department of Medicine, Karolinska Institutet, 171 77 Stockholm, Sweden
[2]Center for Molecular Medicine, Karolinska Institutet, 171 77 Stockholm, Sweden
[3]Unit of Clinical Epidemiology, Department of Medicine, Karolinska University Hospital, L8, 17176 Stockholm, Sweden
[4]Science for Life Laboratory, 17121 Solna, Sweden
[5]Mucosal and Salivary Biology Division, King's College London Dental Institute, London SE1 9RT, UK
[6]Centro de Investigacion Principe Felipe, 46012 Valencia, Spain
[7]Hospital Clinic de Barcelona, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, 08007 Barcelona, Spain
[8]Center for Biomedical Network Research in Respiratory Diseases (CIBERES), 28029 Madrid, Spain
[9]TIMC-IMAG Laboratory, UMR 5525, Centre National de la Recherche Scientifique, Vetagro Sup, Université Grenoble-Alpes, 38400 Saint-Martin-d'Hères, France
[10]Department of Applied Statistics, Operations Research and Quality, Universitat Politècnica de València, Camí de Vera, 46022 Valencia, Spain
[11]Microbiology and Cell Science Department, Institute for Food and Agricultural Sciences, University of Florida, Gainesville, FL 32603, USA
[12]Biological and Environmental Sciences and Engineering Division (BESE), Computer, Electrical and Mathematical Sciences and Engineering Division (CEMSE), King Abdullah University of Science and Technology (KAUST), Thuwal 23955–6900, Kingdom of Saudi Arabia
[13]Lead Contact
*Correspondence: david.gomezcabrero@ki.se (D.G.-C.), jesper.tegner@ki.se (J.T.)
http://dx.doi.org/10.1016/j.cels.2017.05.013

Systems medicine and systems biology have inherent educational challenges. These have largely been addressed either by providing new masters programs or by redesigning undergraduate programs. In contrast, short courses can respond to a different need: they can provide condensed updates for professionals across academia, the clinic, and industry. These courses have received less attention. Here, we share our experiences in developing and providing such courses to current and future leaders in systems biology and systems medicine. We present guidelines for how to reproduce our courses, and we offer suggestions for how to select students who will nurture an interdisciplinary learning environment and thrive there.

## INTRODUCTION

With the arrival and continuing rise of high-throughput technologies generating data reflecting the complexity of biological systems during health and disease, biomedical investigators have accordingly been required either to set up and lead multi-skilled research teams or to rely on collaborations incorporating and representing novel techniques and strategies optimizing their implementation. Systems biology (Chuang et al., 2010; Kitano, 2002) and systems medicine (Gomez-Cabrero et al., 2014b; Hood et al., 2013) are examples of such emerging paradigms in biomedical research.

On the one hand, both systems biology and systems medicine effectively consider three complementary perspectives when approaching problems: the elements of the system, the interactions between them, and finally, the often-non-intuitive emergent properties that these interactions may produce in space and time. Yet, on the other hand, the overarching objectives of systems biology and systems medicine are different. Systems biology aims to achieve a fundamental understanding of and, in some cases, an engineering control over complex biological systems. In contrast, investigators within systems medicine aim for tangible clinical outcomes such as the accurate prediction and assessment of disease, risk, and drug responses using personalized medicine. Fundamental understanding derived from a systems biology investigation does not necessarily lead to a medical advance. Similarly, advancing personalized medicine could be accomplished using predictive statistical approaches that do not require fundamental mechanistic understanding of the underlying biological dynamics. To us, one consequence of these differences is that students of systems biology and systems medicine have different educational needs.

Often, these students are not traditional. Since the implementation of system-based methodologies requires the combination of biological knowledge, clinical skills, the ability to handle biotechnological platforms, and quantitative computational analysis, systems medicine teams benefit from having members who are competent in these different areas. Furthermore, the leader of such a team needs to infuse an open, critical, creative, and respectful atmosphere where different minds and skills can meet on equal terms. This rather new

Please cite this article in press as: Gomez-Cabrero et al., Guidelines for Developing Successful Short Advanced Courses in Systems Medicine and Systems Biology, Cell Systems (2017), http://dx.doi.org/10.1016/j.cels.2017.05.013



**Box 1. The Main Characteristics of Systems Biology and Systems Medicine Courses**

| | Systems Biology | | Systems Medicine | |
|---|---|---|---|---|
| | Characteristics | Possible interests in a course | Characteristics | Possible interests in a course |
| Data Integration | Focus on understanding biological systems.<br>Interest to identify emergent behaviours. | Network analysis | Special interest in including genotype information.<br>Biomarker discovery driven analysis | Feature selection<br>Nature, benefits and limitations of genotype information |
| Statistics and Machine Learning | Outcome, derived by specific methods/parameters, affects conclusions gained | Feature selection for input in follow-up analysis | Considered for clinical applications | Overview, without being too technical, of the possibilities and limitations of these techniques |
| Mechanistic modeling and dynamics | Aims for a **mechanistic** understanding of a system.<br>Need to incorporate biological knowledge<br>Interest to generate or validate hypothesis. | Dynamical modeling techniques<br>Biological knowledge for quantitative trained students<br>Control theory | Large physiological models such as heart or lung<br>Clinical applications | Long-term aims of systems medicine |
| Data | Publicly available data<br>Limited number of replications with a limited genetic background | Available sources: GEO, SRA, etc.<br>Guidelines for experimental design | Restricted access<br>Larger number of samples with large variability<br>Poor annotation of confounders | Data use and access elements: ethical, filling requests, sources,<br>Propose analysis when counfounders do exist<br>Raise awareness of importance of data sharing |

situation creates a non-trivial educational challenge.

Historically, there has been a deeper development of thorough educational programs in systems biology compared to systems medicine. For example, at Princeton, under the guidance of David Botstein (Bialek and Botstein, 2004) among others, serious thinking has been devoted to developing a Ph.D.-level program using a physics-inspired approach to the characterization of living systems (Wingreen and Botstein, 2006). This kind of deep labor-intensive approach requires extensive interaction and development across departments and associated faculty, which is non-trivial. For example, in their pioneering efforts, textbook examples in physics such as electrical circuit problems were replaced using biological examples. Walking through the curriculum in such a manner, rethinking the example problems and skill sets to be communicated to the students, requires genuine collaboration across faculty (http://www.princeton.edu/genomics/botstein/education/). Similarly, recent reviews have summarized the experiences of longer, newer educational programs in systems medicine and have discussed what could and should be included in such programs (Cascante et al., 2014; Cvijovic et al., 2016).

Although less comprehensive, there are also several masters-level programs that teach systems approaches to biology. In many cases, these are open to both engineering and biology students. In practice, such programs include several blocks from what could be referred to as either biology or the physical and computational sciences. This is in contrast to the Princeton example, where an intimate synthesis of biology and the physical sciences has been explicitly designed. This latter approach has been employed in systems medicine. For example, the one-year Clinical Research Program at the Universitat of Barcelona was designed for an audience with multiple backgrounds, ranging from nurses to mathematicians. It has proven to be successful, but the kinds of courses and pieces of knowledge that should or could be included in short master programs remain undefined. It is also unclear whether more fundamental redesigns, similar to the Princeton example, would be advantageous.

Yet, beyond such programs, redesigned or not, there is a need for many students to access shorter overview courses or hands-on courses targeting specific skills. This is particularly valid when considering the needs of physicians and members of the pharmaceutical industry, who need training systems medicine but do not have the time to complete an entire masters-level program. In our hands, we find that there has been an increasing interest in developing short courses that target slightly different audiences: either systems biology in general or systems medicine in particular. Given that experience, our aim in the present paper is to review the educational challenges in current short systems biology or systems medicine courses. Furthermore, when there is sufficient experience, we will highlight unique challenges for shorter courses in systems medicine. Using specific examples from our experiences, we will address the following questions: (1) how should we educate the future leaders and practitioners in systems biology and systems medicine?; (2) how should we provide continuous education to existing professionals, such as clinicians, principal investigators, or senior postdocs in academia or industry?; and (3) how should these differences between the needs of systems medicine and systems biology be reflected in courses and workshops?

## Distinct Challenges Posed by Systems Medicine and Systems Biology and Their Ramifications for Curriculum Design

At a first glance, systems biology and systems medicine appears to be very similar fields. However, they do not have to be. For example, much of the progress in medicine relies on empirical observations and the identification of correlations between biomarkers and individual response to therapy. Understanding these data does not necessarily depend on a quantitative account of the underlying causal biological processes. Below, we highlight four themes, which differentiate the needs of systems biology and systems medicine and, therefore, define their different educational needs. The themes, summarized in Box 1, are as follows: (1) data integration, (2) statistics and machine learning, (3) mechanistic modeling and dynamics, and (4) access to data.

### *Data Integration*

A major development in biomedical research is the accelerated growth of the amount and diversity of available data (Gomez-Cabrero et al., 2014a; Ma'ayan et al., 2014; O'Driscoll et al., 2013) and factual knowledge of biological processes in different contexts (Cano et al., 2014; Subramanian et al., 2005) obtained during the last two decades. This trend has been fueled by several large-scale consortia projects in biology, initiated

Please cite this article in press as: Gomez-Cabrero et al., Guidelines for Developing Successful Short Advanced Courses in Systems Medicine and Systems Biology, Cell Systems (2017), http://dx.doi.org/10.1016/j.cels.2017.05.013





with the Human Genome Project (Lander et al., 2001) and followed by others such as FANTOM (Andersson et al., 2014), ENCODE (Gerstein et al., 2012), UniProt (UniProt Consortium, 2015), Human Proteome (Uhlén et al., 2015), and Blueprint (Adams et al., 2012) projects. Importantly, some projects specifically target human variation, such as the 1000 Genomes Project Consortium (1000 Genomes Project Consortium et al., 2012), and human diseases, such as TCGA (Cancer Genome Atlas Research Network et al., 2013). It is clear that this development presents new opportunities and challenges to the prospects of systems biology and systems medicine.

A shared concern for systems biology and systems medicine is the task of how to integrate such valuable public resources with private data and knowledge. The goals of this integration, however, are different for each field. In our hands (Gomez-Cabrero et al., 2014a), we find that the systems medicine community is more interested in integrating DNA sequences, genetic variants, and clinical data. Integration using predictive models may also be useful for linking biomarkers or other clinical signals to diagnosis, prognosis, and response to therapy. In contrast, systems biologists are more concerned with asking how to integrate different omics data into a network representation capturing the process of interest. Hence, from the standpoint of systems biology, it became necessary in our courses to replace the analysis of single features (e.g., a single genetic variant being differentially expressed and believed to be causally associated with a disease) with the simultaneous analysis of many features (e.g., many genes being de-regulated in a disease but without clear causal association). Many computational tools and approaches to analyze biological systems are available (including diseases [Gustafsson et al., 2014]); examples are network analysis (Jeong et al., 2000; Menche et al., 2015), dynamical modeling (Le Novère, 2015), and Boolean networks (Wittmann et al., 2009), among numerous others. This situation clearly poses unique needs and requirements for short-term systems biology courses.

### Statistics and Machine Learning

Statistics has been revitalized by omics data, including genetic data, because it has a large number of features, relatively few samples, thus requires new and powerful methods for calculating significance. This has been a major challenge and concern within systems biology community (Hawkins et al., 2010). In short, if the feature selection (genes, proteins, metabolites) is erroneous, then the subsequent network analysis will consequently be flawed. Thus, in our experience there is a larger need for technical hands-on treatments capturing this difficulty within the framework of Bioconductor (Huber et al., 2015; Kannan et al., 2016) in systems biology courses. However, analysis based on machine learning and predictive statistical modeling has captured the interest of the systems medicine community (e.g., TCGA derived analysis [Cancer Genome Atlas Research Network et al., 2013]), since such methods can be considered as a shortcut for predicting clinical outcomes and providing integration without requiring mechanistic understanding of why such correlations exist. Hence, we find it is important to provide systems medicine investigators—those with a clinical background in particular—a sound overview of the possibilities and limitations of these techniques. This includes topics such as over-fitting (Sîrbu et al., 2011), prediction versus mechanistic understanding (Gomez-Cabrero et al., 2014c), and data requirements for in many cases very data hungry techniques (Tandon et al., 2006).

### Mechanistic Modeling and Dynamics

The systems biology community has traditionally been deeply interested in dynamical modeling techniques and control theory, which in many cases are neither required nor feasible in clinical contexts (Gomez-Cabrero and Tegnér, 2017, Tegnér and Gomez-Cabrero, 2017). The language and concepts could potentially be useful in considering disease development and progression, but given the mathematical technicalities and ease of misuse if not properly understood or if attempted with insufficient data, we find this aspect less critical in short courses targeting the systems medicine community. Therefore, we find that investigators of systems biology require more robust quantitative model based skills (to make correct use of system-based tools) compared to the more clinical topics addressed in systems medicine (Strunz et al., 2016; Wang et al., 2015). Moreover, we have found it to be of paramount importance to stress the relevance of incorporating broad biological and, in some cases, medical knowledge (to make correct interpretations and usage of the analytical outcome) in systems biology research, which otherwise runs the risk of becoming purely investigations of applied and computational mathematics (Nussinov, 2015; Tegnér et al., 2009).

### Access to Data

Access to data and control of its quality, including standardization, has been a very active area in bioinformatics and systems biology. While there remains much work to do, progress has been significant compared clinical data. Such data is, in most if not all cases, accessed ad hoc in specific projects and poorly documented, and standardization and quality control are all outstanding issues. This state of affairs, together with legal and ethical issues, constitutes a major roadblock for progress in systems medicine (Tegnér and Abugessaisa, 2013). Hence, this is a topic that requires quantitative analysis, biological understandings, clinical skills, and ICT (Information and Communication Technologies) aspects of computer science that need to be discussed in short systems medicine courses (Cascante et al., 2014). Specifically, we find it useful to raise the awareness of this to clinicians and compare and contrast the status quo in medicine with what has been achieved in the omics space (Tegnér and Abugessaisa, 2013).

In summary, to share our overarching view of the education landscape, we have organized some of the current and past educational programs in systems biology and systems medicine in Figure 1 along two axes. The horizontal axis portrays the aim of the training module: from short introductory courses (or very specific courses) on the left to detailed extended educational plans on the right-hand side. The vertical axis denotes the differences between systems biology and systems medicine; systems medicine has been, for example, defined as "the application of systems biology approaches to medical research and medical practice" (European Commission, 2010), indicating that the scopes and goals are different for each discipline. We would like to remark that courses differentiate between systems biology and systems medicine on the basis of their final goals (basic science versus translational applications). Importantly, there are successful courses,

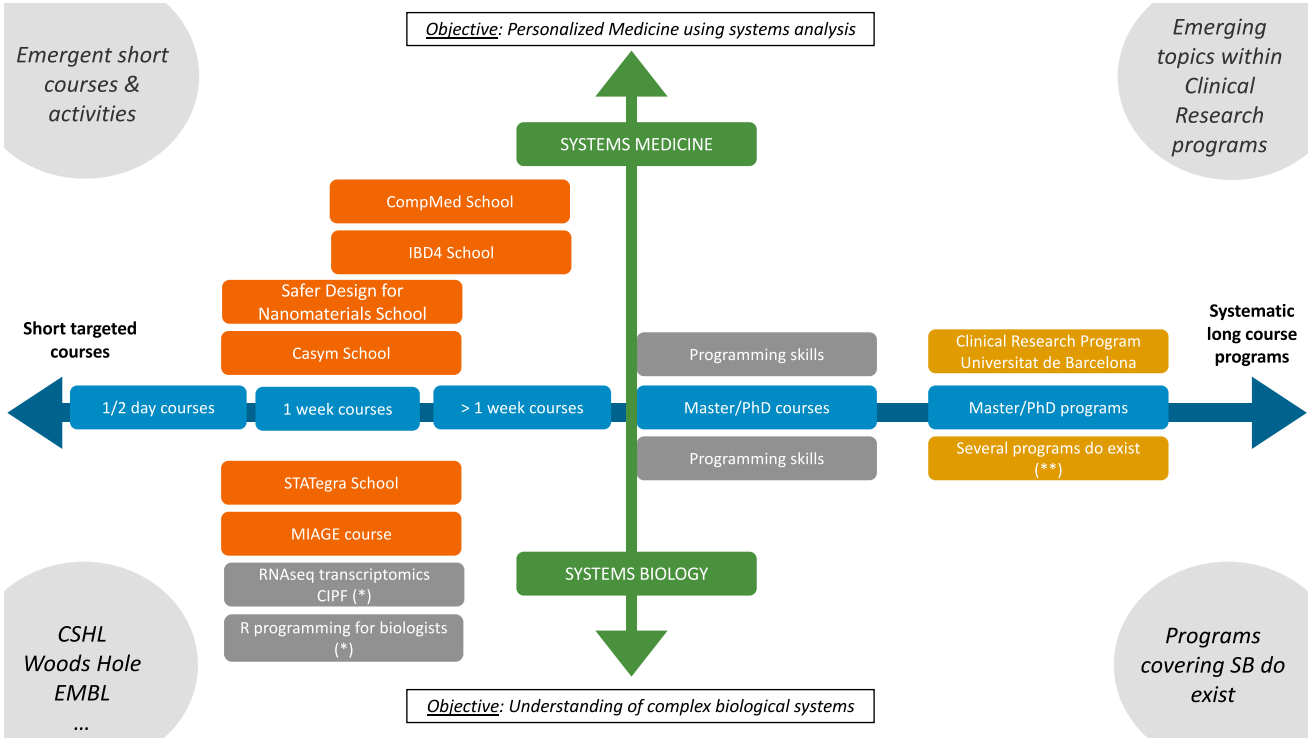


**Figure 1. Educational Efforts in Systems Biology and Systems Medicine**
The two-dimensional schematic plot stratifies courses/programs according to whether the primary objective is systems biology or systems medicine (vertical axis) and the duration of the course/program (horizontal axis). Asterisk (*) denotes courses that are not considered as SBM courses still necessary background for many SBM researchers. Two asterisks (**) denotes the existence of several educational programs addressing the topic (especially in the context of SBM).

such as the R course and the RNA sequencing (RNA-seq) course, which provide important technical skills that do not necessarily fall into either the systems biology or the systems medicine group (in gray in Figure 1).

Short courses in systems biology (left in Figure 1) provide a mechanism to lecture students not only in novel methodologies or theories, but also in topics that are not yet incorporated in the regular educational curriculum (e.g., multi-omic statistical integrative methods in clinical research). Existing publications on systems biology education describe necessary efforts in the context of long-term educational projects (Cascante et al., 2014; Cvijovic et al., 2016)—training the future working force—but they have not yet developed specific guidelines for short-term educational efforts such as summer schools.

## Social Challenges Shared by Short Systems Biology and Systems Medicine Courses

While systems biology and systems medicine courses have different goals and aims, they do share specific characteristics. From now on, we will use the acronym SBM when, despite their differences, we refer to both systems biology and systems medicine courses. In this section, we first aim to characterize the differential nature in short-time SBM courses in relation to other short-time courses, and next, we elaborate those further into specific guidelines. As general characteristics, we have identified the following three.

### *Collaborative Nature*

In graduate courses, students collaborate in activities or tasks; ideally, a collaborative bond is developed during the duration of the course. In a short course, time to establish bonds is limited, and educational and social activities have to be deliberately organized in order to generate collaborative relations as soon as possible. Especially participating individuals are expected to have complementary backgrounds, and collaboration is expected and necessary. Multidisciplinarity is a special characteristic of SBM courses that is not required in specific courses (such as the R and RNA-seq courses discussed in Conclusions).

### *Language Harmonization and Team Formation*

We have witnessed the pervasiveness of two phenomena: the formation of multi-skilled research teams operating in a single group and, alternatively, reliance on external collaborators with training in several areas that augment the local research team. To set up, lead, and sustain a multi-skilled team either locally or externally highlights the challenge of setting a common language and shared definitions. This a very practical issue for both systems biology and systems medicine. For instance, during the Synergy-COPD project (Gomez-Cabrero et al., 2014b), a systems medicine project, researchers were required to harmonize definitions of probability networks and clinical criteria defining different stages of COPD to ensure that clinicians, statisticians, bioinformaticians, mathematicians, computer scientists, and biologists in the project were on the same page and referring to the same concept or, for that matter, that they grasped the differences in vocabulary. In short-term projects, lack of language harmonization may hinder

Please cite this article in press as: Gomez-Cabrero et al., Guidelines for Developing Successful Short Advanced Courses in Systems Medicine and Systems Biology, Cell Systems (2017), http://dx.doi.org/10.1016/j.cels.2017.05.013



CellPress

**Box 2. Examples of SBM Courses**

| Course | Type | Web-site | Description | Duration (Days) | Student Heterogeneous background | Programming skills | Poster Session | Multi-skilled groups | Project | Hands-on | Lectures recorded | Business Development | Topics (summary) | Software & Programming |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Casym School | Systems Medicine | Link | Showcases of success stories of systems medicine projects including a critical discussion of challenges and opportunities ahead. Co-financed by FEBS. | 5 | | | | | | | Table S1 | | Network analysis<br>Systems Medicine case studies | R, Rstudio, Cytoscape, SQL |
| STATegra School | Systems Biology | Link | Prepaing students in the analysis of multi-omics data and to generate long-term educational resources. We consider data-integration at the very core of Systems Biology (and Medicine). | 5 | | | | | | | Table S2 | | Omic analysis tools<br>Integrative analysis tools<br>Visualization tools | R, Rstudio, Cytoscape, Python, Paintomics |
| CompMed School | Systems Medicine | Link | Analytical tools and methodologies for P4 Medicine. The goal is to train clinicians, researchers, and engineers in healthcare optimization using computational methods of modelling and simulation relevant to chronic diseases. CompMed is supported by Erasmus Mundus Programme (EACEA) and CASyM. | 14 | | | | | | | | | Knowledge databases<br>Mechanistic modeling<br>Network Medicine<br>COPD case study | R, Cytoscape, SQL |
| IBD4 | Systems Medicine | Link | Big Data computing and analytics for health and wellbeing,. Enhance innovation and entrepreneurial awareness among participants. Supported by EIT Health, Erasmus Mundus Programme and CASyM. | 10 | | | | | | | | | Case study: Health, wellbeing, obesity<br>Experimental design<br>Sample collection and storage<br>Cloud computing, data analytics<br>Visualization | R, Cytoscape, SQL |
| SDN | Systems Biology | Link | Challenges of safer eco-design for nanomaterials. Reorientation of toxicity testing towards evaluation of the systemic responses of toxicity pathways. | 6 | | | | | | | | | Life-cycle analysis<br>Transformation in the environment<br>Ecosystem and human exposure<br>Toxicology and risk assessment<br>Technology transfer<br>Business development | |
| MIAGE | Systems Biology | Link | Perform an integrative analysis of different omic data types in order to model the regulation of gene expression: from the pre-processing of omics (special focus on NGS platforms) to the study of integration examples covering different integration strategies. | 5 | | | | | | | | | Data pre-processing and analysis for each omic; Overview of integration strategies and software; Matching omic features; Integration examples: analysis and visualization | R, Paintomics, Cytoscape, Rgmatch |

Links to videos are provided in Tables S1 and S2. All courses are fully detailed in Complete Course Description.

the completion of the project if not addressed at the outset. Having such a shared understanding moderates unrealistic expectations on either the computational or biology and biomedical side.

Additionally, although a few individuals could be trained equally in both biology and quantitative skills, it is not to be expected for a mathematical modeler to have the same insights in the disease that a clinician has or for a clinician to develop complex mathematical models. Still, it remains necessary for researchers with quantitative training to learn biology and for researchers with biological knowledge to gain individual quantitative skills (Cascante et al., 2014; Cvijovic et al., 2016).

### *Carefully Designed Intended Learning Outcomes that Reflect Student Backgrounds*

If a student with a biological background attends a course on quantitative analysis, it is likely that the student will require investing extra time (compared to a student with quantitative skills) to learn and master all the knowledge of the course. In a graduate or undergraduate course spanning over several months, that is feasible; however, in a short course, that may not be feasible, and therefore, intended learning outcomes need to be aligned to the duration of the course.

### *Careful Student Selection Criteria*

In our experience and despite our efforts to explain "*the goals, necessary background and scope*" of every SBM short-term course we have had thus far, numerous students attend without having the necessary background for the course. In this case, there is no time for the cancellation of participation (e.g., many summer schools are abroad). For this reason, we think a special emphasis on student selection is necessary. To include detailed motivation letters, interviews over the internet, and project descriptions in addition to the cv is even more crucial in SBM courses.

## Responding to Social and Scientific Challenges through Practical Considerations in Course Design

Above, we have discussed what distinguishes systems biology from systems medicine; we recommend that these distinctions are carefully reflected in any systems biology or medicine course's curriculum. We also discussed these courses from the more personal aspect of the students involved; we recommend that necessary capacities for participants to be reflected in the prerequisites and expectations of a course. Below, we consider practical considerations of a course itself and demonstrate how all of these matters may be brought together into a coherent whole. This is summarized in Box 2, and complete course descriptions can be found in the Supplemental Materials.

### *Importance of Working in Multi-disciplinary Groups*

If the course is (even partly) aimed to teach the SBM (and not only to teach a unique specific technique), it is essential in SBM courses to create an environment similar to that of a systems biology or systems medicine research team. To this end, the idea of working in multi-skilled teams is crucial. To accomplish such a goal, we recommend the following criteria:

*Student Selection*. It is more relevant to have a set of students covering many knowledge backgrounds than to have the best students covering in depth a unique background. During the short courses, students are expected to work in teams, and those teams are required to have backgrounds in biology, chemistry, physics, statistics, programming, engineering, clinical sciences, etc. It can only be achieved by incorporating such criteria during the selection. Additionally, other selection criteria, such as the prioritization of multi-skilled students, may be considered. In our experience, an optimal or desired proportion would be a 1:1 relationship between life science and computational biology profiles. However, in practice, we have observed that in systems biology courses, the applications tend to pile up as a 1:2 proportion, whereas among the applicant to a systems medicine course, the ratio could well be reversed (and

Please cite this article in press as: Gomez-Cabrero et al., Guidelines for Developing Successful Short Advanced Courses in Systems Medicine and Systems Biology, Cell Systems (2017), http://dx.doi.org/10.1016/j.cels.2017.05.013



several of the life science participants are MDs or MDs by training).

*Activity for Initial Bonding*. To foster collaboration between students, it is relevant to have activities that allow them to interact from day 1. In our experience, part of the first day should be devoted to activities such as a poster session, games in groups (e.g., quiz), activities to expose individual current research, and, importantly, activities to define working groups that would select (or would be provided) scientific questions to be addressed during the remainder of the course.

*Added Value by Providing a Highly Structured Curricula*. In introductory biology courses, it has been shown that a well-designed course structure is associated with improved achievements (Haak et al., 2011). In our experience, this applies principally in the context of short courses. While the structure must be very specific to each course, we provide some general observations and specific recommendations.

*Preliminary Material*. Short courses are, by definition, very limited in time, and anything students may prepare before the course increases the individual outcomes. Required and proposed material must include research papers, books, and (very importantly) videos (such as those in Massive Open Online Courses [Searls, 2014]). If the preliminary material is organized into sub-topics, it will be easier for the students to realize how to spend their time to fill knowledge gaps before the course. The preliminary material should also help to define the baseline level of knowledge assumed by the courses and illustrate what knowledge is necessary.

*Lectures*. Active learning is evidently associated with improved achievements while learning biology (Haak et al., 2011). However, students in short courses benefit from general lectures describing the field because they provide a cohesive and comprehensive overview of the topic of interest with some case studies. For example, it is very useful early on to acquire a grasp on what kind of questions investigators ask, what is the nature of the answers, which problems are open while still being feasible to attack, and finally, what are the modes of validation in the different contexts ranging from biological mechanisms, algorithms, clinical findings, and software. This is an interesting challenge to communicate such insights to non-experts in a sensible manner, sufficiently specific without being lost in details. Overall, we recommend avoiding very detailed descriptions, if not being clearly linked to a general point, but instead providing references to them for the interested student. The idea is that the student thereby gains a better overall understanding and is served with the information of "where to look" if he or she aims for more detail. Additionally, there can be more focused lectures on applying the general information provided into more specific case studies.

*Hands-On Learning*. Part of the course should teach precise applications and illustrate different uses of systems biology and systems medicine. We believe that the insights and value of those applications can only be learned by active learning (Bonwell and Eison, 1991; Haak et al., 2011) and its multiple implementations (project method, activity based learning, group work, etc.). In our experience, it is better to conduct the hands-on activities in working groups of three to six students; each working group should incorporate at least one representative of each knowledge field. The nature of the hands-on learning activities would be particular to the course topic.

*Reviewing Course Material*. Although lectures and hands-on might include introductory sections, it is almost unavoidable that some students will get lost on some parts of the course or misinterpret information. Although instructors should encourage discussion, not all students will participate with the same intensity. We have seen two practices which are essential for letting students benefit the most from hands-on learning activities: a sufficient number of instructors should be present in the class to attend personal or group questions during practical work, and hands-on learning activities should conclude with an exercise where students are asked to provide and discuss their results while instructors comment on key aspects of the results of exercise.

*Group and Wrap-up Projects*. In addition to hands-on learning activities devoted to problem solving or completing predefined exercises, it is important for the students to explore the concept of systems biology and systems medicine freely. To this end, students (organized in groups) may define, during the first or second day, a long-term biomedical or biological question to be addressed during the course. We use cooperative learning (Slavin, 1980) as a guideline in this activity, considering that it structures a positive interdependence between students that is fundamental in systems biology. Alternatively, students can propose a wrap-up project where they combine different aspects of the course material to consolidate their understanding of analysis pipelines and strategies. Typically, we propose a simple but complete systems biology data analysis project, where raw data are provided as a starting point, and biological interpretation of analysis results needs to be delivered as an end point.

*Entrepreneurship*. We suggest that courses include sessions that explore the potential of entrepreneurship in systems biology and systems medicine. In these sessions, students are asked to apply both conceptual knowledge (lectures) and practical skills (from hands-on activities) to ideas and projects that could be competitive in the market. Including such sessions will not always be possible, but even introductory lectures providing insights of opportunities beyond academia will be most valuable. Examples of SBM courses including them are the BioHealth Computing Schools described below.

## Additional Relevant Pointers to Consider

*Lecturers*. The challenge of teaching systems biology requires having instructors with an adequate background and specific skills. Practically, this means that teachers must be able to communicate with students coming from diverse disciplines. We recommend that teachers are trained to supervise and organize group dynamics for multi-disciplinary groups. Also, the number of students per instructor should not exceed 15.

*Characteristics of the Environment and Venue*. Short courses are meant to be intensive, so potential interruptions must be minimized. To avoid educational breaks, it is key to have an optimal course organization and to provide the students appropriate resources to perform their tasks. In this regard, preliminary inspection of internet connection limitations is necessary, together with the required software and hardware installations. We also recommend, if possible, having

available spaces for the groups to work without disturbing each other. A possible, but not compulsory, format is to conduct the course in an isolated location to reduce the occurrence and effects of possible distractions.

*Recording Lectures*. We consider recording lectures important for two reasons. First, if students have the opportunity to review preliminary material and view lectures from a previous course, they may come with specific aims and questions (Searls, 2014). Second, during the course, students may review the material at home if the classes were recorded. As an example, we provide links to resources generated between 2015 and 2017 courses (see Box 2 and Tables S1 and S2).

*Computer Resources*. Students may bring their own laptops, which makes it easier for them to reproduce any exercise after the course is finished but limits to the use of tools available in all operative systems (Windows, Unix, OS). Alternatively, the course may provide the students with computers that all share the same operating system. A proposed solution is to provide to all students a bioLinux Live CD or DVD including all necessary tools; in this case, the programs run directly from the CD or DVD, making them slower, but it serves as an ad hoc solution for the use of unix-only tools. In most of the courses we provide, we ask students to bring their own laptops.

## Conclusions

As already indicated by Francois Jacob, "Every object that biology studies is a system of systems." (Jacob, 1974). Understanding the complexity of biological systems represents the greatest intellectual and experimental challenge yet faced by any biologist. While systems biology thinking has generated revolutions in population biology, ecology, and evolutionary studies, it is only recently that biomedicine has adopted a systems approach. The enormous growth in genomics and molecular sciences makes this possible.

Here, we argue that the rapid expansion in SBM tools, resources, and data explosion makes a permanent need for alternative educational activities in the form of short courses. Short courses bring the need of teaching novel tools in educational efforts to the forefront, avoiding the time lag that is existent in the more established educational curriculum. Furthermore, short courses provide opportunities for many individuals, such as senior postdocs or clinicians, in academia or industry who would otherwise be unable to devote themselves for an entire program.

Finally, we would like to stress the need of generating a long-term resource as an outcome of every course. This can be achieved by publishing recorded sessions and used material; as an example of such efforts, the Bioconductor community has a long track of publishing all slides and hands-on material used in almost every course (Huber et al., 2015). Additionally, it would benefit having a central reference point of all material produced by the community such as The Systems Medicine Web Hub (http://www.systemsmedicine.net/).

## SUPPLEMENTAL INFORMATION

Supplemental Information includes two tables and complete course description and can be found with this article online at http://dx.doi.org/10.1016/j.cels.2017.05.013.

## ACKNOWLEDGMENTS

The authors would like to acknowledge helpful comments from the editor. J.T. was supported by EU FP7 305033 CASyM and King Abdullah University for Science and Technology. D.G., F.M., S.T, A.C., and J.T. were supported by EU FP7 306000 STATegra.

## REFERENCES

Adams, D., Altucci, L., Antonarakis, S.E., Ballesteros, J., Beck, S., Bird, A., Bock, C., Boehm, B., Campo, E., Caricasole, A., et al. (2012). BLUEPRINT to decode the epigenetic signature written in blood. Nat. Biotechnol. *30*, 224–226.

Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461.

Bialek, W., and Botstein, D. (2004). Introductory science and mathematics education for 21st-century biologists. Science *303*, 788–790.

Bonwell, C.C., and Eison, J.A. (1991). Active Learning: Creating Excitement in the Classroom (The George Washington University).

Cano, I., Tényi, Á., Schueller, C., Wolff, M., Huertas Migueláñez, M.M., Gomez-Cabrero, D., Antczak, P., Roca, J., Cascante, M., Falciani, F., and Maier, D. (2014). The COPD Knowledge Base: enabling data analysis and computational simulation in translational COPD research. J. Transl. Med. *12* (*Suppl 2*), S6.

Cancer Genome Atlas Research Network, Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A., Hoadley, K., Triche, T.J., Jr., Laird, P.W., Baty, J.D., et al. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. *368*, 2059–2074.

Cascante, M., de Atauri, P., Gomez-Cabrero, D., Wagner, P., Centelles, J.J., Marin, S., Cano, I., Velickovski, F., Marin de Mas, I., Maier, D., et al. (2014). Workforce preparation: the Biohealth computing model for Master and PhD students. J. Transl. Med. *12* (*Suppl 2*), S11.

Chuang, H.-Y., Hofree, M., and Ideker, T. (2010). A decade of systems biology. Annu. Rev. Cell Dev. Biol. *26*, 721–744.

European Commission (2010). From systems biology to systems medicine. Workshop, C. Kyriakopoulou, B. Mulligan, ed. http://ec.europa.eu/research/health/pdf/systems-medicine-workshop-report_en.pdf.

Cvijovic, M., Höfer, T., Aćimović, J., Alberghina, L., Almaas, E., Besozzi, D., Blomberg, A., Bretschneider, T., Cascante, M., Collin, O., et al. (2016). Strategies for structuring interdisciplinary education in Systems Biology: an European perspective. NPJ Syst. Biol. Appl. *2*, 16011.

Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.-K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R., et al. (2012). Architecture of the human regulatory network derived from ENCODE data. Nature *489*, 91–100.

Gomez-Cabrero, D., and Tegnér, J. (2017). Iterative Systems Biology for Medicine – Time for advancing from network signatures to mechanistic equations. Curr. Opin. Syst. Biol. *3*, 111–118.

Gomez-Cabrero, D., Abugessaisa, I., Maier, D., Teschendorff, A., Merkenschlager, M., Gisel, A., Ballestar, E., Bongcam-Rudloff, E., Conesa, A., and Tegnér, J. (2014a). Data integration in the era of omics: current and future challenges. BMC Syst. Biol. *8* (*Suppl 2*), I1.

Gomez-Cabrero, D., Lluch-Ariet, M., Tegnér, J., Cascante, M., Miralles, F., and Roca, J.; Synergy-COPD consortium (2014b). Synergy-COPD: a systems approach for understanding and managing chronic diseases. J. Transl. Med. *12* (*Suppl 2*), S2.

Gomez-Cabrero, D., Menche, J., Cano, I., Abugessaisa, I., Huertas-Migueláñez, M., Tenyi, A., Marin de Mas, I., Kiani, N.A., Marabita, F., Falciani, F., et al. (2014c). Systems Medicine: from molecular features and models to the clinic in COPD. J. Transl. Med. *12* (*Suppl 2*), S4.

Gustafsson, M., Nestor, C.E., Zhang, H., Barabási, A.L., Baranzini, S., Brunak, S., Chung, K.F., Federoff, H.J., Gavin, A.C., Meehan, R.R., et al. (2014). Modules, networks and systems medicine for understanding disease and aiding diagnosis. Genome Med. *6*, 82.

Haak, D.C., HilleRisLambers, J., Pitre, E., and Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. Science *332*, 1213–1216.

Hawkins, R.D., Hon, G.C., and Ren, B. (2010). Next-generation genomics: an integrative approach. Nat. Rev. Genet. *11*, 476–486.

Hood, L., Flores, M.A., Brogaard, K.R., and Price, N.D. (2013). Chapter 23—Systems Medicine and the Emergence of Proactive P4 Medicine: Predictive, Preventive, Personalized and Participatory.

In Handbook of Systems Biology (Elsevier), pp. 445–467.

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. Nat. Methods *12*, 115–121.

Jacob, F. (1974). The Logic of Living (Allen Lane).

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., and Barabási, A.L. (2000). The large-scale organization of metabolic networks. Nature *407*, 651–654.

Kannan, L., Ramos, M., Re, A., El-Hachem, N., Safikhani, Z., Gendoo, D.M., Davis, S., Gomez-Cabrero, D., Castelo, R., Hansen, K.D., et al. (2016). Public data and open source tools for multi-assay genomic investigation of disease. Brief. Bioinform. *17*, 603–615.

Kitano, H. (2002). Systems biology: a brief overview. Science *295*, 1662–1664.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al.; International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Le Novère, N. (2015). Quantitative and logic modelling of molecular and gene networks. Nat. Rev. Genet. *16*, 146–158.

Ma'ayan, A., Rouillard, A.D., Clark, N.R., Wang, Z., Duan, Q., and Kou, Y. (2014). Lean Big Data integration in systems biology and systems pharmacology. Trends Pharmacol. Sci. *35*, 450–460.

Menche, J., Sharma, A., Kitsak, M., Ghiassian, S., Vidal, M., Loscalzo, J., and Barabási, A.-L. (2015). Disease Networks. Uncovering disease-disease relationships through the human interactome. Science *347*, 1257601.

Nussinov, R. (2015). Advancements and challenges in computational biology. PLoS Comput. Biol. *11*, e1004053.

O'Driscoll, A., Daugelaite, J., and Sleator, R.D. (2013). 'Big data', Hadoop and cloud computing in genomics. J. Biomed. Inform. *46*, 774–781.

Searls, D.B. (2014). A new online computational biology curriculum. PLoS Comput. Biol. *10*, e1003662.

Sîrbu, A., Ruskin, H.J., and Crane, M. (2011). Integrating heterogeneous gene expression data for gene regulatory network modelling. Theory Biosci. *131*, 95–102.

Slavin, R.E. (1980). Cooperative Learning. Rev. Educ. Res. *50*, 315–342.

Strunz, S., Wolkenhauer, O., and de la Fuente, A. (2016). Network-assisted disease classification and biomarker discovery. Methods Mol. Biol. *1386*, 353–374.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Tandon, R., Adak, S., and Kaye, J.A. (2006). Neural networks for longitudinal studies in Alzheimer's disease. Artif. Intell. Med. *36*, 245–255.

Tegnér, J., and Abugessaisa, I. (2013). Pediatric systems medicine: evaluating needs and opportunities using congenital heart block as a case study. Pediatr. Res. *73*, 508–513.

Tegnér, J., and Gomez-Cabrero, D. (2017). Editorial overview. Curr. Opin. Syst. Biol. *3*, 12–14.

Tegnér, J.N., Compte, A., Auffray, C., An, G., Cedersund, G., Clermont, G., Gutkin, B., Oltvai, Z.N., Stephan, K.E., Thomas, R., and Villoslada, P. (2009). Computational disease modeling - fact or fiction? BMC Syst. Biol. *3*, 56.

Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. Science *347*, 1260419.

UniProt Consortium (2015). UniProt: a hub for protein information. Nucleic Acids Res. *43*, D204–D212.

Wang, R.S., Maron, B.A., and Loscalzo, J. (2015). Systems medicine: evolution of systems biology from bench to bedside. Wiley Interdiscip. Rev. Syst. Biol. Med. *7*, 141–161.

Wingreen, N., and Botstein, D. (2006). Back to the future: education for systems-level biologists. Nat. Rev. Mol. Cell Biol. *7*, 829–832.

Wittmann, D.M., Krumsiek, J., Saez-Rodriguez, J., Lauffenburger, D.A., Klamt, S., and Theis, F.J. (2009). Transforming Boolean models to continuous models: methodology and application to T-cell receptor signaling. BMC Syst. Biol. *3*, 98.

1000 Genomes Project Consortium, Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56–65.