Bringing Plant and Computing Scientists Together to Solve Grand Challenges

@

Cold Spring Harbor Laboratory
April 7-9, 2008

Sponsored by
The iPlant Collaborative:
a cyberinfrastructure-centered community
to enable a new plant biology

Monday evening (April 7)

Dinner 5:30 – 7 PM

7:00 Introduction and Overview, Rich Jorgensen (Director, The iPlant Collaborative)

7:30 Discussion

I. *Opportunities and Challenges* (Chair, Gregory Andrews, CS, Arizona)

7:45 Francine Berman, Director, San Diego Supercomputer Center,
Opportunities and Challenges in Cyberinfrastructure Development

8:15 Discussion

8:30 Break

9:00 Lee Allison, State Geologist & Director, Arizona Geological Survey
Building a Cyberinfrastructure Coalition - the Geoscience Information Network

9:30 Discussion

9:45 End of session

Tuesday morning (8:30 - noon)

II. *Compelling grand challenges* for the plant and computing sciences: working together (Chair, Steve Long, Illinois; facilitator)

8:30 Elliott Meyerowitz, Biology, CalTech
Defining solvable grand challenges in plant growth and development

9:00 Discussion

9:15 Eric Mjolsness, Computer Science, UC Irvine
Challenges in computational modeling of growth and development

9:45 Discussion

10:00 Coffee Break

10:30 John Willis, Duke University
Grand Challenges in Ecology, Evolution, Biodiversity and Organismic Biology

11:00 Discussion

11:15 Dennis Shasha, Computer Science, NYU
Dealing with Scale in Visualization and Machine Learning

11:45 Discussion

Noon – 1:30 Lunch

Tuesday afternoon (1:30-4:30pm)

III. *Discovery Environments*: prototype cyberinfrastructures for discovery, collaboration, and computational thinking in biology (Chair, Sudha Ram, MIS, Arizona)

1:30 Lincoln Stein, Bioinformatics, Cold Spring Harbor Laboratory; iPlant's Integrated Solutions Team and community Grand Challenge Teams: a collaboration to build prototype cyberinfrastructures ("Discovery Environments") to solve the community's most compelling grand challenges

1:50 Discussion

2:30 Brief Break

IV. *Break-out groups* (3 concurrent) Solving Grand Challenges: Data, Algorithms and Models

2:45 PM – 4:45 PM

1) *Primary metabolism, physiology, and organismic biology, from molecules to whole organisms*
Moderator: Gwen Jacobs, Montana State; Facilitator: Steve Rounsley, Arizona; Discussion panel: Steve Long, Gloria Coruzzi, David Salt, Sue Rhee, Carolyn Lawrence, Klaas van Wijk, Jodi Schwartz

2) *Evolution and development of form and function, from molecular, cellular and developmental biology to biodiversity, comparative biology, systematics and phylogenetics.*
Moderator: Mohan Tanniru, Oakland College; Facilitator: Kobus Barnard, CS, Arizona; discussion panel: Andrew Bangham, Edgar Spalding, Jody Banks, Todd Vision, Eric Kramer, Peggy Ozias-Akins, Bruce Kirchoff

3) *Biotic and abiotic interactions, from biochemistry and cell biology to ecosystems*.
Moderator: Laurie Kirsch, Pittsburgh; Facilitator: Doreen Ware, CSHL; discussion panel: Fumi Katigiri, Rob Last, David Rand, John Nason, Mark Schildhauer, Steve Slater

We request that a community member(s) in each session volunteer to be a note-taker.

Representatives from each group should be chosen at the end of the discussion period to compile a summary and identify a spokesperson to present a 10 minute oral summary in the evening session. Also, a written summary should be prepared by each group and posted on the conference wiki by the end of the conference.

Questions for discussion in break-out groups and on the conference wiki:

1) *What is the nature of a grand challenge that is both compelling and tractable?*
2) *What are the essential infrastructural datasets that iPC should ensure are available to the community for these grand challenges?* (Where are these datasets available? Would you like these to be linked to each other, and if so how? For these datasets, should iPC act as the primary repository, as a mirror, or as an interface? Which of these sets require curation? How frequently

should these sets be refreshed? What sort of documentation and metadata do you need to have for these datasets? How would you want the quality of the data to be documented?)

3) *What algorithms and analytic methods should iPC ensure are available to the community for these grand challenges?* (Are these currently available and if so where? If these methods/algorithms are computationally intensive, how should compute resources be allocated equitably?)
4) *What new types of algorithms, analytical models, and visualizations do you think are necessary to create in the iPC?* Can you summarize these and show how they may be useful to the plant biology community at large.
5) *Are there 'foundational tools and capabilities' that are needed to support diverse grand challenges?*
6) *How will education, outreach, and training be addressed in these grand challenges? What knowledge and skills do the next generation need to work on such multidisciplinary questions?*
7) *What sort of social networking tools do we need to enable effective collaboration?*
8) *How do we engage the computing research community to make them aware of cutting edge research opportunities in their disciplines?* How do we motivate computing researchers to participate in this collaborative?
9) *What do you think would help a diverse grand challenge team (diverse in terms of background and geography) collaborate effectively?*
10) *What computer-based tools do you think would help the grand challenge teams in developing their questions and pursuing their research?*

4:45 PM Wine and cheese party in foyer of Grace Auditorium

6PM Dinner

Tuesday evening (7:30 – 10pm)

Reports of breakout groups and discussion by participants: integration of perspectives (co-chairs, Lincoln Stein, CSHL, and Sudha Ram, Arizona)

7:30 Group 1 report by discussion panel leader

7:40 Group 2 report by discussion panel leader

7:50 Group 3 report by discussion panel leader

8:00 Discussion

9:00 Brief break

Group representatives should meet together as one group to plan compilation of a single summary, after the conference, to be posted on the conf wiki for community discussion

V. Keynote Speaker (chair, R. Jorgensen)

9:15 Richard Jefferson, CEO, CAMBIA BiOS Initiative
Vision and Strategy: Structuring grand challenges in plant science to help solve the grand challenges of society.

9:45 Discussion

Wednesday morning (9:00 AM - noon)

VI. Education, Training and Outreach: **Training the Next Generation of Biologists in Computational Thinking** (Chair, Susan Singer [iPlant EOT Adv. Comm. Chair]; Facilitator, Martha Narro [iPlant EOT Director])

9:00 Suzanne Westbrook, Computer Science, Arizona: Computational Thinking

9:10 Dave Micklos, Cold Spring Harbor Laboratory: Using CyberInfrastructure for EOT

9:20 Susan Singer, Carlton College: Integrating Research with EOT

9:30 Discussion

10 Coffee Break

VII. How will the community participate? What is the process?

Grand challenge workshops and teams; proposals; role of Board of Directors; Grand Challenge Foundational Tools

10:30 Rob Last, Michigan State (Chair, iPlant Board of Directors)
Steve Rounsley, Arizona (iPlant Management Team)

11:00 Discussion

Noon – Lunch

1:30 – 4 PM (Optional)

VIII. Self-forming discussion groups **(proto - grand challenge teams?**)

1:30 A number of rooms of various sizes will be provided - sign up sheets posted Tuesday to create and join groups - organized and run by the participants, facilitated by project personnel

4PM Departure

(Meals and accommodations available for participants departing Thursday morning, at additional charge through CSH Meetings)