Introduction

As you can see on the class schedule, the final project is divided into several deliverables. In order to ensure adequate attention and time to each part of the project process, they are due over the course of the semester. The first several tasks are designed to have you explore your interests, examine the scientific literature, and seek out effective partnerships with other students. These early milestones are included in the problem sets (and due at the same time). The later, more significant, milestones are due as separate assignments. This project summary document collects several brainstorming techniques and outlines the grading criteria that will be used to evaluate your project. Milestone-specific handouts with more details are included on the course website.

Milestones

The five main deliverables of the project are as follows (copied from the course schedule):

• I: Project preferences and initial brainstorming, due Mon. Sept. 27 (after week 3)
• II: Literature search and pre-proposal, due Mon. Oct. 13 (after week 5)
• IIIa: Proposal draft, due Mon. Oct. 25 (after week 7)
• IIIb: Proposal reviews, due Mon. Nov. 1 (after week 8)
• IIIc: Review response, due Mon. Nov. 8 (after week 9)
• IV: Midcourse report, due Mon Nov. 22 (after week 11)
• Va: Final report, due Mon. Dec. 6 (after week 13)
• Vb: Project presentations, due Mon. Dec. 9 (after week 14)

Advice on selecting a final project

The first part of any research project is coming up with a good, innovative, concrete, and feasible idea. There is no single recipe for getting a good idea for a project, and our best ideas frequently come in unexpected ways. While you are brainstorming, there are several resources available: During lectures, we will discuss several research directions that can be pursued as final projects. Going through your notes can give you more ideas. A brief look at the calendar of remaining lectures can also give you ideas about current research directions, and the lecture notes from last year (linked from the class website) can give you more details on most of these and can help you build your own research projects. The problem sets will provide possible starting points for projects that extend the algorithms and programs you have already written in new research directions. And of course, browsing recent publications in Nature, Science, PLoS Biology, Genome Research, Nucleic Acids Research, PNAS, PLoS Computational Biology, the Journal of Computational Biology, PubMed, and Google Scholar is a great way to get ideas of recent research ideas, datasets, and results that you can expand upon for your project.

In addition, we have scheduled the following session to give you feedback on your ideas: Friday 10/29 at 4pm: MIT Stata Center G882. Postdocs and graduate students from the MIT Computational Biology group (comp-bio.mit.edu) will answer questions, discuss project ideas, and give advice and feedback. You are of course also welcome to discuss project ideas with the course staff during their office hours throughout the term.
Grading criteria

Your projects will be graded on overall quality, but to give you a sense of the criteria used, here’s a fictitious breakdown of points (the categories are not fictitious, the points are):

- **Originality** (~5pts) How original / novel is the idea? If people say “oh, that’s a good idea” when you describe your project, you’re all set! If they say “oh, just like that paper from so-and-so in 2001”, you must bring a new twist somehow!

- **Challenge** (~5pts) How challenging was the undertaking? If something was easy and you made it harder for yourself, you don’t get any extra points. If it was hard, but you made it easy by taking intelligent shortcuts, you still get full credit. Lastly, if you undertook something challenging in an area that’s novel to you, for the joy of learning something new, you get some extra points.

- **Relevance** (~5pts) Is this relevant to the course? Are you using things you learned about in the lectures/recitations/problem sets? Is this something that we could have used as an example in one of the lectures? Or is this only a vaguely justifiable tangential connection of something you were working on already?

- **Achievement** (~10pts) This is the big one. What did you actually accomplish in your project and what is your actual contribution to the field? If you promised us something amazingly original, super challenging, and central to computational biology, but didn’t deliver anything, you’re not in great shape (see formula below). If you managed to achieve the goals you set forward in your project proposal and milestones, and these are original and ambitious and relevant, you’re solid!

- **Presentation** (~10pts, oral and written) Lastly, research is not just about doing great work, it’s about being able to communicate it. **Written.** If you spent 15 pages describing in general terms why motif finding is important, but we can’t figure out what you actually did for your project, we obviously won’t be able to reward your contribution. If your paper is clear and crisp, you’re all set! **Oral.** Similarly, if at the end of your talk the entire class is wondering what that graph meant, or what your project was about, you’re not doing great. If you manage your time well, speak well, use clear slides and display items, and put some thought into your presentation, you’ll get more points on this, and we’ll be able to appreciate your entire work all the better!

Of course, despite the illusion of a point breakdown by the fictitious point assignments above, it’s really not an additive score, so the final formula may look more like this:

- **Total = Min(O,C,R) * A + P**

Lastly, these categories are just a guideline we use to make our grade assignments as objective as possible. Of course, an exceptional performance on any one category will probably spill over and make the overall score higher (just like a subpar presentation may suck some points out from the other categories as we won’t be able to understand your work).

Overall, the course staff will discuss every final project, every presentation, and your overall contributions. We’ll use a big fancy table and complicated formulae to come up with a final score, but will review that score thoroughly and make sure it reflects the overall quality of your project. We’ll give you feedback on the project, whether you ask before or after the end of the class (but we won’t disclose the individual scores for each rubric). Lastly, as you prepare to write up your work for publication, we can of course give you additional feedback in the future, and help you continue to be active members of the computational biology community.