

Agriculture and Agri-Food Canada - GCP Vancouver Meeting minutes

Minutes of the video conference between AAFC and SP4 members of the
GenerationCP, August 17, 2005

(prepared by Fran Clarke)

Contents

- [1 Participants](#)
- [2 ICIS \(International Crop Information System\)](#)
- [3 Generation Challenge Program \(GCP\)](#)
- [4 Current Technologies Being Worked On](#)
 - [4.1 David DeKoeber](#)
 - [4.2 Travis Banks](#)
 - [4.3 Nick Tinker](#)
 - [4.4 Weikai Yan](#)
 - [4.5 Graham McLaren](#)
 - [4.6 Akin Akintunde](#)
 - [4.7 Masaru Takeya](#)
 - [4.8 Richard Bruskiewich](#)
 - [4.9 Guy Davenport](#)
 - [4.10 Manual Ruiz- CIRAD](#)
 - [4.11 Simon Reinhard](#)
 - [4.12 Thomas Metz](#)
- [5 Future Plans](#)
- [6 Other](#)

Participants

- **AAFC:**
 - **Swift Current:** Fran Clarke, John Clarke, Ron Knox Shawn Yates; Brad Meyer
 - **University of Saskatchewan:** Curtis Pozniak
 - **Winnipeg:** Travis Banks
 - **Fredricton:** David DeKoeyer
 - **Ottawa:** Nick Tinker, Weikai Yan
- **GCP via Burnaby:**
 - **IRRI** (International Rice Research Institute), Philippines: Richard Bruskiewich, Graham McLaren, Thomas Metz;
 - **CIMMYT** (International Maize and Wheat Improvement Center), Mexico: Guy Davenport;
 - **ICARDA** (International Center for Agricultural Research in the Dry Areas), Syria: Akin Akintunde
 - **NIAS** (National Institute of Agrobiological Sciences) Japan: Masaru Takeya
 - **CIRAD** (Agricultural Research Centre for International Development/ Centre de coopération internationale en recherche agronomique pour le développement, [\[1\]](#)) France: Manual Ruiz
 - **CIP** (International Potato Centre), Peru: Simon Reinhart

ICIS (International Crop Information System)

- Nick Tinker feels we should consider a place for work on crop information systems as we develop new outcome projects. This would help us develop common platforms for data mining and data exchange both within AAFC and with clients and partners. It would also make help us participate in International collaborations in systems design. This is an important area that needs to be part of our science strategy.
- Many researchers within AAFC have either developed their own information system, or are using a commercial package, such as AgroBase

- Akin indicated that at ICARDA, they use a combination of AgroBase and ICIS for their needs. The raw evaluation data from their field trials is entered with the ICIS workbook into ICIS after using AgroBase to do the statistical analysis.
- A variety of different analysis tools and information systems can be used within individual research programs, and the data can easily be tailored to go into ICIS to provide a safe, central repository of data for further use and access by multiple users
- There is a need for bioinformatics experts, developers and administrators in order to implement a full scale information system such as ICIS for AAFC.
- Graham indicated that the new alliance between IRRI and CIMMYT will help develop more ICIS tools for wheat and maize.
- The modeling layer of the Generation Challenge Program (GCP) is being built into ICIS, to ensure interoperability between them.
- David commented that they are interested in raw data collection. Graham explained that ICIS can be used to integrate data from the field. He used Nunhems as an example, who use ICIS for data collection, and then 3rd party tools for statistics, making labels, etc.
- Akin is working on the LIMS (Laboratory Information Management System) project and will send Shawn a copy plus post it to CropForge.

Generation Challenge Program (GCP)

- Collaboration with the GCP:
 - Nick indicated that he sent the minutes of the last meeting to Campbell Davidson
 - Campbell responded that he is attending a Generation Challenge Program meeting in late summer
 - There is a suggestion that AAFC fund a programmer at another institution such as IRRI to work on tools of joint interest. Another possibility is an MII-type project, where the GCP would provide a grant and AAFC would match it, with the activity conducted in Canada.
 - Graham mentioned that there are many ways that collaboration between AAFC and the GCP can occur, and that they will know more about the possibility of AAFC being able to lead a GCP

project in the fall. (perhaps at the GCP Annual Review meeting in Rome, Sept. 29 – Oct.1).

- The next round of competitive grants for the GCP are in the spring of 2006
- Graham and Richard said that there has been concern among the GCP representatives about the absence of AAFC and Canada. AAFC would probably be a welcome partner.
- The GCP is looking for interoperability between information systems
- The GCP is being built as a framework of open source tools.
- Nick expressed an interest in using GDPC (Genome Diversity Phenotype Connection) from Ed Buckler's lab at Cornell to make use of data exchange between database systems.
 - He finds that GDPC browser is a useful front-end
 - Richard indicated that GDPC is a specialized protocol for a subset of domain data types, as opposed to the broader scope, general models being developed by the GCP
 - There is overlap between GCP and GDPC.
 - Terry Casstevens of GDPC was funded several times to participate in GCP meetings (as a Cornell representative, Cornell University being a partner in the GCP).
 - GDPC has hard coded modules that are not updated easily
 - GDPC does not cover gene models
 - The GCP has worked on integrating many of the models from the GDPC into their own. (Terry's GDPC models were one starting point for the GCP modeling effort)
 - The developers of GDPC have met with representatives from the GCP about collaboration but at this point they seem to have different funded goals so relatively little formal progress toward convergence of efforts has taken shape. As result, the GCP has moved forward on their broader agenda, as they cannot wait. However, GCP programmers are working on establishing a loose coupling (mapping) of GDPC into the GCP platform.
- The GCP has been using Eclipse [\[2\]](#) for some of its software design. It is an open-source software designing development platform, which defines a Rich Client Platform (RCP) that makes development of interoperable and

extensible applications very easy. In addition to being a software engineering development environment, the underlying software (user interface) framework (“Eclipse/RCP”) can be used to develop third party applications. Many of the projects being developed by the GCP are going to be implemented as Eclipse RCP plug-ins when completed.

- The GCP have embedded MOBY web services interoperability (networking) technology [\[3\]](#) into the GCP platform.
- Information pertaining to the GCP and the GCP Models are listed on the CropWiki [\[4\]](#) with the latest UML diagrams and associated classes posted to Cropforge, under the GCP Middleware project, Demeter-DMS subproject.

Current Technologies Being Worked On

David DeKoefer

- Has been working on developing BioGroupWare for phenotypic data and also breeding projects
- Used for managing genomic projects
- Receiving help from training graduates at the University of New Brunswick on creating tools
- He is working on developing a potato information system
- Technical support is a constraint
- Has used data models in design phase ???
- Could be possible collaboration with the CIP (Simon Reinhart’s group)
 - DEBA software for refractions ???
 - Could integrate their common tools

Travis Banks

- Working on a WGRF project with Darryl Somers
 - 300 Canadian hexaploid wheat varieties

- 300 SSR markers on each
- Has yet to start coding anything that is database specific
- Has been working with Darryl Somers on a comparative mapping tool that is not static like C-map
 - Is a web-base SVG comparative mapping tool that is GDPC and CMAP compatible
 - Choose maps and drag on the screen
- Developed a web-based database/software tool for cataloging gel photos.

Nick Tinker

- Primary biological interest is in Oats
- Developed a comparative map viewer for oat maps
 - Works well for their program, but not published
- Borrowed the ICIS philosophy to create his own database system. This includes an oat pedigree database (avena.agr.gc.ca), a phenotype database (DUDE – see below), and a genotype database with a web-base front end for browsing or populating virtual gels (GELATO). These systems are “simplified” from ICIS and may provide a stepping stone for a user to develop a full ICIS implementation.

Weikai Yan

- Developed DUDE
 - Query on various variables in database
 - Entry level ICIS-like database that could be imported to ICIS
- Biplot analysis
 - g x e analysis
 - Meta Data descriptions
 - Developed “GGE Biplot software”

Graham McLaren

- Coordinates the ICIS community
- Managing crop improvement and phenotypic information
- Coordinates the GCP Platforms
- Bioinformatics – how to make better use of the information in the databases (i.e. estimation of improved breeding values...)

Akin Akintunde

- Administers ICIS for 5 different crops (barley, lentil, durum, breadwheats and chickpea)
- GCP – organizing molecular information from these crops

Masaru Takeya

- Working on the functional genomics domain model for the GCP
- NIAS is leading a GCP project for the development of a comparative crop gene expression repository with data mining tools

Richard Bruskiewich

- Bioinformatics of rice management
- Comparative cereal analysis of rice, wheat and maize
- Domain modeling
- MOBY
- Comparative stress gene catalogue
- Collaboration with Mussaro

Guy Davenport

- Comparative genomics

- GCP – data format for input into database
 - Genotyping and mapping data
 - Looking at ICIS and GERMINATE [\[5\]](#) for storage

background:

Germinate was funded by BBSRC in attempt to make use of GENE-MINE products as the GENE-MINE database

GENE-MINE was funded by the EU with a remit to produce a database for genotyping and phenotype data. NCGR based the GENE-MINE database on ICIS and a new model for generic genotyping data.

The GENE-MINE grant ended before the database was complete. However Germinate was already funded, so towards the end of the GENE-MINE we took the best ideas from GENE-MINE and made Germinate.

Both Germinate and ICIS have similar generic ways of storing data. However, germinate relies heavily on Postgres features, where as ICIS is fairly DBMS independent.

CIMMYT/IRRI is currently working through a comparison of Germinate and ICIS for genotyping data, which we hope will result improvements in ICIS if necessary. Thomas Metz is in charge of this project and it is documented on the ICIS wiki at [\[6\]](#)

- Developing new Eclipse RCP tools for visualization of maps and traits (e.g. new version of CMTV)
- NCGR and CIMMYT have developed a Java web-based comparative mapping tool (CMTV) that can be found at [\[7\]](#). CIMMYT has plans to develop a new version of CMTV with Eclipse RCP that will be integrated with GCPC (the middleware software, not necessarily the GDPC browser) and the GCP platform CIMMYT is developing.

Manual Ruiz- CIRAD

- Genotype and mapping
- Mapping and modeling sequence maps
- Connect MOBY
- Comparative genes ???

Simon Reinhard

- Biometrics and bioinformatics
- Genebank and bio ???

Thomas Metz

- Migrate IRRI genebank database system with IRIS
- Collaborative platforms (ICISWiki at CropWiki [\[8\]](#); CropForge [\[9\]](#))
- Institutionalizing the recognition of open source software
- Management of internal data

Future Plans

- Plant and Animal Genome Meeting in San Diego in Mid-January
 - Workshop could be held either before or after this meeting
 - Vancouver would work best