IBI* SERIES WINNER

Students as Collaborators in Systems Biology Research

Quantitative Trait Mapping, an IBI prize–winning module, immerses students in the role of a systems biology researcher.

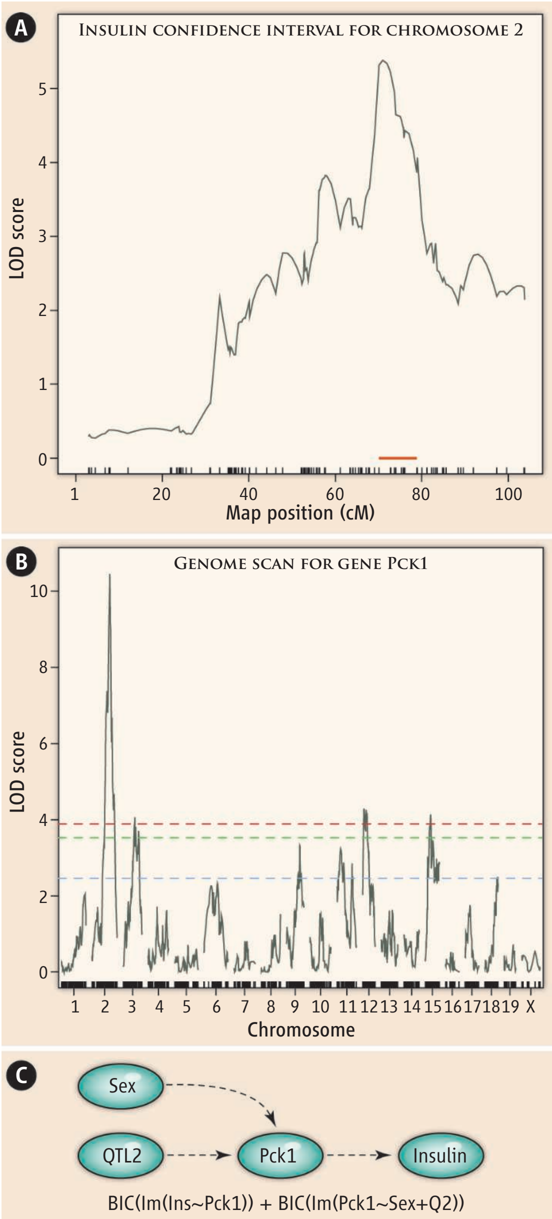
Susan McClatchy,[1] Deborah McGann,[2] Robert Gotwals,[3] Amanda Baskett,[4] Gary Churchill[1]†

In his 1854 speech at the University of Lille, Louis Pasteur stated "In the fields of observation, chance favors only the prepared mind." Preparation for scientific inquiry is critical for the future of research, yet opportunities are rare in the early stages of education. This discrepancy can be resolved by immersing students in genuine research activities as early as possible. We have created an experience for high school students to engage in and contribute to ongoing research. We prepare students to formulate and test hypotheses using computational tools and data collected in our laboratory, or available from public repositories.

The Quantitative Trait Mapping (QTM) module (fig. S1) is a core component of Independent Studies in Computational Biology (ISCB), a two-semester systems biology research course offered to talented high school students by The Center for Genome Dynamics at The Jackson Laboratory (*1*). Introductory high-school genetics and statistics are prerequisites. We employ a hybrid strategy of online learning and in-class instruction in which students and teachers from three publicly funded magnet schools participate via Web conference. Students and their teachers become active partners in our group's research and some have published their work in peer-reviewed journals (*2*, *3*).

Students and teachers meet in their classrooms during the week, and all sites attend a weekly Web conference with a Jackson Laboratory researcher, who acts as their scientific mentor. In the first semester, students read and present information from scientific literature. We select review articles, summaries of featured literature, and reports of experimental findings from leading journals. Students present and discuss papers in journal club, honing their critical thinking and communication skills. Each student writes a brief literature review on a topic of their choosing. In parallel, we provide students with instruction and hands-on practice in statistical analysis of quantitative traits. Students analyze published data from quantitative trait loci (QTL) mapping studies (e.g., QTL Archive; qtlarchive.org) by means of R/qtl software (*4*); they learn how to employ Web-based resources to dig deeper, obtaining up-to-date information on genetic variation and functional annotation. The semester culminates with a written NIH-style grant proposal that integrates material from these exercises.

In the second semester, we begin the research phase of the course. Student teams of two to four are given relevant data but must decide on an analytical strategy that will address their hypothesis. This is a challenging step for students and involves Web conferences with their scientific mentor, who is a practicing research scientist in our lab, to arrive at clear hypotheses that are testable with the data in hand. The students' creative expression of their own ideas must be balanced against the scope of available data resources. This hypothesis-driven approach prepares the student with a conceptual framework for finding biological meaning amid the network of statistical relations that can emerge from large-scale data analysis (see the figure).



A student-initiated project begins with a hypothesis. In this example, the student identified a quantitative trait locus (QTL) for plasma insulin on chromosome 2. (**A**) The support interval for the QTL contains dozens of genes. Using database resources including Mouse Genome Informatics (www.informatics.jax.org), she generated a gene list along with functional annotation and allelic variants. She then carried out a systematic evaluation of each gene. On the basis of her hypothesis, she mapped gene expression traits in adipose tissue. (**B**) She evaluated genes using structural equation models to identify causal intermediates of the effect on plasma insulin. (**C**) She identified Pck1, an enzyme in the pathway that converts triglycerides to glucose, as a candidate with a potential role in diabetes and proposed a follow-up study that could provide validation for her finding.

CREDIT: ADAPTED FROM CASEY THORNTON

[1]The Center for Genome Dynamics, The Jackson Laboratory, Bar Harbor, ME 04609, USA. [2]Maine School of Science and Mathematics, Limestone, ME 04750, USA. [3]North Carolina School of Science and Mathematics, Durham, NC 27705, USA. [4]Rockdale Magnet School for Science and Technology, Conyers, GA 30012, USA.

*IBI, *Science* Prize for Inquiry-Based Instruction; www.sciencemag.org/site/feature/data/prizes/inquiry/.

†Corresponding author: gary.churchill@jax.org

Published by AAAS

Examples of student-initiated investigations include the following:

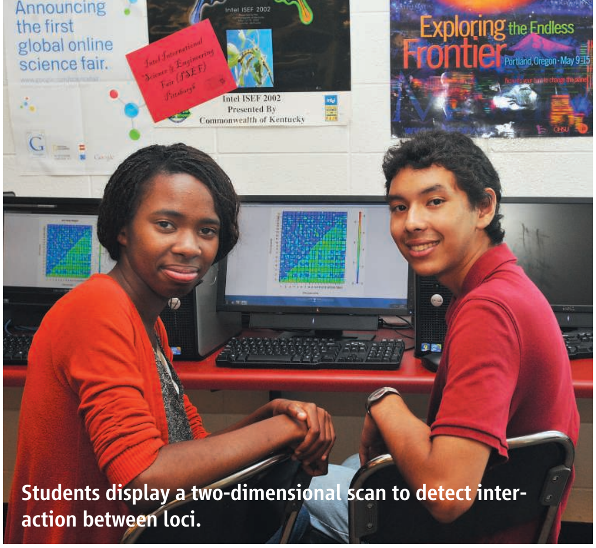
- Activation of brown adipose tissue reduces obesity. Gene expression data were examined from mice fed either standard or high-fat diets. Causal inference methods were used to explore relations between adiposity and the expression of specific brown fat–associated genes.
- Adipocytes communicate to pancreatic beta cells. The relation of adipokines to pancreatic gene expression and plasma insulin levels was investigated by using data from a mouse intercross population.

Students communicate findings in weekly oral progress updates with the scientific mentor. Web conferences provide a venue for addressing students' questions and evaluating student progress and continued development of communication skills. We assess student understanding of background information and review critical concepts on an as-needed basis. E-mail exchanges among students, teachers, and mentors clarify interpretation of results and strategic approach. Classroom teachers, develop their own assessment tools for course assignments (see sample assessment in supplementary materials).

Student projects frequently require refinement or redirection, which can be a frustrating and difficult experience for high-performing students accustomed to having all of the correct answers. They learn that, in scientific inquiry, success is not a foregone conclusion and that there are neither predictable outcomes nor answers at the back of the book. And yet, we have witnessed acceptance of failure leading to perseverance. Often, seemingly false starts lead to unexpected results and novel insights as students experience the process of discovery firsthand (see the photo). The semester concludes with a final oral research presentation and a written report of student findings and conclusions.

We have considered how the QTM module might be expanded to reach a wider audience, including college undergraduates. Toward that end, we provide background reading material, recorded lectures, links to data and software resources, and examples illustrating data-analysis techniques (www.sysgenonline.org). Instructors can supplement this material by guiding students to selections from the scientific literature, selecting data for analysis, and assisting students in the use of Web resources. The didactic content of the module can save class time for discussion and interaction.

Although online learning modules can be widely disseminated, they cannot replace the personal elements of one-on-one interactions with a scientific mentor, a critical component of success. The scientist provides overall direction for the students' projects, as well as data and the expertise to interpret it. Educated and prepared teachers are also essential to the process. They evaluate students' progress, help students find information and resources, debug R code, and review students' practice presentations. The levels of personal attention and dedication required from the scientist and teachers are limiting factors in course expansion. We invite educators and scientists to contact us and join our weekly Web conferences to learn how to implement the module.

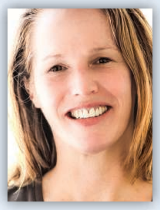
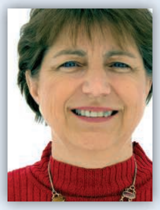


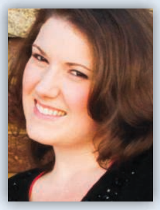
Students display a two-dimensional scan to detect interaction between loci.

Although the footprint of our course has been small (~70 students have completed or are currently enrolled in the course), its impact on these students has been enormous. Testimonials from course alumni, while anecdotal, indicate the power of this research experience in shaping education and career paths. One former student reported that "I'm graduating in Quantitative Biology … because of this class!" and another stated that "The course taught me how to pursue scientific problems from a computational mindset." Students are not the only beneficiaries of the course. As an indication of the effect of this experience on educators, one of our instructors commented, "I anticipated the course would have a great impact on our students, but could not have known how much the experience would influence me as a teacher, re-igniting a drive and desire to ask questions and explore."

Immersing students in scientific research as early as possible bridges the divide between practices in education and research. Early exposure to authentic systems biology–research experiences produces students with "prepared minds" who understand and value scientific research as a career choice, and who are comfortable speaking the languages of biology, mathematics, and computing.

About the authors




Susan McClatchy is the outreach coordinator for the Center for Genome Dynamics at The Jackson Laboratory. **Deborah McGann** teaches chemistry and biology at the Maine School of Science and Mathematics. **Robert Gotwals** is a chemistry and research instructor at the North Carolina School of Science and Mathematics. **Amanda Baskett** teaches research and microbiology at the Rockdale Magnet School for Science and Technology. **Gary Churchill** is a statistical geneticist at The Jackson Laboratory and director of the Center for Genome Dynamics, a National Center for Excellence in Systems Biology.

References and Notes

1. R. Von Smith *et al.*, *NCSSSMST J.* **15**, 10 (2009).
2. K. R. Shockley, D. Witmer, S. L. Burgess-Herbert, B. Paigen, G. A. Churchill, *Physiol. Genomics* **39**, 172 (2009).
3. M. Feng *et al.*, *Hypertension* **54**, 802 (2009).
4. K. W. Broman, S. Sen, *A Guide to QTL Mapping with R/qtl* (Springer Science+Business Media, New York, 2009).

Acknowledgments: Research reported in this publication was supported by the National Institute of General Medical Sciences, NIH, under award no. P50GM076468. We gratefully acknowledge feedback and suggestions from M. A. Handel, I. Hurley, and S. Munger. We thank K. Broman for his excellent R/qtl tutorials. A special thank you goes to R. Von Smith for originating the independent studies in computational biology course.

Supplementary Materials
www.sciencemag.org/cgi/content/full/340/6136/1061/DC1

10.1126/science.1229906