RICE GENETIC DIVERSITY PLATFORM TO ACCELERATE GENE DISCOVERY AND PRECISION BREEDING

Hei Leung
International Rice Research Institute
DAPO 7777, Metro Manila
Philippines

ABSTRACT

Genetic diversity is the foundation of the genetic improvement of crop plants. Knowledge of multiple facets of rice genetic diversity from the molecular to the phenotypic is essential for effective conservation and use to meet current and future needs. Under the GRiSP research agenda, we seek to harness collective research capacity to achieve innovations that are not possible by individual institutions alone. We envision major innovations of significant scale and scope. These include:

- integrating management of the world’s largest collection of rice genetic resources;
- establishing a global public genetic diversity research and gene discovery platform;
- modernizing trait evaluation using high-throughput precise phenotyping;
- designing plant genotypes for climate-change scenarios;
- producing a more efficient rice plant (C₄) for the future.

In this talk, I will highlight ongoing work and initiatives under GRiSP that will benefit from close collaboration with research institutions in Japan.

Closing the gene-phenotype gap with systematic genotyping and phenotyping. As a foundation for breeding and gene discovery research, we are building a genetic diversity research platform to enable efficient use of rice diversity. Underpinning this platform are several essential components:

- Detailed genotypic and phenotypic evaluation of a common set of genetic resources. This involves the use of a collection 2,000 diverse rice lines in genome-wide association studies to discover gene-phenotype relationships (Tung et al. 2010). The 2,000 lines are being genotyped by a 1 M SNP chip. By evaluating this germplasm collection for multiple traits in different environments, we will determine the gene-phenotype relationships, and genotype x environment interactions. In parallel, multi-trait mapping populations are being developed to enable high resolution mapping to validate gene-phenotype relationships predicted by association mapping. These include the Nested Association Mapping (NAM) populations developed by CIAT and AfricaRice, and the Multi-parent Advanced Generation InterCross (MAGIC) populations developed by IRRI.
- Decoding a significant portion of the International Rice Gene Bank (10% of accessions) to identify rare alleles for use in breeding.
- Integrating the sequence and phenotype information to enable the design of new virtual genomes for high-value varieties.

Precision breeding to develop climate-ready rice. Climate change is expected to significantly affect the abiotic and biotic environments where rice is grown. Changes in temperature and availability of water in quantity (too much or too little) and in quality (salinity) will directly impact rice productivity. Such climatic conditions are already prevalent in regions of the world where food security is a major concern. For developing climate-ready rice, we need to assemble a battery of stress tolerance genes and put them into
ready-to-use breeding stocks. Major genes and large-effect QTL have been identified to confer tolerance to submergence, salinity, and drought conditions. Examples include genes for submergence tolerance (Sub1) and salinity-tolerance (Saltol), and major QTL for sustaining yield under drought stress. Research on heat tolerance or avoidance is progressing, leading to identification of useful genetic variation. The study of early-morning flowering pioneered in Japan illustrates the use of new traits for adaptation to new environments. Due to the dynamic nature of pathogen and insect populations, it is more difficult to predict the impact of climate change on the biotic environment. New approaches to improve resilience to the variable pathogen and insect populations are therefore needed. Recent advances in our understanding of non-host resistance may lead to new strategies to achieve stable resistance to multiple diseases (Alyffe et al 2011).

**Breaking the yield ceiling.** A key achievement made in modern plant breeding is the dramatic change in harvest index (HI), where the grain yield constitutes over 50% of the total biomass (HI >0.5). However, further improvement of HI beyond 0.55 has proved difficult. To further increase grain yield, it is necessary to explore new approaches to increase total biomass and to enhance photosynthetic and respiratory efficiency. A GRiSP Workshop on Yield Potential was held in August 22-26, 2011 at CIAT, Colombia to discuss strategies on raising the yield potential of rice. From this workshop, five research areas emerge as important: 1) combine large-effect genes for yield component traits, 2) accumulate QTLs controlling yield potential traits using Marker-Aided Recurrent Selection and Genomic Selection, 3) exploit yield enhancing loci from wild species, 4) enhance the level of heterosis, and 5) breeding for physiological traits. The identification of large-effect QTL affecting yield components in Japan is particularly relevant to this effort.

Each of these research activities requires collaboration to bring together complementary expertise to achieve the objectives. Leveraging rice research in Japan will be crucial to success. The International Rice Genome Sequencing Project (IRGSP) led by Japan produced the first reference genome of rice that has impacted every aspect of rice genetics research. The large collections of rice mutants (Tos17 lines, FOX lines) and specialized genetic stocks (chromosomal substitution lines, near-isogenic lines) produced in Japan will continue to be the drivers of gene discovery. Many genes of potential agronomic importance have already been identified by different laboratories in Japan. These genes can be evaluated in multiple genetic backgrounds and environments through different GRiSP networks. GRiSP provides a mechanism to promote such collaboration and to accelerate the translation of scientific research into practical impact.

**KEYWORDS**
climate change, genome sequencing, productivity, stress tolerance

**REFERENCES**


Rice genetic diversity platform: accelerate gene discovery and precision breeding

Hei Leung
International Rice Research Institute

GRiSP Theme 1: Genetic Diversity and Gene Discovery

- integrating management of the world’s largest collection of rice genetic resources;
- establishing a global public genetic diversity research and gene discovery platform;
- modernizing trait evaluation using high-throughput precise phenotyping;
- designing plant genotypes for climate-change scenarios;
- producing a more efficient rice plant (C₄) for the future.

Public Genetic Diversity Research Platform

<table>
<thead>
<tr>
<th>Drought tolerance</th>
<th>Grain quality</th>
<th>Problem soils</th>
<th>C₄ Rice</th>
</tr>
</thead>
</table>

Current problems: Drought tolerance, Grain quality, Varied soils, C₄ Rice

Future challenges: Phenotype-genotype association

Public Genetic Diversity Research Platform

Use Gene Function

Conserved Germplasm
Breeding Lines
Specialized Genetic Stocks

Use Gene Function

Conserved Germplasm
Breeding Lines
Specialized Genetic Stocks

Rice SNP Consortium for gene discovery

- Developed high-density genotyping Affy arrays with 1 M SNPs
- Includes newly discovered SNPs from >150 genomes and from other projects
- Genotype 2000 rice lines spanning range of diversity
  - http://www.ricesnp.org
- Partners include Cornell, USDA, AfricaRice, CIRAD, Bayer CropSciences, Syngenta, CIAT

Rice SNP Consortium for gene discovery

- Precise phenotyping of traits in target environments for an integrated Rice Diversity Platform

Discussion outline

- Determine gene-phenotype relationship.
- Precision breeding to develop climate-ready rice:
  - Abiotic stresses
  - Biotic stresses
- Breaking the yield ceiling
- Blue-sky research: making rice productive but with lesser environmental impact

Rice SNP Consortium for gene discovery

- 3000 diverse rice lines clustered by molecular markers

Genetic Resources: Genotype/Phenotype information

- Minimal knowledge of most materials
  - Solution: Data on whole collection to predict performance - sequence collection
  - Knowledge of genotype / phenotype / agronomic value
  - Number of lines

Genetic Resources: Genotype/Phenotype information

- Detail required to evaluate agronomic performance
1M SNP Chip design from pool of >27M SNPs
(from 150+ resequenced genomes, OryzaSNP, and OMAP projects)

- 1,028,547 SNP features using 4 or 6 probes each
- 23,855 SNP targets for QC
- 170,240 CHI-sensitive sites
- 57,753 Invariant 25-mer probes
- 1,280,395 Total targets assayed by 6,967,550 probes

Rice diversity panel: structure of ~2,000 rice samples

- 500 Cornell University (Susan McCouch, NSF project)
- 1440 samples from the IRRI nominations split into
  450 Indica
  350 trop. Japonica
  250 Aas
  150 temp. aopica
  100 aromatic
  20 O. glaberrima
  50 O. barthii
  70 O. rufipogon and nivara

Association analysis with 44K SNPs

Association analysis for plant height

Progress in IR64 assembly

De novo assembly, per chromosome for contigs from mapped + unmapped reads that align to Nipponbare

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Estimated kb</th>
<th>Total aligned</th>
<th>Total # of contigs</th>
<th>% Nb cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr01</td>
<td>49</td>
<td>4555</td>
<td>49391</td>
<td>84.77</td>
</tr>
<tr>
<td>chr02</td>
<td>51</td>
<td>4176</td>
<td>42856</td>
<td>87.62</td>
</tr>
<tr>
<td>chr03</td>
<td>51</td>
<td>4482</td>
<td>49328</td>
<td>90.58</td>
</tr>
<tr>
<td>chr04</td>
<td>51</td>
<td>3502</td>
<td>41293</td>
<td>76.27</td>
</tr>
<tr>
<td>chr05</td>
<td>51</td>
<td>3345</td>
<td>54516</td>
<td>85.10</td>
</tr>
<tr>
<td>chr06</td>
<td>51</td>
<td>3507</td>
<td>44765</td>
<td>80.81</td>
</tr>
<tr>
<td>chr07</td>
<td>53</td>
<td>3386</td>
<td>41841</td>
<td>79.71</td>
</tr>
<tr>
<td>chr08</td>
<td>53</td>
<td>3158</td>
<td>40641</td>
<td>82.34</td>
</tr>
<tr>
<td>chr09</td>
<td>53</td>
<td>2941</td>
<td>44188</td>
<td>81.51</td>
</tr>
<tr>
<td>chr10</td>
<td>53</td>
<td>2834</td>
<td>36501</td>
<td>79.44</td>
</tr>
<tr>
<td>chr11</td>
<td>53</td>
<td>3038</td>
<td>44914</td>
<td>77.42</td>
</tr>
<tr>
<td>chr12</td>
<td>53</td>
<td>2933</td>
<td>43281</td>
<td>73.57</td>
</tr>
<tr>
<td>Total</td>
<td>3488</td>
<td>54516</td>
<td>305263273</td>
<td>81.99</td>
</tr>
</tbody>
</table>

Genebank sequencing - 10k project

Apply low-cost sequencing by next generation technology

- Working with BGI-Shenzhen to sequence 10,000 genomes
- First 3,000 DNAs shipped to BGI for sequencing
- Tests using indexed libraries and pooling on 48 samples gave an average of 7x coverage with a range of 3.6 to 15.5x
Genetic resources underpinning gene-phenotype association

Breeding populations (~750)
- 25 elite breeding lines
- 20 OryzaSNP panel
- 16 MAGIC founders

Association genetics
- Diversity panel
2000 lines, divided into 5-6 sub-populations for phenotyping
- Yield Potential
- Genetic stocks for validation
- Permanent populations to validate gene-phenotype predictions

Phenotyping for multiple traits at specialized phenotyping platforms

Yield Potential
- Common target of breeding programs
- Yield-related traits
- Done at field phenotyping hubs

Breeding populations (~70)
- 25 elite breeding lines
- 20 OryzaSNP panel
- 16 MAGIC founders

Phenotyping for multiple traits at specialized phenotyping platforms

Global MAGIC: indica x japonica S1 lines

Testing genotyping by sequencing (GBS) for high-resolution genome scans
- GBS can potentially provide low-cost high-resolution SNP scans by multiplex sequencing
- Buckler lab, Cornell University

96-plex GBS method:
Elshire et al. 2011

Library Prep: ApeKI digestion, ligation, PCR, cleanup, QC

Sequencing:
96 samples/lane on HiSeq2000

Analysis pipeline:
align tags, call SNPs

Elshire et al. 2011 PLoS One

Testing GBS in rice at Cornell University

GBS in S1 MAGIC population
- 75k SNPs
- 12k SNPs
- 7k SNPs
- 3k SNPs
- 872 SNPs

GBS in MAGIC population
- 96 samples S1 MAGIC
- 96 samples S1 MAGIC plus

Pilot study at Cornell (with Susan McCouch and Chih-Wei Tung)

MAGIC lines from IRRI:
- 184 samples of S4 MAGIC indica
- 96 samples S1 MAGIC
- 96 samples S1 MAGIC plus

Mike Thomson, Chih Wei Tung et al.
GWAS analysis using General Linear Model (GLM) in TASSEL identified several new QTLs.

C. Raghavan et al. IRRI

Rice Nested Association Mapping (NAM) populations
Founders: O. sativa cultivars popular in Africa & LA All crossed to IR64 (near F6-F7 now)
High resolution mapping

Mathias Lorieux et al. CIAT/Africa Rice

Discussion outline
• Determine gene-phenotype relationship.
• Precision breeding to develop climate-ready rice.
  o Abiotic stresses
  o Biotic stresses
• Breaking the yield ceiling.
• Blue-sky research: making rice productive but with lesser environmental impact

Convergence of problems already here in some rice production environments
Consequences of Climate Change:
Rice systems will experience more...
  • Drought
  • Submergence
  • Salinity
  • Heat waves

Challenges of Climate Change = Challenges in unfavorable environments

Progress as of wet season 2011
• Approx. 1.12 million farmers in South Asia
• Major activities:
  * 100 tons of Swarna-Sub1 was distributed by Bihar government under seed village programme.
  * Large scale promotion by State Governments of UP, Bihar and Orissa
  * DAE in Bangladesh distributed 19 tons seed for multiplication by the farmers in 44 districts. Total coverage: 54 districts out of 64 in country

U. Singh et al. IRRI
BR28-SalTol = Available and being tested in the field
IR64-SalTol = Breeder seeds
BR11-SalTol = Breeder seeds
BR29-SalTol = Almost completed

Two in one - Saltol and Sub1

- Combining Saltol and SUB1 in one genetic background seems feasible with no apparent negative impacts on agronomic traits, and this will help develop more stable varieties adapted to coastal zones
- Advanced breeding lines (~20) with both salinity and submergence tolerance sent to all the NARES partners for evaluation and selection under their target environments.

Mega varieties with SalTol

BR28-SalTol
BR11-SalTol

IR64-SalTol = Breeder seeds
BR29-SalTol = Almost completed

Discovering large-effect QTLs for grain yield under drought

- Major QTLs in the background of high-yielding but drought-susceptible mega varieties IR64, MTU1010, and Swarna identified
- QTLs introgression in IR 64 and Swarna near completion
- Efforts initiated for introgression of RM 520, RM 324 drought QTLs in Swarna-Sub1

Potential heat tolerance donors

Dular IR28 PSBRc94
Ganga Rata IR50 RJT174
Ganjay IR8866-30-1-4-2 Sahel108
Gerdeh Fars Jijai Niki Tak Ratia
Giza178 (Egypt) Kala Chawanran Todorokiwase (Japan)
IR2006-p-12-12-2-2 Khara Ganja Toor
IR2307-247-2-2-3 N22 (India)

Potential heat tolerant Korean varieties
Chengcheong KeunSeom Milyang23 TR22183

Climate change on biotic stresses

1. "Bio-stations"
   - Multiple sites to sample diverse climatic environments
   - Specialized genotypes as biological probes to monitor host-pest/pathogen interactions

2. Mobilize rapidly resistance resources
   - Breeding-ready gene pool
   - Stay one-step ahead

QTL mapping

- Two QTL are responsible for spikelet fertility under high temperature.

Two QTL are responsible for spikelet fertility under high temperature.

<table>
<thead>
<tr>
<th>Chromosome 1</th>
<th>Location: 150.5 cM</th>
<th>LOD: 4.63</th>
<th>R²: 12.6%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Additive effect</td>
<td>5.66</td>
<td>Dominance</td>
<td>-16.61</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chromosome 4</th>
<th>Location: 67.5 cM</th>
<th>LOD: 6.66</th>
<th>R²: 17.6%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add. effect</td>
<td>-0.81</td>
<td>Dom. Effect</td>
<td>27.66</td>
</tr>
</tbody>
</table>

Source: A. Kumar, IRRI

Installation of Vantage Pro 2 Weather Station

At PhilRice Izusan 9 – 27 Jan 2011
Resistance of a UK-derived plant to RTV

A GRiSP Workshop on Yield Potential
August 22-26, 2011, CIAT, Colombia
- Combine large-effect genes for yield component traits,
- Accumulate QTLs controlling yield potential traits using Marker-Aided Recurrent Selection and Genomic Selection,
- Exploit yield enhancing loci from wild species,
- Enhance the level of heterosis, and
- Breeding for physiological traits.

C4 Rice Project

Aligning research collaboration under GRiSP

Acknowledgment

Financial Support