

---

**Generation CP Subprogramme 4**  
**Platform Task 28-10**  
**(Functional) Genomics Tool Integration**

---

Richard Bruskiwich,  
Marcos Costa, Masaru Takeya



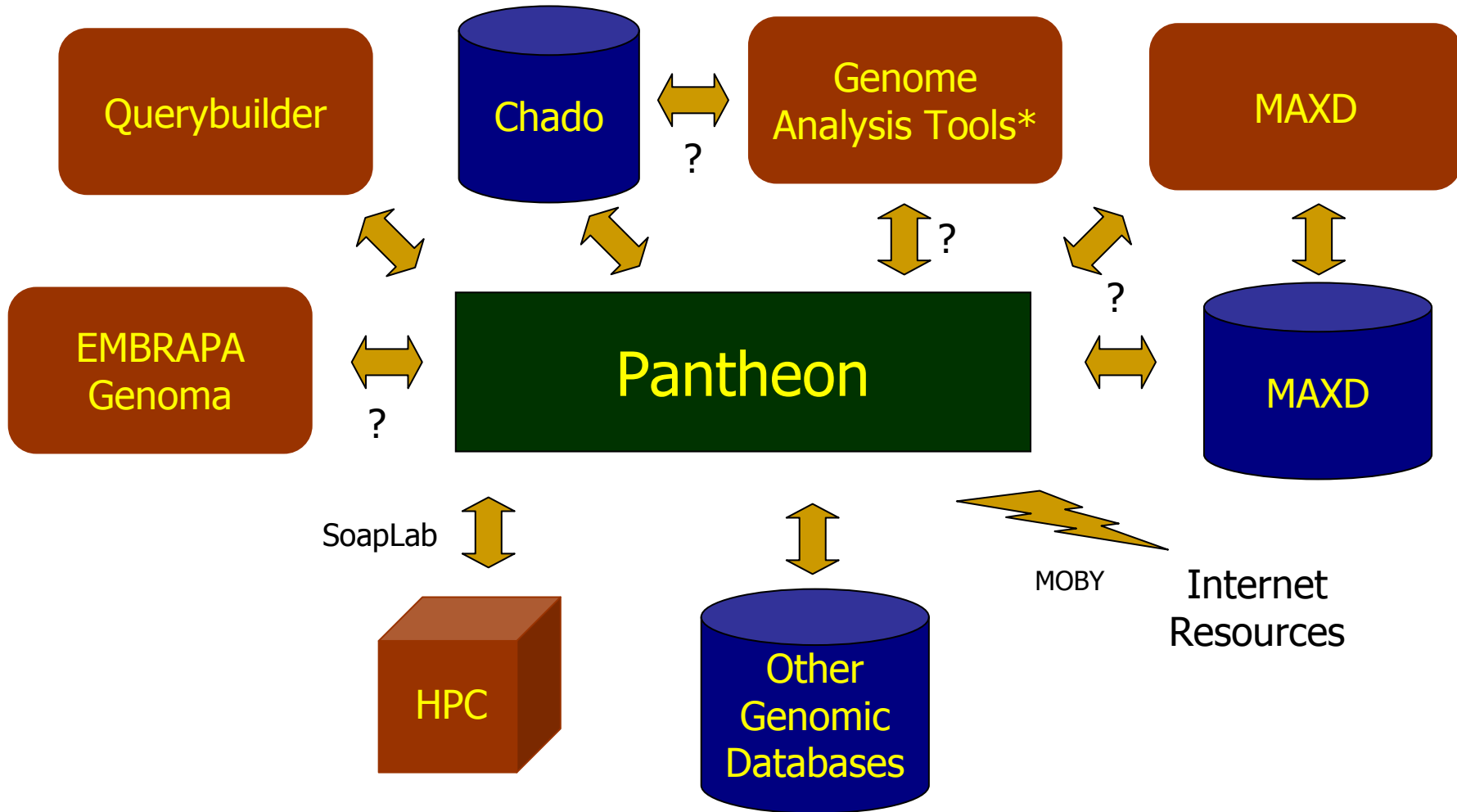
---

# Outline

- General overview of project 28-10
  - Sequence inventory tool (overlap with task 31)
  - GMOD schema (Chado) & tool integration; other tools (slight overlap with project 28-5)
  - HPC/SoapLab (overlap with task 27)
- Microarray data repository & tools (Masaru)
  - MAXD development (overlap with task 32)
- Genomics workbench development (Marcos)
  - EMBRAPA Genoma adaptation

Project Leader	Richard Bruskiwich (IRRI)
Collaborator(s)	Natalia Martins, <i>Marcos Costa</i> (EMBRAPA) + <i>Georgios Pappas</i> <i>Masaru Takeya</i> (NIAS)
Project Title	Standalone Tool Adaptation for Functional Genomics
Rational	Java software tools exist for functional genomics– GMOD, TIGR, MAXD and EMBRAPA (Genoma) – that should be adapted as plug-ins to the GCP platform. This will require some remedial Java software engineering. Although the GCP ortholog gene catalog and gene expression repository tasks are delivering a portion of this need, there remains some software adaptation work for specific tools not adequately covered by these two other tasks.
Project Category	X Standalone application integration
Deliverable Product(s)	Specific standalone tools targeted for integration into GCP platform: <ul style="list-style-type: none"> <li>• The Apollo genome browser tool from GMOD</li> <li>• The TIGR microarray data management and analysis tools, especially, the MADAM LIMS tool and TIGR Multiple Experiment Viewer (MeV); the Manchester “MAXD” tool suite will also be assessed for use in the project</li> <li>• The EMBRAPA “Genoma” tool suite</li> </ul>
Proposed implementation strategy & technology	<ul style="list-style-type: none"> <li>• Use cases and deficiencies (of existing tools relative to GCP needs) applicable to target tool integration will be documented</li> <li>• Documentation and source code for target tools will be examined to identify current tool-to-data source integration Java interfaces and classes; technical design requirements for (plug-in) tool integration with GCP middleware will be identified</li> <li>• Remedial Java software engineering will be undertaken to implement tool integration into GCP platform</li> </ul>

# Vision



---

# Chronology

- Project coordination discussions held throughout year in Netherlands, Philippines, Japan
  - Data template, database, tool integration and web services
- Several University of Philippines students engaged in software development (main target of IRRI funds...)
  - Sequence inventory tool, HPC, MAXD, Apollo
- Martin Senger assisted SoapLab and MOBY services development for IRRI, EMBRAPA and NIAS
- MAXD (mostly task 32) – report by Masaru
- Genoma – report by Marcos

---

# (IRRI) Project Status

- Sequence data template designed
- Prototype sequence inventory tool (using template) being used by GCP gene catalog postdoc & being refined by responsible student (using SO/GO ontology Chado database); can connects to NCBI for sequence data
- GCP genomics model refined for implementation in Pantheon
- Chado database wrapping started (maybe CIRAD should complete?)
- Students working on GMOD tool integration (... slowly...)
- Paracel BLAST and EMBOSS wrapped with SoapLab on HPC; platform clients under development
- Progress/strategy/implementation partly documented on the GCP Wiki

# Future Work

(April-May IRRI)

- UP students on “summer break” and have pledged to pay more attention to the project, hopefully leading to more/refined prototypes of tool integration
  - Sequence inventory tool refined/debugged by end user feedback from GCP postdoc
  - Chado wrapped further with GCP domain model using Hibernate and applied to tool integration task
    - Querybuilder (general searches on genomic data?)
    - GMOD tools: Apollo(?)
    - JalView(??)
  - Soaplab client for HPC connection integrated into platform?
  - MAXD integration with NIAS (maybe via ISYS?)

---

# Functional Genomics Experiment (FUGE) and Functional Genomics Ontology (FUGO)

- Public standards for experimental modeling and ontology-driven documentation of experiments
- [fuge.sourceforge.net](http://fuge.sourceforge.net) and [fugo.sourceforge.net](http://fugo.sourceforge.net)