

The Genome Solver Website: A Virtual Space Fostering High Impact Practices for Undergraduate Biology

Anne G. Rosenwald^{1*}, Janet S. Russell², and Gaurav Arora¹

¹Department of Biology, ²Center for New Designs in Learning and Scholarship, Georgetown University, Washington, DC 20057

INTRODUCTION

The Human Microbiome Project (HMP) seeks to investigate the diversity of microbial organisms in and on the human body (3, 4). From current findings, we now know that the organisms found are much more diverse than previously appreciated, that there are many hundreds of new genes described—ones that have no matches in the current databases—and that there is a unique interplay between the members of the microbiome and the human host with impacts on human health (6). With these new data come new opportunities for undergraduate research that need not be limited to independent study projects but can be based in traditional coursework.

Because sequencing is now relatively fast and inexpensive, and because many reference genomes are now available, it is possible to sequence and assemble a moderately sized bacterial genome in a matter of days, yielding massive amounts of data, so much so that professional bioinformaticians are unable to keep up. However, the tools for analysis of prokaryotic genomes are available on the internet and are relatively easy for undergraduates to master. Instead the challenge lies with faculty inexperienced in teaching bioinformatics. To meet this challenge we developed the Genome Solver (GS) website (<http://www.genomesolver.org>), a community of practice for faculty support and student learning (see Fig. 1). At the GS site, faculty can learn about face-to-face training opportunities or use the virtual training materials such as primers on the pertinent science, tools, and techniques. Faculty can also find colleagues engaged in similar work as well as relevant curricular materials. The GS site also helps facilitate sharing of student work, so that students can engage in peer-to-peer review. Finally, genomics experts also have a presence on GS and can weigh in on problems or discuss methods.

Thus, the GS space serves as a virtual community around the study of prokaryotic genomes. Bioinformatics studies, such as those presented in GS, are authentic

research problems, but ones that don't require extensive equipment and space. Rather, the only tools necessary are computers with an internet connection, making research possible for nearly everyone so that faculty can more easily include student research in their curricula.

The recent *Vision and Change in Undergraduate Biology Education: A Call to Action* report recommends that all biology students participate in authentic research experiences (<http://visionandchange.org/files/2011/03/Revised-Vision-and-Change-Final-Report.pdf>). The benefits are many as demonstrated by numerous studies, including the SURE and CURE surveys (1, 2, 5). Further, students themselves identify undergraduate research as a “high impact practice” (National Survey of Student Engagement [NSSE] report http://nsse.iub.edu/NSSE_2007_Annual_Report/docs/withhold/NSSE_2007_Annual_Report.pdf). GS fosters other high impact practices, including common intellectual experiences, learning communities, intensive writing, and collaborative assignments. All of these practices lead to practical competence, deep engagement with and deep learning of the material, personal-social development, and the ability to integrate knowledge across disciplinary boundaries.

PROCEDURE

The interactive GS site is easy for faculty and students to use. The site is built on the Buddy Press platform (<http://buddypress.org>), a free software package that allows

*Corresponding author. Mailing address: Department of Biology, Georgetown University, 37th and O Sts., NW, 406 Reiss Science Center, Washington, DC 20057. Phone: 202-687-5997. Fax: 202-687-5662. E-mail: rosenwaa@georgetown.edu.

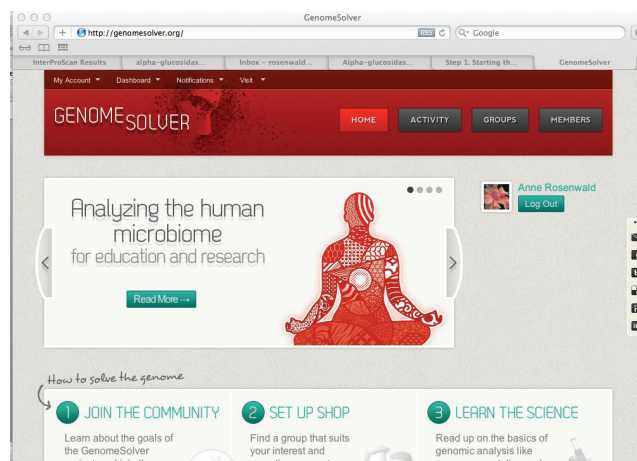


FIGURE 1. The Genome Solver interface.

sharing of materials among users. We have created how-to videos explaining site navigation, introductory materials that faculty can use to orient students to the HMP, links to bioinformatics tools and instructions on how to use them, some basic exercises using the tools, and a web feed feature which automatically aggregates new information about the HMP from both the scientific literature and the popular press on the GS site.

Perhaps the most important feature of the GS site is members' ability to create groups that support their own interests. Groups can bring together faculty, students, and experts engaged in a common study (intestinal microflora, for example) regardless of their institution or geographic location. The groups can be either fully public ones that anyone can join or ones that can be joined by invitation only, at the discretion of the group's creator. In the latter category, for example, are groups for faculty only, where assessment materials and their analysis will be available. We plan to continually update the "nuts and bolts" of the GS site, but importantly the content will be enriched by the contributions of the GS community members. Also, although we currently offer face-to-face workshops to train faculty, we plan for virtual workshops in the future hosted on the GS site that can be continuously updated in this fast-changing field.

The GS community supports many different sub-disciplines within biology, including:

- *Microbiology* – where students learn about the diversity of microbial life;
- *Biochemistry* – where students investigate biochemical pathways in detail;
- *Ecology* – where students explore the different environmental niches found in and the on the human body;
- *Phylogenetics* – where students can use sequence data to compare organisms to one another and explore similarity and differences; and
- *Evolution* – where students can examine aspects of horizontal gene transfer, "pangenomes," and selection.

In addition, by engaging in bioinformatics studies, students explore the intersection of math, statistics, computer science, and biology, a fitting demonstration of the increasingly interdisciplinary nature of scientific endeavors.

We are currently engaged in creating exercises that can be used as is to encourage faculty to explore HMP data, but faculty are also encouraged to modify the existing exercises and add their own curriculum modules to enrich the offerings available for all. An example of an exercise we developed is discussed briefly here, but is available in more detail on the GS web site (<http://genomesolver.org/groups/curriculum-modules/documents/>). A former student discovered through examination of the automatic annotation of a *Chlamydia* genome available from the HMP that genes from

bacteriophages that infect *Chlamydia* species are found in the genomes of all human *C. pneumoniae* isolates sequenced to date. Based on the student's investigation, we have created an exercise ("Evidence for Horizontal Gene Transfer...") to gather related phage and bacterial sequences for a gene that encodes a protein referred to as PRIP (putative replication initiation protein) and create a phylogenetic tree. The tree demonstrates that the bacterial sequences cluster with each other, while the phage sequences, which also cluster with each other, are more diverse, allowing for students to hypothesize that the transfer of a phage sequence to the bacterial genome happened once before the divergence of the different bacterial strains.

CONCLUSION

The study of human health and disease is appealing to undergraduates, many of whom are interested in careers in the health professions. The research community is generating a new view of human health, one that takes into account the microbes that co-exist with us. Undergraduate students can contribute to this new view by performing research with data and tools made accessible to them by their faculty. The GS site aims to put this experience into the hands of all interested faculty. Curricula can range from short units in a pre-existing class to an entire class based on analysis of the tools (see our Microbial Genomics Biology 388 class on the website) to independent research projects. It is important to note that the work students do with GS can be extended into wet lab hypothesis testing (see our *Salmonella* Pilot Studies on the web site). Although we are featuring the recent exciting work about the HMP, the site can be used by anyone with an interest in life at the microbial level. Finally, the design of the GS site is intended to foster practices that are high impact for all of us—students, faculty, and experts.

ACKNOWLEDGMENTS

The Genome Solver Project is funded by the National Science Foundation (DUE-1123016). Justin Secor, with Rob Pongsajapan and Marie Selvanadin, designed the GS interface and logos; and Annalisa Adams recorded the "how-to" videos (all from the Center for New Designs in Learning and Scholarship (CNDLS) at Georgetown University). Thanks also to Theresa Schafly (CNDLS) for carefully reading the manuscript. Our partners in the Genome Solver Project are Ramana Madupu (J. Craig Venter Institute) and Jennifer Roeklein-Canfield (Department of Chemistry, Simmons College). The authors declare that there are no conflicts of interest.

REFERENCES

1. Denofrio, L. A., B. Russell, D. Lopatto, and Y. Lu. 2007. Linking student interests to science curricula. *Science* **318**:1872–1873.

2. **Gregg-Jolly, L. A., R. Kington, D. Lopatto, and J. E. Swartz.** 2011. Benefits of intertwining teaching and research. *Science* **331**:532.
3. **Human Microbiome Project Consortium.** 2012. A framework for human microbiome research. *Nature* **486**:215-221.
4. **Human Microbiome Project Consortium.** 2012. Structure, function and diversity of the healthy human microbiome. *Nature* **486**:207–214.
5. **Lopatto, D.** 2007. Undergraduate research experiences support science career decisions and active learning. *CBE-Life Sci. Educ.* **6**:297–306.
6. **Nelson, K. E., G. M. Weinstock, S. K. Highlander, K. C. Worley, H. H. Creasy, J. R. Wortman, et al.** 2010. A catalog of reference genomes from the human microbiome. *Science* **328**:994–999.