March 2011 Update

Milestones

As of 2011 Feb 24, there were 41,414 batches completed and a further 3,151 being processed. As of 14/03/11, WCG returned 140,395,316 results on the HCC project – thank you all for your continued support and contribution.

HWI has output approximately 6,359 new batches and thus the updated number is 65,000 batches. We are also working on a new proposal for the Help Conquer Cancer II – focusing on systematically identifying molecular markers that would improve early diagnosis and prognosis prediction for multiple cancers.

Studying changes in a crystallization trial over time

As part of the effort of building a statistical model of protein crystallization, we recently reviewed 9,000 images from 1,501 crystallization experiments at multiple points in time, generating statistics of how a the state of a protein changes over time. This data is slowly being supplemented with results from the *Help Conquer Cancer* project. We are interested in the sequence of changes that occur in both crystallizing and non-crystallizing experiments. Tables 1 and 2 show the top five sequences discovered in both cases.

By building knowledge of these changes into our model of crystallization, our model can then assess the complete time-sequence of a protein-cocktail trial, using all images from the experiment as pieces of evidence. Figure 1 shows how this model divides crystallization trials into crystallizing and non-crystallizing outcomes.

Figure 1: How the time-series model divides crystallization trials by outcome

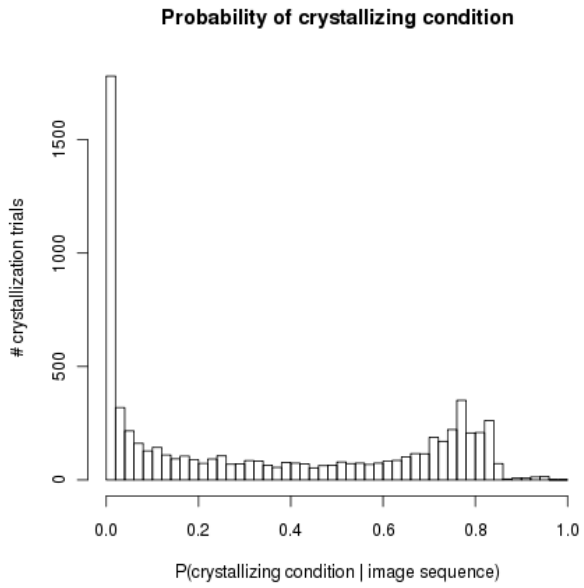


Table 1: Top five most frequent non-crystallizing reactions

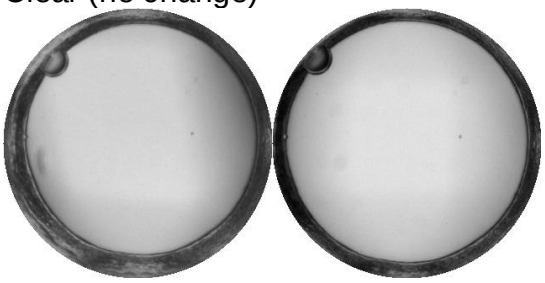
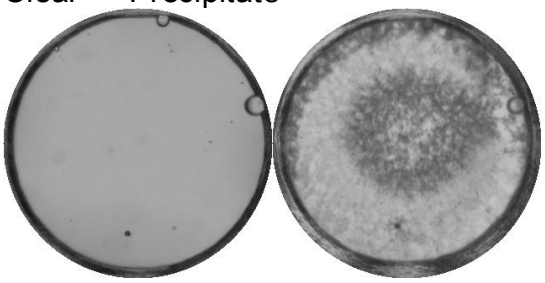
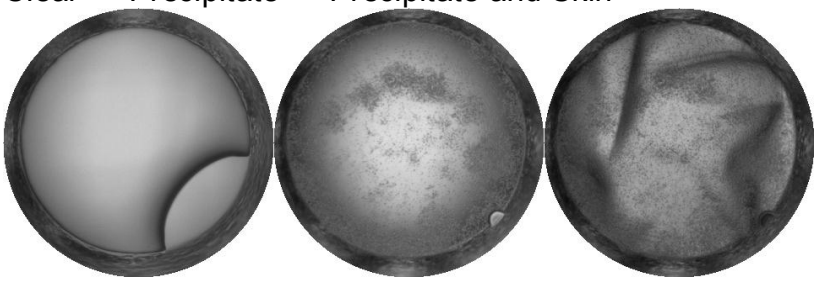
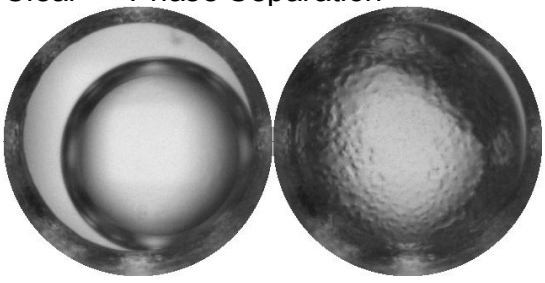
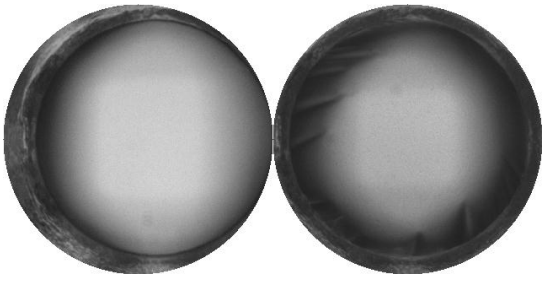
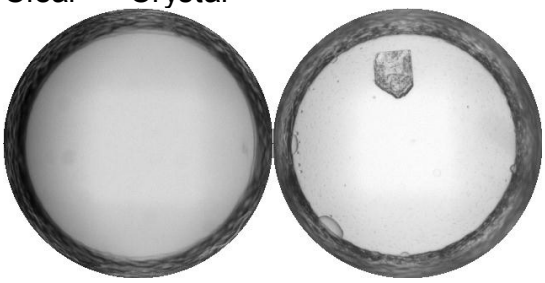
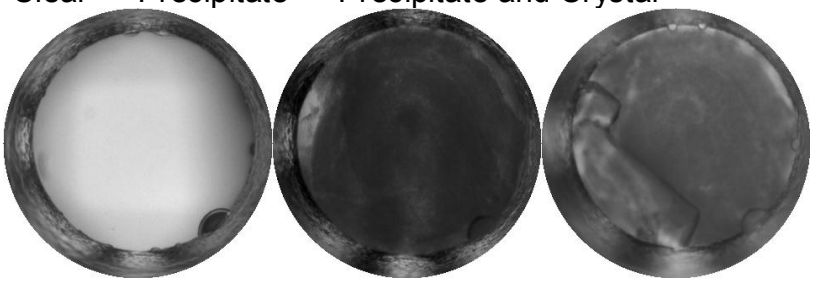
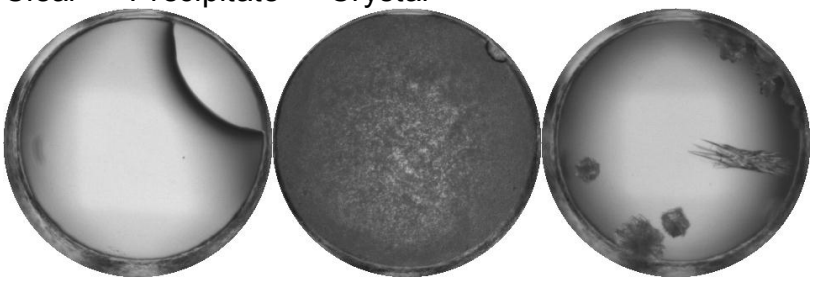
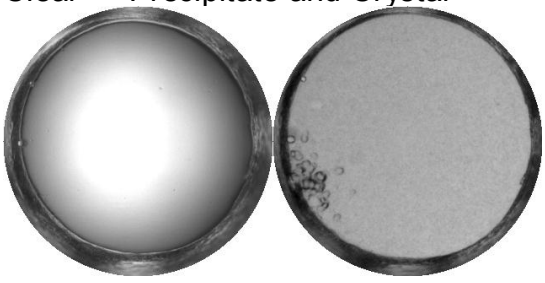
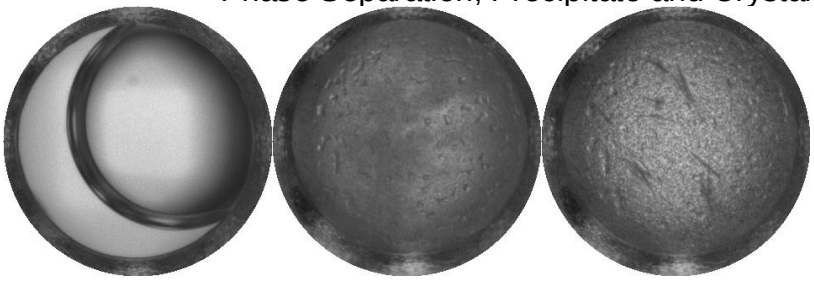
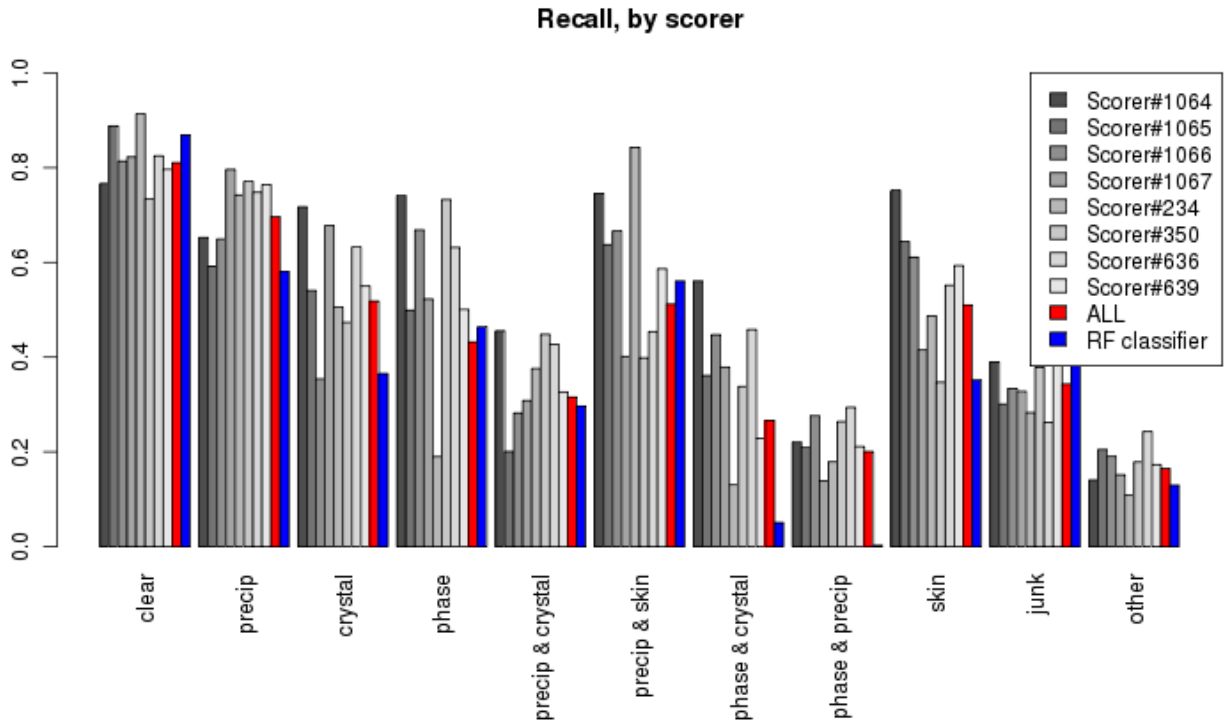
Clear (no change) 	35%
Clear → Precipitate 	29%
Clear → Precipitate → Precipitate and Skin 	8%
Clear → Phase Separation 	6%
Clear → Skin 	3%

Table 2: Top five most frequent crystallizing reactions

Clear → Crystal 	20%
Clear → Precipitate → Precipitate and Crystal 	17%
Clear → Precipitate → Crystal 	5%
Clear → Precipitate and Crystal 	4%
Clear → Phase Separation and Precipitate → Phase Separation, Precipitate and Crystal 	3%

Comparing the classifier to human experts

Image features computed on the Grid are used to build an image classification system. We measure the accuracy of our classifier by comparing its output against scores assigned by human experts. Some images are difficult to interpret, causing experts to score these images differently. We recently took a real-world sample of images scored by multiple experts, and measured our classifier against them (RF Classifier in the Figure below). The results show our classifier does a comparable job to the human experts in many categories (e.g., *clear*), with room for improvement in others (*phase separation and crystal*, *phase separation and precipitate*). The results also show that some categories are very difficult to judge, even by human standards.



Other progress

It is exciting to celebrate IBM Watson's triumph in Jeopardy. As one of the WCG projects, we will be benefiting from additional research support (http://www.worldcommunitygrid.org/about_us/viewNewsArticle.do?articleId=148). Thank you.

Other progress in our research is highlighted in a newly launched web site for the Cancer Gene Encyclopedia project: http://ophid.utoronto.ca/cgep. We have released version 1 of our portal for microRNA:gene target prediction – mirDIP – http://ophid.utoronto.ca/mirDIP.

A very exciting new results in pancreatic cancer research have been recently published in *Cancer Research:* Chang, Q., Jurisica, I., Do, T., and Hedley, D. W. Hypoxia predicts aggressive growth and spontaneous metastasis formation from orthotopically-grown primary

xenografts of human pancreatic cancer, *Cancer Res*, 2011, In press. Summary of other publications can be found at http://www.cs.utoronto.ca/~juris/publications.htm.

Thank you,

C. A. Cumbaa, Yulia Kotseruba and I. Jurisica