Strategic Collation of Evaluation Data for Priority Trait Screening

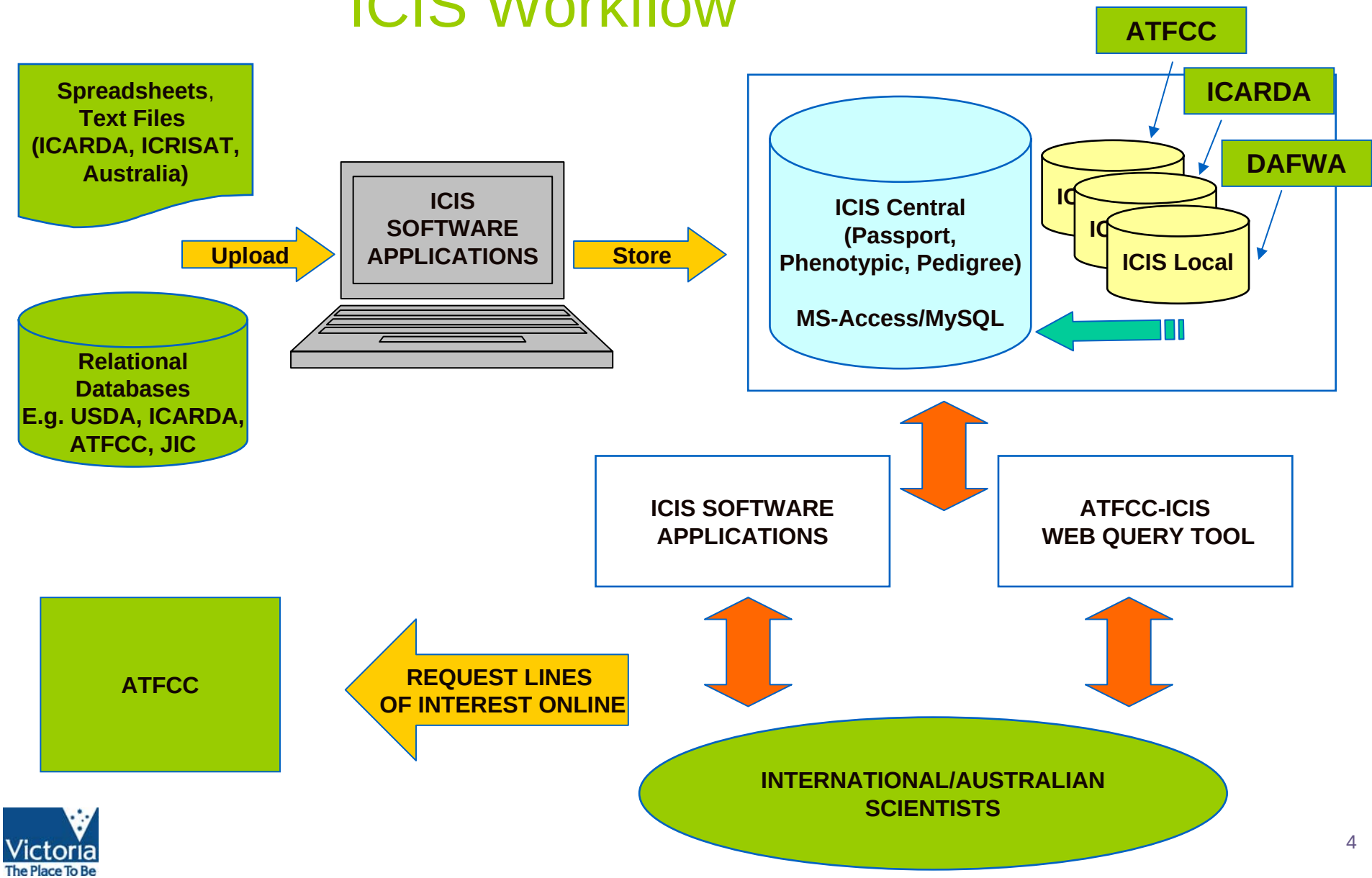
Bob Redden, Ranjan Balachandra,
Kevin Murray, Selwyn Ellis.

Overview

- 4 Year DPI / GRDC Project
  - Jan 2003 – Dec 2006.
- Objective - to collate Australian evaluation data for the lentil and chickpea, pea and faba bean germplasm for priority trait breeding.
- Key outputs
  - Relational databases for location, passport and evaluation data for lentil, chickpea, pea and faba bean.
  - Web-based search engine to interrogate the data.
- These will enable a virtual world gene bank online.

Global Cooperating Organisations

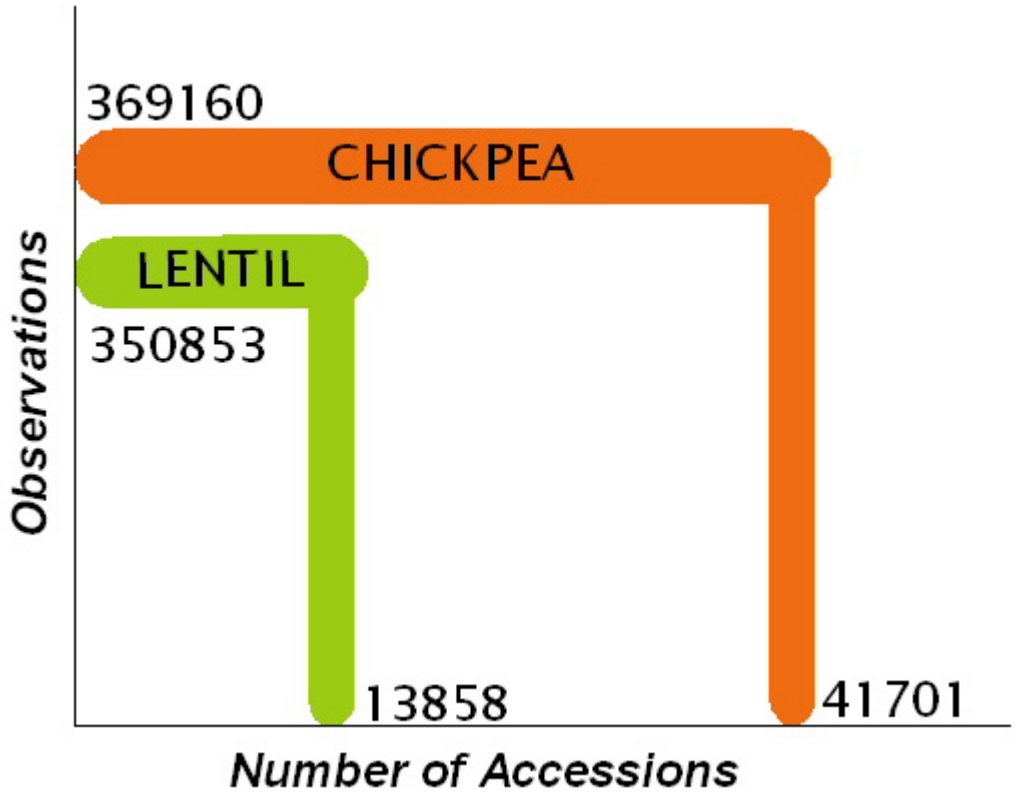
ICIS Workflow



Data Sharing and Collation

- Organisation of data into relational databases using ICIS database structures with storage of data from multiple sources.
- Negotiations with ICARDA, ICRISAT, USDA and European gene banks for evaluation data.
- Visits to ICARDA/ICRISAT and European genebanks to assist access to data.
- Making databases available publicly through a web search engine for users to perform multiple trait queries.
- Interaction with and feedback from end users.

Statistics


369160
CHICKPEA
LENTIL
350853
Observations
13858
41701
Number of Accessions

Plant Breeder Feedback

2004 tele-conference with breeders, with on-line database.

Participants;breeders of chickpea, lentil, faba bean and pea in Vic, NSW, SA, and WA.

2006, seminar/workshops at 6 locations,
Perth, Adelaide, Horsham, Melbourne- DPI biotechnology Bundoora, and Univ. Melbourne, and Tamworth

Key suggestions;
- On-line requesting of germplasm.
- Site filters to exclude data, e.g. with low disease levels.
- Display graphical distribution of trait values.
- Linkage with FIGS database for abiotic stress tolerance.
- Linkage of accessions with seed/plant images.
- Combine passport and trait queries.

Database Organisation

- Data for lentils and chickpea were loaded into ICIS databases.
  - Unique identification of germplasm accessions and linkages of different names for the same accession.
  - Standardised a procedure for the upload of evaluation data.
- Best solution = upload entire crop data files from international sources instead of just those pertaining to the ATFCC's collections.
  - Project objectives were expanded to have single germplasm evaluation databases for worldwide evaluation data instead of just those matching the ATFCC's accessions.
  - Huge potential for the ATFCC to become a considerable part of a world gene bank network.

ICIS Chickpea database summary

- Contains 382,282 data items on 107 traits.
- Traits include:

LOCATION	PLANTSIZE	Stem pigmentation	Seed weight
SAMPLE STATUS	PLANTHABIT	Seed splitting	Seed hydration
INVENTORY WEIGHT	Hypocotyl color	Sowing time	Cotyledon colour
INVENTORY	Natural life span	Pods per peduncle	Seed coat colour
Herbicide tolerance	Pod borer	Processing type	Botrytis reaction
Iron deficiency	Colletotrichum Reaction	Seed yield	Experiment Design
Frost damage	Seed surface	Rust reaction	Ascochyta reaction
Cold reaction	Seed shape	Protein content	Dehulling
Cyst nematode reaction	Dots on seed coat	Pod pigmentation	Herbicide damage
Callosobruchus chinensis	Seed color	Pod shedding	Hard seed
Leaf Miner Reaction	Tertiary branches number	Shattering	Stem length
Seeds per area	Apical secondary branches number	Growth habit	Pea Enation Mosaic Virus
Pods per Flowering Node	Basal secondary branches number	Seeds per pod	PC
Peduncles per plant	Apical primary branches number	Seeds per plant	Selected
TIME TO POD	Basal primary branches number	Pods per plant	Catalogue
Plant hairiness	Plant pigmentation	Flowers per peduncle	Location
FLOWERING DURATION	Plant width	Notes on data	Designation
Primary Branches	Experiment	Moisture content	Year
Leaflet Number	Seed type	Maturity	Leaf size
Pod Length	Sclerotinia reaction	Lodging	Harvest index
Picture/Image	Leaf pubescence	Plant height	Canopy width
CORE SUBSET	Environmental stress	Fusarium reaction	Biological yield
Drought tolerance	Virus reaction	Flower colour	GID
Leaf type	Seed coat pattern	Pod Height	LATITUDE
Leaf color	Tendril length	Emergence / establishment	LONGITUDE
Pod concentration	Orobanche reaction	Vigour	COUNTRY [ISO CODE]
Year and yield	Staw yield	TIME TO FLOWERING	

Shared Database Management

- Procedure for shared management of ILIS (Lentil) has been submitted to and approved by ICARDA.

- Similar procedure for shared management submitted to ICARDA and ICRISAT (in progress) for a central IChIS (Chickpea) database.

- Future evaluation data from lentil, chickpea and faba bean breeding programs will be uploaded from Australian sources and passed on to ICARDA and ICRISAT annually.

- Australia (DPI Vic) to host a central copy of a world pea database.

Web Search Engine

- A web-based search engine prototype for multiple trait searching on crop germplasm and retrieval of matching accessions, was evaluated by breeders (January 2005) and feedback was collated.

- Iterative interrogation using multiple trait search parameters.

- Use of the open source technology (e.g. Java programming language, MySQL) for application development.
  - In line with ICIS' technological direction.

Web Search Engine (continued)

- Key functionalities:
  - Filtering by trait values (e.g. good disease resistance).
  - Filtering by site means (e.g. lines from high-yielding sites).
  - Discretization of quantitative data (1-9 scale, normalisation).
  - Graphical display of trait value distributions.
  - Achieve a shortlist of accessions matching key trait values.
  - Request lines of interest online from the ATFCC and overseas.

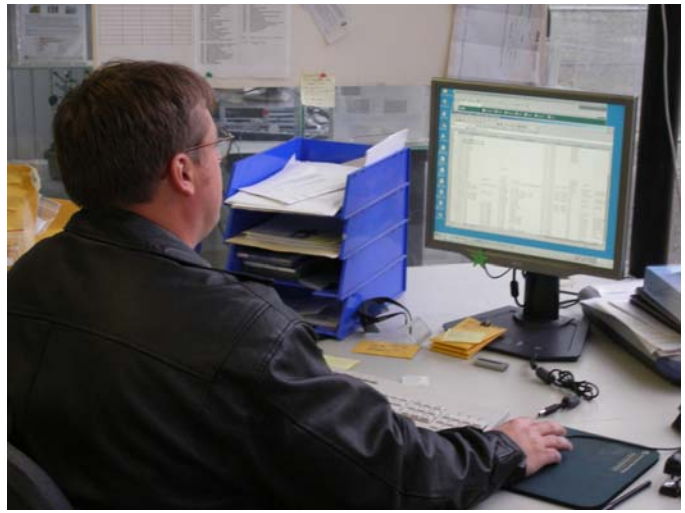
Immediate Priorities (by Dec 2006)

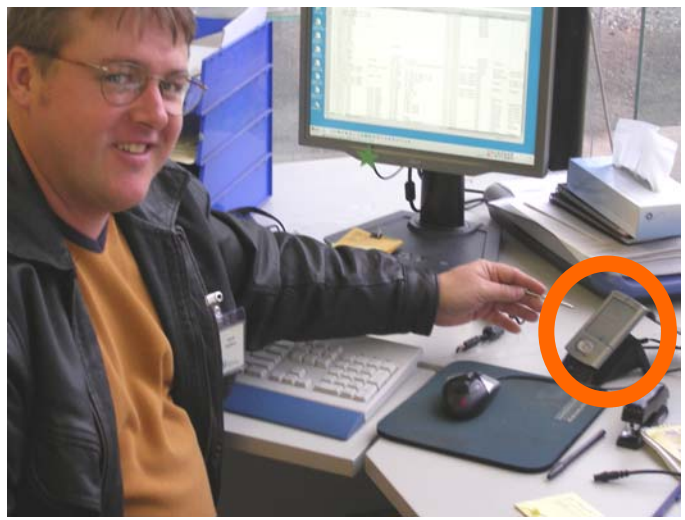
- Incorporation of pea and faba bean data into ICIS databases.
- Implementation of web application and databases in DPI Vic server for public access.
- Functionality to view raw data as well as specifying pedigree search filters.
- Additional Filters
  - Trial site filters (e.g. lat-long coordinates, country, site codes, etc).
  - Pedigree search filters (e.g. progenitors/descendants of landraces).

Future Directions

- Enhanced collaboration.
  - Stronger linkages with ICARDA and IRRI for development of additional functionalities for search engine development including research into mining global distributed databases.
  - Enhancements to user interfaces

- Propose to the Australian gene banks within the framework of the National Genetic Resource Centre, on developing a national integrated gene bank information system.

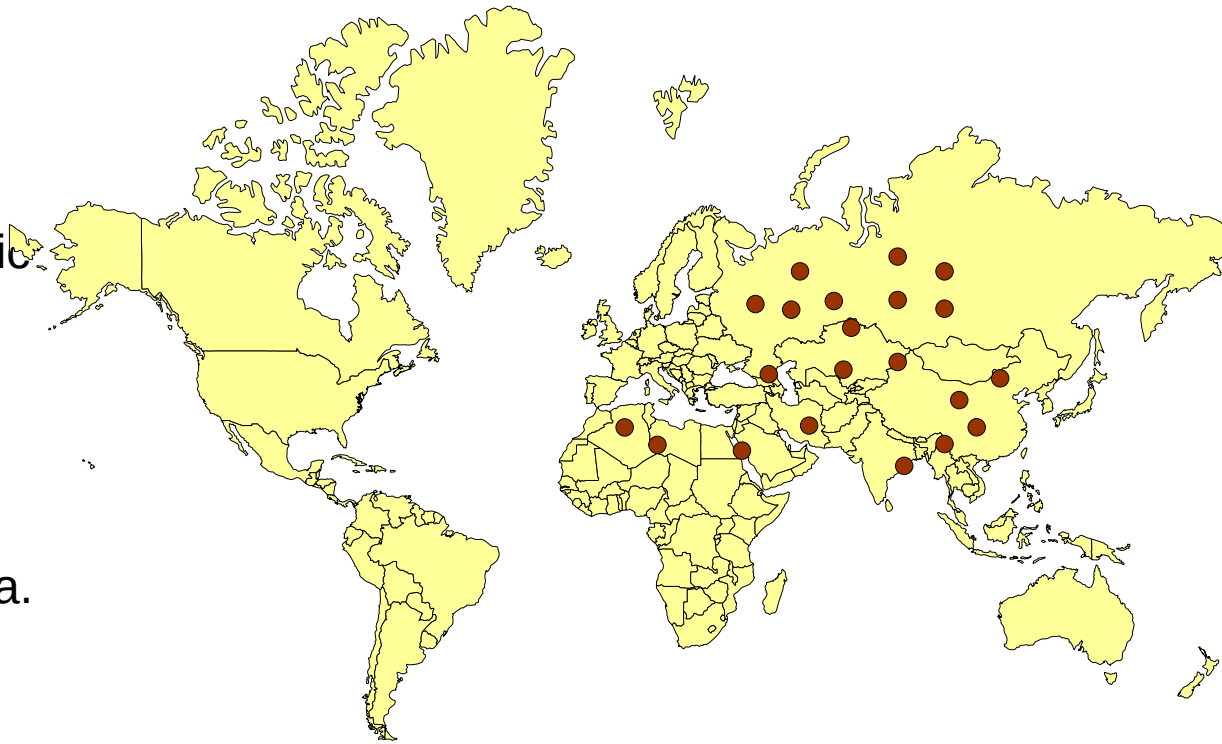




Future Directions (continued)

- Integration of GPS coordinates for origins of accessions with geographic and environmental abiotic stress GIS maps.

- Incorporation of molecular fingerprint and marker data.

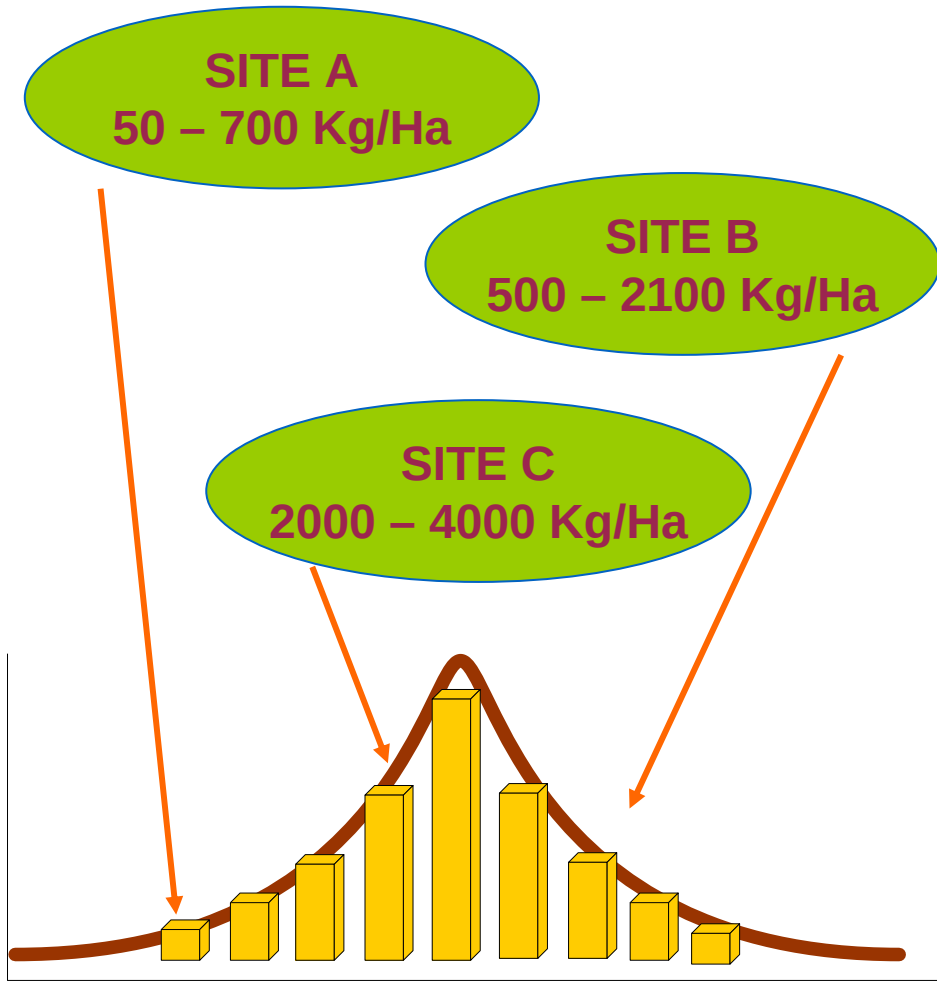


Demo

Database demonstration

Discretization of Quantitative Trait Data

- In order to compare the relative performance of accessions over multiple trials, a standardisation procedure is necessary.
- Absolute values are converted to fit into discrete scores on a 1-9 scale based on their performance relative to other accessions in individual trials.
- Replication of this procedure over multiple experiments and traits.



Benefits of Web-enabled ICIS databases

- Benefits gained from expansion of initial project objectives.
  - Fits in the ATFCC's vision to be a key part of an international group.
  - Expands querying possibilities to globally available data.
- Breeders have the ability to make world's best choice.
- ATFCC Benefits.
  - Enhanced work efficiency in the ATFCC.
    - Handling requests from breeders.
    - Ensuring quality of germplasm collection.
  - Client-driven web-facilitated selection of gene bank accessions.

Benefits (continued)

- Breeders.
  - Identify suitable candidates for immediate use.
  - More efficient utilisation of genetic resources.
  - 24/7 access to the database regardless of time zone.
- Data from various sources available through a single point of access.
- Use of relational databases
  - industry standard using best practices.