SHORT COMMUNICATION

RGMIMS: a web-based Laboratory Information Management System for plant functional genomics research

Leakha Henry · Kerrie Ramm · Qian-Hao Zhu · Narayana M. Upadhyaya

Abstract A robust Laboratory Information Management System (LIMS) is required for the efficient handling of data generated from large-scale insertional mutagenesis projects. The Rice Gene Machine Information Management System (RGMIMS), a web-based modular LIMS, developed in a rice functional genomics laboratory at CSIRO, currently has four core modules: Seed Management, Transformation Management, Plant/Progeny Management, Phenotype Management, and an ad hoc querying module. RGMIMS manages, preserves and tracks large inventories of transgenic germplasm and enables efficient and accurate record keeping of the large quantities of experimental data. RGMIMS automates and seamlessly integrates multi-step experimental processes. A web user interface, incorporating barcoding utilities, enables rapid data capture and tracking of biological resources. Ontologies from Gramene and Plant Ontology consortium are used to describe mutant phenotypes. RGMIMS supports generic research processes in plant mutagenesis and could readily be adapted to general LIMS for high-throughput plant research.

Keywords Insertional mutagenesis · Laboratory Information Management System (LIMS) · Ontology · Phenotype · Rice Gene Machine Information Management System (RGMIMS) · Rice · Transformation · Web application

Abbreviations

LIMS	Laboratory Information Management System
PACLIMS	Phenotype Assay Component LIMS
RGMIMS	Rice Gene Machine Information Management System
TPM	Transgenic plant monitor

L. Henry (✉)
CSIRO Mathematical and Information Sciences, ANU Campus, Canberra, ACT 2601, Australia
e-mail: leakha.henry@csiro.au

K. Ramm · Q.-H. Zhu · N. M. Upadhyaya (✉)
CSIRO Plant Industry, Cnr. Clunies Ross Street and Barry Drive, GPO Box 1600, Canberra, ACT 2601, Australia
e-mail: narayana.upadhyaya@csiro.au

L. Henry · K. Ramm · Q.-H. Zhu · N. M. Upadhyaya
New South Wales Agricultural Genomics Centre, Wagga Wagga, NSW 2678, Australia

Results

Availability and contacts

RGMIMS is available freely to public-sector laboratories via a standard material transfer agreement with CSIRO. Demonstration software and supplementary information are available through http://www.pi.csiro.au/fgrttpub/rgmims.htm. Either the first or the last authors can be contacted for further information.

🖄 Springer

RGMIMS platform

RGMIMS is implemented in Java (http://www.java. sun.com), a platform-neutral open-source software, and developed on the Microsoft Windows's platform (http://www.microsoft.com). It has been tested using Windows XP and SUSE Linux operating systems, with Internet Explorer browser 6.0 and Mozilla FireFox 2.0 client browsers. RGMIMS uses a client/ server architecture: a rich browser client (a Java Applet and Java Server Pages, http://java.sun.com), Apache Tomcat server (with a Java servlet acting as a broker, http://tomcat.apache.org) to execute the business logic and mediate the communication between the client and the back-end relational database (Microsoft SQL Server 2000, http://www.microsoft.com).
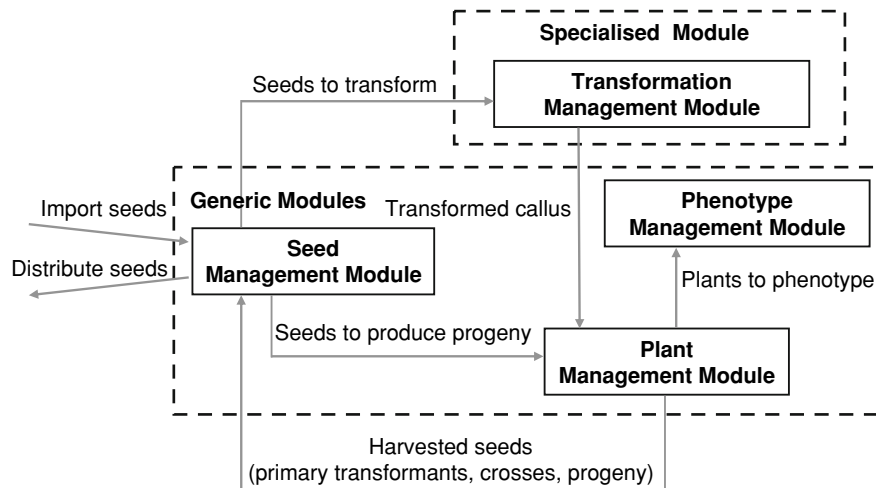

Fig. 1 RGMIMS has four core modules: Seed, Transformation, Plant and Phenotype. Each module supports a related set of functionality used at different stages of the operational insertional mutagenesis workflow. The modules are integrated and share information between them to reduce data redundancy (the arrows show the interconnections between the modules and the flow of information). However, the modules are loosely coupled so that each can be used as a stand-alone such as the seed module, or in combination such as the seed and plant modules or the seed, plant and phenotype modules
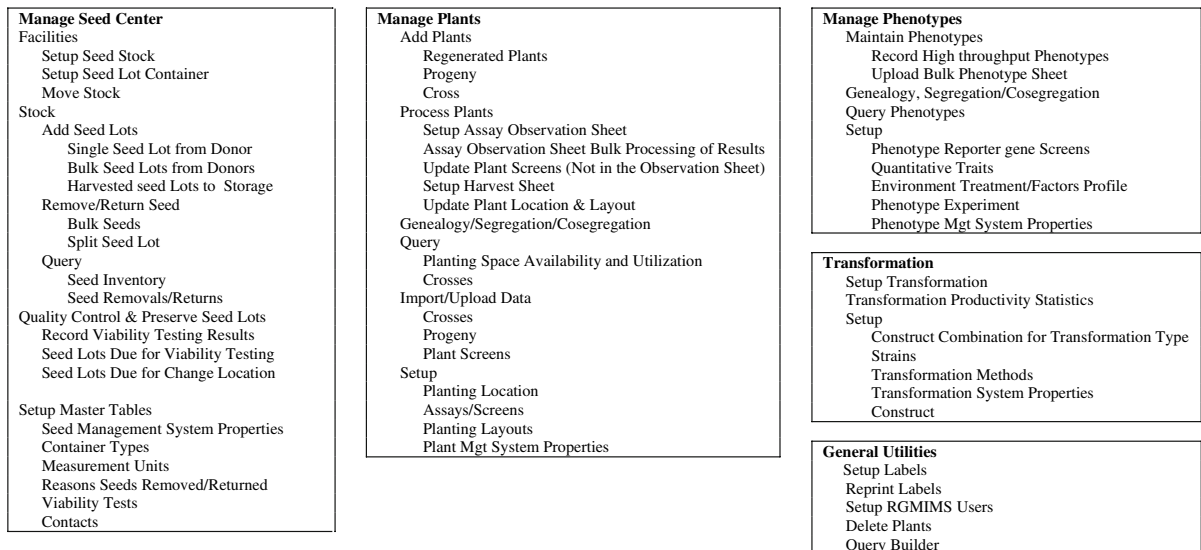
Manage Seed Center
Facilities
  Setup Seed Stock
  Setup Seed Lot Container
  Move Stock
Stock
  Add Seed Lots
    Single Seed Lot from Donor
    Bulk Seed Lots from Donors
    Harvested seed Lots to Storage
  Remove/Return Seed
    Bulk Seeds
    Split Seed Lot
  Query
    Seed Inventory
    Seed Removals/Returns
Quality Control & Preserve Seed Lots
  Record Viability Testing Results
  Seed Lots Due for Viability Testing
  Seed Lots Due for Change Location

Setup Master Tables
  Seed Management System Properties
  Container Types
  Measurement Units
  Reasons Seeds Removed/Returned
  Viability Tests
  Contacts

Manage Plants
Add Plants
  Regenerated Plants
  Progeny
  Cross
Process Plants
  Setup Assay Observation Sheet
  Assay Observation Sheet Bulk Processing of Results
  Update Plant Screens (Not in the Observation Sheet)
  Setup Harvest Sheet
  Update Plant Location & Layout
Genealogy/Segregation/Cosegregation
Query
  Planting Space Availability and Utilization
  Crosses
Import/Upload Data
  Crosses
  Progeny
  Plant Screens
Setup
  Planting Location
  Assays/Screens
  Planting Layouts
  Plant Mgt System Properties

Manage Phenotypes
Maintain Phenotypes
  Record High throughput Phenotypes
  Upload Bulk Phenotype Sheet
Genealogy, Segregation/Cosegregation
Query Phenotypes
Setup
  Phenotype Reporter gene Screens
  Quantitative Traits
  Environment Treatment/Factors Profile
  Phenotype Experiment
  Phenotype Mgt System Properties

Transformation
Setup Transformation
Transformation Productivity Statistics
Setup
  Construct Combination for Transformation Type
  Strains
  Transformation Methods
  Transformation System Properties
  Construct

General Utilities
  Setup Labels
  Reprint Labels
  Setup RGMIMS Users
  Delete Plants
  Query Builder

Fig. 2 Menu structure of integrated modules of RGMIMS

Key features of RGMIMS

RGMIMS has four core modules—Seed Management Module, Transformation Management Module, Plant Management Module and Phenotype Management Module (Fig. 1). The available menu structure is presented in Fig. 2.

Seed management module

This module manages information about seed receipt, handling, storage, inventory levels, quality and preservation. Seed acquisition is recorded and seeds are

assigned a storage location (Fig. 3). Accession number, passport information, taxonomy of imported seeds are captured. NCBI's taxonomy (http://www.ncbi.nlm.nih.gov) is used to standardise the description of plant species. Seed storage facilities are organised hierarchically in terms of stores, containers and packets. The facilities can be reorganised graphically using "drag and drop" features. The seed inventory provides information on seed availability and an audit trail of seed movements. Seed lots can be split into sub-lots, each with a unique accession and storage location. Seed viability review results are recorded to ensure the quality of seed in storage; RGMIMS automatically flags any seed lots due for

Fig. 3 Seed acquisition from donors. (**a**) Details of the imported seed lot are recorded, such as donor accession number, passport information and imported quantity; (**b**) a taxonomy is selected from the NCBI Taxonomy browser to describe the species of the imported seeds; (**c**) a donor is selected from an electronic diary of contacts; (**d**) a storage location is assigned to the seeds from the pick list of customised stores and containers. A measurement unit is selected from the pick list of customised units (not shown)

quality control based on seed age and previous viability results. To preserve seeds in storage and prevent their deterioration and loss, the system automatically flags seed lots due for movement to more ideal store rooms, based on seed age and the customised period for each store.

Transformation management module

This module handles information about the transformation systems used to produce genetic modifications in the plants under study (Fig. 4). It records details of tissue culture when it is a component of the transformation strategy. Details and results of transformation protocols are recorded, including gene constructs, selected callus/plant lines. The efficiency of the tissue culture processes and transformations

can be easily monitored, from the initiation of callus to the number of transgenic lines obtained.

Plant management module

This module manages information about the generation, screening, monitoring and harvesting of the transgenic as well as non-transgenic plants. The production of regenerated lines and the propagation of plants through crossing and selfing are recorded. An 'observation sheet' is used to screen and record the results of a variety of assays that are applied to the lines. A 'harvest sheet' is used to monitor plant harvest, to print barcode seed packet tags for harvested plants and to record the yields of seed lots from single or pooled plants. A pedigree tree is dynamically generated showing plants, parents,

Fig. 4 Transformation management. (**a**) New transformation details (constructs, transformation method, target tissue/quantity) are recorded with logged user as "person handling" as default; (**b**) once entered, existing transformation can be retrieved to update/edit transformation progress; (**c**) productivity statistics can be retrieved for different transformations, callus/plant lines, persons, dates or constructs
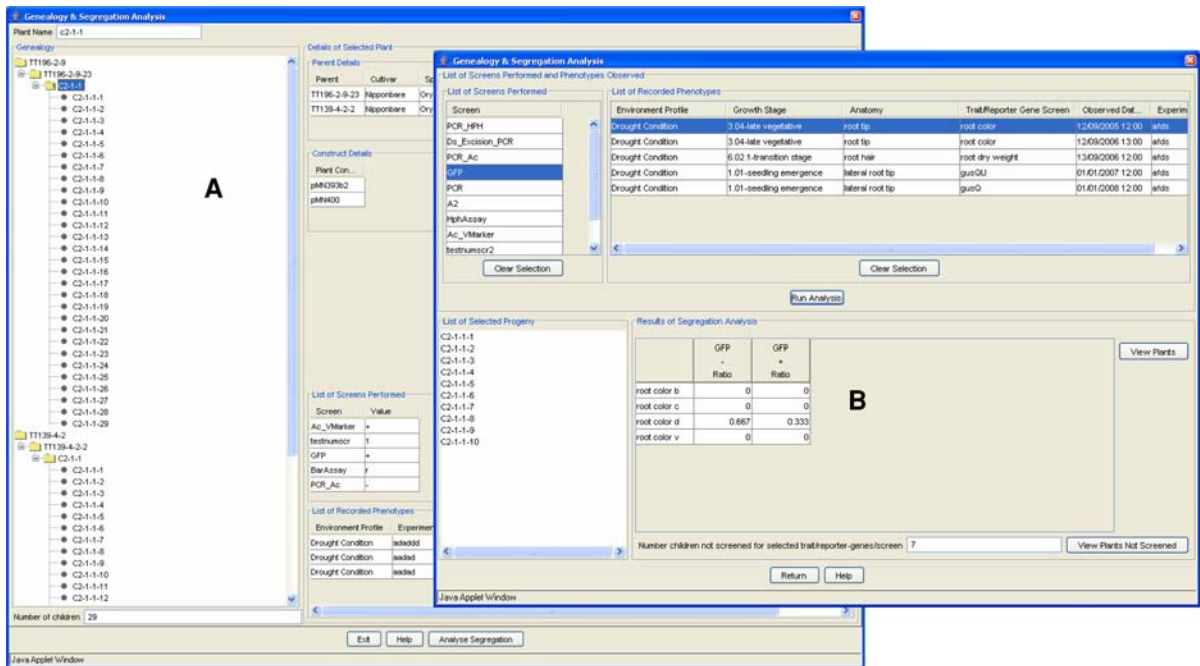
Fig. 5 A pedigree tree. (**a**) Parents of F$_1$ plants, sibling(s) and progeny are displayed on the left-hand panel and the details of molecular tests and phenotypic screening are displayed in the right-hand panel; (**b**) the results of a co-segregation analysis of the selected progeny, involving a qualitative molecular screen and qualitative trait, is displayed

sibling(s) and progeny, as well as details of any molecular tests and phenotypic screening (Fig. 5). This tree permits segregation analysis where inheritance pattern is summarised.

RGMIMS captures the hierarchical organisation of the growth facilities in terms of geographical location, plots/glasshouses and rows/benches. RGMIMS enables efficient space usage of the growth facilities. For instance, newly produced lines can only be located in growth facilities with available space and the occupied space within the growth facilities is released automatically when plants are processed as harvested, dead or discarded. Reorganisation of plants within the growth facility is permitted, and including space reservation in advance.

Phenotype management module

This module handles the phenotypic characterisation of the transgenic plants, where many screens with a variety of parameters are carried out on every plant. To facilitate the sharing of information with collaborators and the wider research community, a combination of generally accepted biological ontologies from Gramene (http://www.gramene.org; Jaiswal et al. 2006) and Plant Ontology Consortium (http://www.plantontology.org; The Plant Ontology^{TM} Consortium 2002) are used to describe phenotypes in standardised terms. A 'Phenotype sheet', is used to characterise and evaluate transgenic plants (Fig. 6). It captures the results of plants screened for various physical and molecular traits, in different plant parts/organs at different developmental stages described using the relevant ontologies.

Discussion

Large-scale plant insertional mutagenesis projects, such as the CSIRO Plant Industry's Rice Gene Machine Project (http://www.pi.csiro.au/fgrttpub/, Upadhyaya et al. 2006), produce vast amounts of diverse experimental data and resources. In our experience, data management with paper records, electronic spreadsheets and database tables (Microsoft Access^{TM} and SQLServer 2000^{TM}) were inefficient, produced poor quality data and afforded

Fig. 6 A 'Phenotype sheet'. This sheet is used to perform large-scale characterisation and evaluation of transgenic plants: (**a**) The results of plants screened for various physical and molecular traits, at various anatomies, developmental stages and time points described using the relevant ontologies are captured; (**b**) the traits and reporter screens to be observed at various growth stages and anatomy are set up; (**c**) the ontology describing the trait to be observed is selected from the Plant Ontology Consortium's 'Trait Ontology'; (**d**) the selected trait ontology can be visualised as a tree

limited accessibility and information sharing. Integrating and managing experimental data and resources present major challenges for these projects. Several Laboratory Information Management Systems (LIMS) have been developed to address these challenges, (Scott et al. 2003; McLaren et al. 2005; Donofrio et al. 2005); however, only Phenotype Assay Component LIMS (PACLIMS, Donofrio et al. 2005) offers a web-based, platform-independent solution. These systems do not appear to possess comprehensive functionality to manage the data and resources generated throughout the insertional mutagenesis process—from transformation and tissue-culture, through breeding, molecular screening, characterisation and harvesting, to the preservation

and quality control of resources, and integration of commonly-accepted ontologies describing phenotypes and taxonomy. The RGMIMS has been developed in close collaboration with its users: research scientists, laboratory managers and technicians. Via the web, RGMIMS manages and tracks hundreds of thousands of biological resources in a collective repository, provides comprehensive electronic record management, ensures easy access to a large body of data, and achieves automation and seamless integration of insertional mutagenesis workflow.

RGMIMS is a multi-user system with access controlled by username and password. Graphical user interfaces, integrated with barcode scanners and electronic weighing scales, were designed to increase

speed and quality of data entry. Online and printed help is available. Results from analytical instruments and data from legacy systems can be imported to, or exported from, RGMIMS via spreadsheet. Query tools provide ready access to relevant information; users can formulate ad hoc queries without recourse to traditional query languages.

There are scripts to automatically download ontologies and taxonomy from public FTP sites and to upload it into RGMIMS. The user could schedule the running of these scripts at predefined times.

Traceability is maintained to ensure regulatory compliance, e.g. transformations and mutant plant lines are traceable to their source seed lots. Resources are tracked rapidly through unique identifiers, automatically generated barcode labels, and their holding locations such as growth facilities and seed storage locations.

RGMIMS is configurable to specific laboratory requirements. It can be customised to user preferences including plant naming conventions, measurement units, growth facilities, and contact details of biological resource donors or recipients. As a platform, RGMIMS could be readily adapted to support a wide variety of plant functional genomics studies.

References

Donofrio N, Rajagopalon R, Brown D et al. (2005) 'PAC-LIMS': a component LIM system for high-throughput functional genomic analysis. BMC Bioinformatics 6:94

Jaiswal P, Ni J, Yap I, Ware D, Spooner W, Youens-Clark K, Ren L, Liang C, Zhao W, Ratnapu K, Faga B, Canaran P, Fogleman M, Hebbard C, Avraham S, Schmidt S, Casstevens TM, Buckler ES, Stein L, McCouch S (2006) Gramene: a bird's eye view of cereal genomes. Nucleic Acids Res 34:D717–D723

McLaren CG, Bruskiewich RM, Portugal AM, Cosico AB (2005) The International Rice Information System. A platform for meta-analysis of rice crop data. Plant Physiol 139:637–642

The Plant Ontology™ Consortium (2002) The Plant Ontology™ Consortium and plant ontologies. Comp Funct Genomics 3(2):137–142

Scott R, Clark C, Mathews A, Yallop S, Davey MR, Lazzeri P, Mutasa-Göttgens E (2003) Transgenic plant monitor (TPM): a relational database for the management and analysis of information in plant tissue culture and transformation laboratories. Mol Breed 11:121–125

Upadhyaya NM, Zhu Q-H, Zhou X-R, Eamens AL, Hoque MS, Ramm K, Shivakkumar R, Smith KF, Pan S-T, Li S, Peng K Kim SJ, Dennis ES (2006) *Dissociation* (*Ds*) constructs, mapped *Ds* launch pads and a transiently expressed transposase system suitable for localized insertional mutagenesis in rice. Theor Appl Genet 112:1326–1341