SC_20100629

iPG2P Steering Committee Meeting
June 29, 2010; 1 to 2 pm CDT

Present: Chris Jordan, Matt Vaughn, Dan Stanzione, Jeff White, Steve Goff, Karla Gendler, Dan Kliebenstein, Ruth Grene

Notes/Agenda:

Item 1: Review Action Items
None

Item 2: CI Development Update
Vaughn gave an update on the progress of the prototype work being done on "Ruth's workflow". Using Twiddla, he demonstrated the scope of work for Ruth Grene, Greg Abram, Karla Vega (and Dan). Lecong Zhou has written a reference implementation of the workflow in R but it is very linear and not generic. Array analysis is out of scope for this and thus the starting point is a list of enriched or interesting genes. There will be support for Aranet and hopefully GeneMania. The workflow could be implemented in Vistrails, Taverna, etc. The plan is for week 6-7 of development to make more generic (add more components) and make parallel (branching patterns).

Stanzione updated the SC on the status of the release scheduled for this week. We are currently fixing a round of bugs on the UHTS workflows. The site has been taken down for maintenance and upgrade and could be up later tomorrow or Thursday morning. We will then open it up for friendly user testing. Vaughn added that we will have test data for the DE and he is working on basic tutorials for a user to get a feel for what has been developed. For iPTol, there has been a major interface overhaul. With the new UI, it feels more like being guided through the workflows and a user can do things simultaneously.

Item 3: GLM Update
Vaughn reported that Susan Lindsey has stub code checked into the code repository and we are looking good for having something for ASPB. For the 2-way epistasis test, the entire problem set may be constrained by I/O issues (i.e. read/write to disk). Liya Wang has implemented up to forward model selection in MatLab and is working on epistasis now.

Gendler reported that the GPU group has implemented forward regression with multiple SNPs in CUDA. They have evaluated the performance based on 100K SNPS each of length 191 and execution time is 62ms which includes data transfer to the GPU and back to the host machine, but it doesn't include reading data from the input file on the host machine. Their next step is to implement the pairwise version. They need to find a way to visualize the output and Stanzione suggested that using Liya's MatLab code would be fine. It would provide verification of the results also.

Next meeting time: Tuesday, July 13, 2010, 2pm EDT

Adjournment: