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 (C_4) for the future.

In this talk, I will highlight on-going 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-traits 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 to confer tolerance to submergence, salinity, and drought conditions. Examples include genes for submergence tolerance (*Sub1*) and salinity-tolerance (*Saltol1*), 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 diseasess (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 identifcaiton 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 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.

.

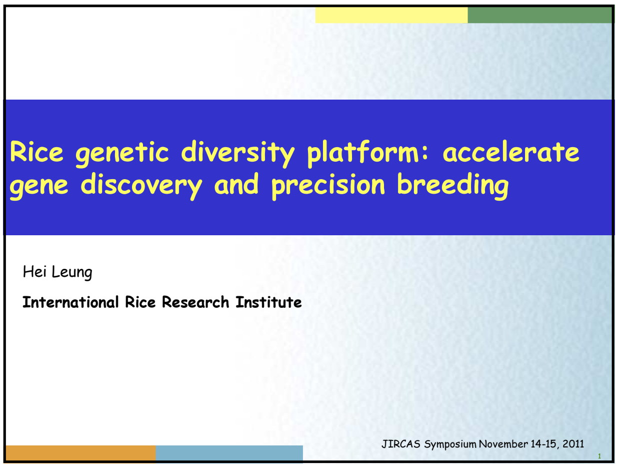
KEYWORDS

climate change, genome sequencing, productivity, stress tolerance

REFERENCES

Tung, C.W., K. Zhao, K. Wright, L. Ali, J. Jung, J. Kimball, W. Tyagi, M. Thomson, K.L. McNally, H. Leung, H. Kim, S.N. Ahn, A. Reynolds, B. Scheffler, G. Eizenga, A. McClung, C. Bustamante and S. McCouch, 2010. Development of a research platform for dissecting phenotype-genotype associations in rice (*Oryza* spp.). Rice, 3, 205-217.

Ayliffe, M., R. Devilla, R. Mago, R. White, M. Talbot, A. Pryor, H. Leung, 2011. Non-host resistance of rice to rust pathogens. Molecular Plant Microbe Interactions, 24, 1143-1155.

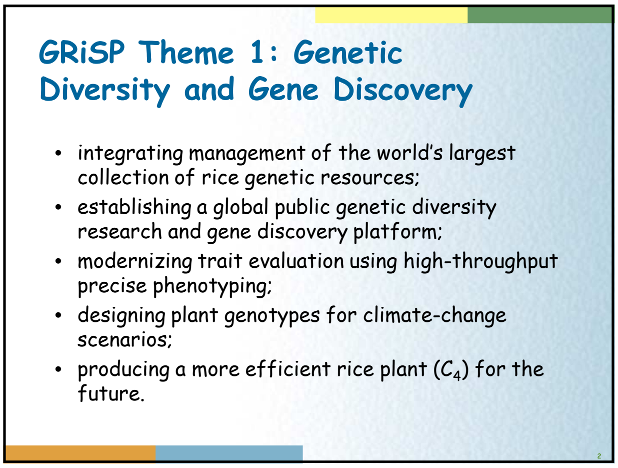


Rice genetic diversity platform: accelerate gene discovery and precision breeding

Hei Leung

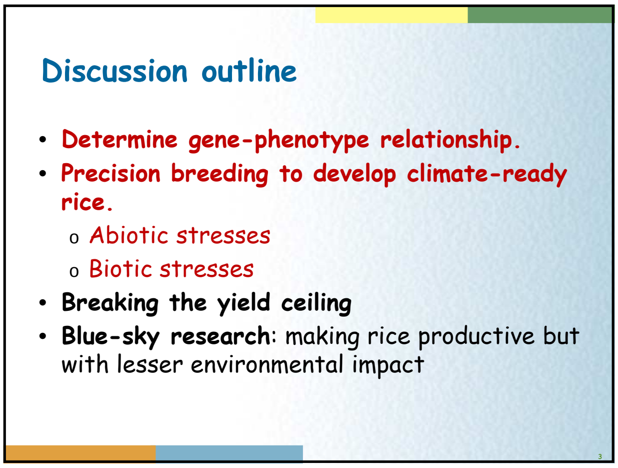
International Rice Research Institute

JIRCAS Symposium November 14-15, 2011



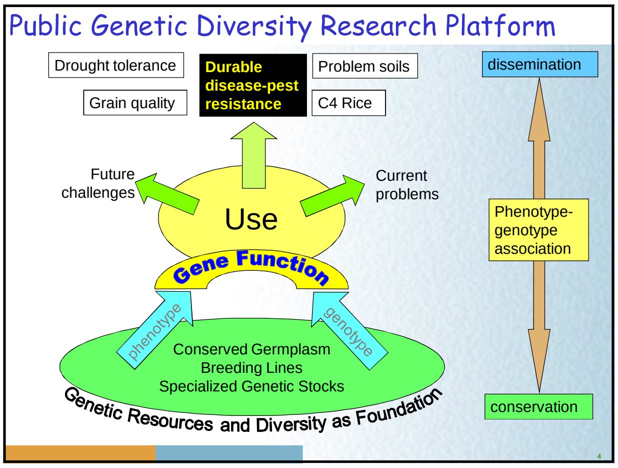
GRiSP Theme 1: Genetic Diversity and Gene Discovery

- 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 (C_4) for the future.



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



Public Genetic Diversity Research Platform

Drought tolerance

Grain quality

Durable disease-pest resistance

Problem soils

C4 Rice

dissemination

Future challenges

Use

Current problems

Phenotype-genotype association

Gene Function

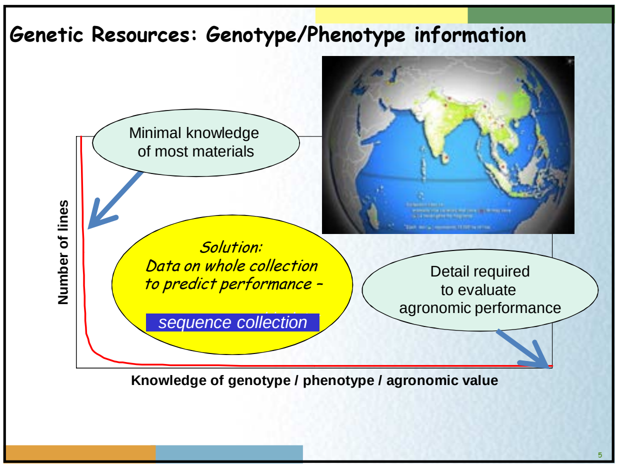
phenotype

genotype

Conserved Germplasm
Breeding Lines
Specialized Genetic Stocks

Genetic Resources and Diversity as Foundation

conservation

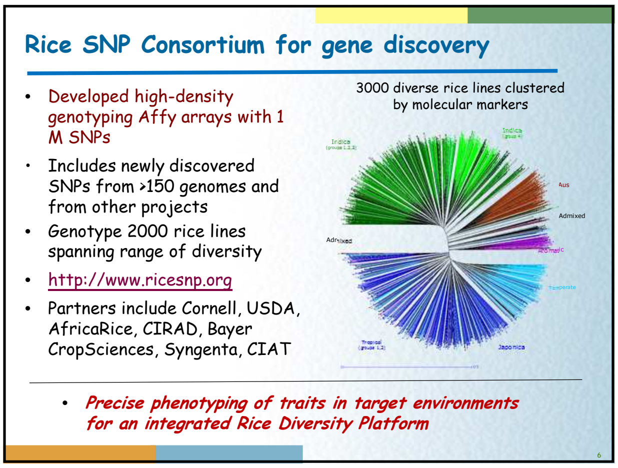


Genetic Resources: Genotype/Phenotype information

Minimal knowledge of most materials

Solution:
Data on whole collection to predict performance –
sequence collection

Detail required to evaluate agronomic performance



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

3000 diverse rice lines clustered by molecular markers

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

1M SNP Chip design from pool of >27M SNPs

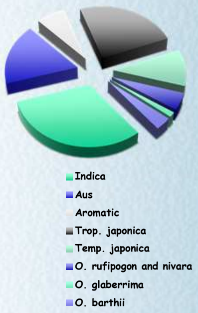
(from 150+ resequenced genomes, OryzaSNP, and OMAP projects)

1,028,547	SNP features using 4 or 6 probes each
23,855	SNP targets for QC
170,240	CH$_3$-sensitive sites
57,753	Invariant 25-mer probes
1,280,395	*Total targets assayed by 6,967,550 probes*

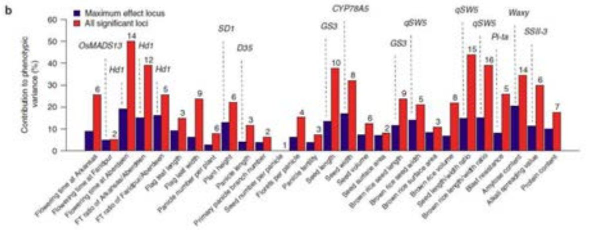
S McCouch, M Wright

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

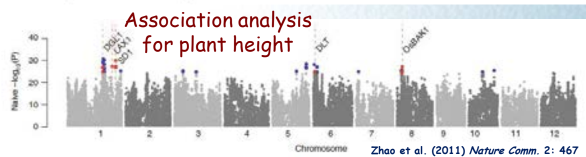
- 500 Cornell University (Susan McCouch, NSF project)
- 1440 samples from the IRRI nominations split into
  - 450 Indica
  - 350 trop. Japonica
  - 250 Aus
  - 150 temp.aponica
  - 100 aromatic
  - 20 *O. glaberrima*
  - 50 *O. barthii*
  - 70 *O. rufipogon* and *nivara*



Association analysis with 44K SNPs



Association analysis for plant height



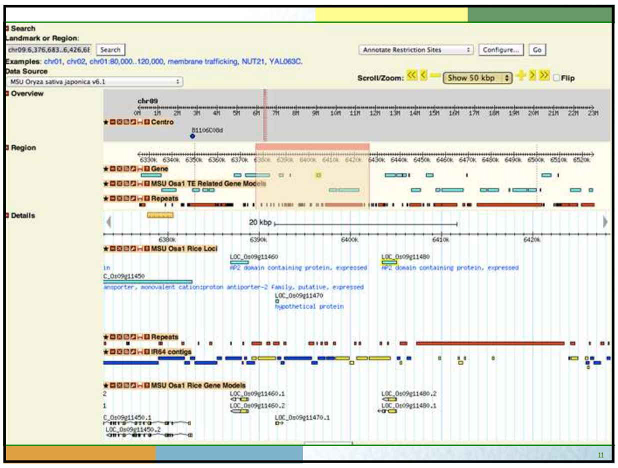
Zhao et al. (2011) *Nature Comm.* 2: 467

Progress in IR64 assembly

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

Chrom	optimum k	n50	max contig size	total aligned size of contig to Nb	Total # of contigs	# of contigs aligned to Nb	# of contig not aligned to Nb	Expected Nb chrom size	% Nb cover
chr01	49	**4555**	49391	36679538	41160	16312	24848	43268879	84.77
chr02	51	4176	42856	31483319	35355	14221	21134	35930381	87.62
chr03	51	4482	49328	32976864	35605	13927	21678	36406689	**90.58**
chr04	51	3502	41293	26905278	39670	14883	24787	35278225	76.27
chr05	51	3345	**54516**	25441365	41430	16254	25176	29894789	85.10
chr06	51	3507	44765	25249764	37048	13625	23423	31246789	80.81
chr07	53	3386	41841	23671352	36533	12385	24148	29696629	79.71
chr08	53	3158	40641	23416680	38494	13758	24736	28439308	82.34
chr09	53	2941	44188	18756480	33952	10704	23248	23011239	81.51
chr10	53	**2834**	36501	18378387	36966	10439	26527	23134759	79.44
chr11	53	3038	44914	22075634	36024	13049	22975	28512666	77.42
chr12	53	2933	43281	20228612	35892	11997	23895	27497214	**73.57**
Total		**3488**	**54516**	**305263273**	**448129**	**161554**	**286575**	**372317567**	**81.99**

R. Mauleon et al. IRRI
Rice SNP Consortium



Search

Landmark or Region:

chr09:6,376,683..6,426,68 Search

Examples: chr01, chr02, chr01:80,000..120,000, membrane trafficking, NUT21, YAL063C.

Data Source

MSU Oryza sativa japonica v6.1

Annotate Restriction Sites | Configure... | Go

Scroll/Zoom: Show 50 kbp ☐ Flip

Overview

Region

Details

LOC_Os09g11460
AP2 domain containing protein, expressed

LOC_Os09g11480
AP2 domain containing protein, expressed

C_Os09g11450
ansporter, monovalent cation:proton antiporter-2 family, putative, expressed

LOC_Os09g11470
hypothetical protein

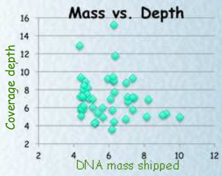
Genebank sequencing – 10k project

IRG Traditional Germplasm
100,000 cultivated accessions

Iterative sampling
5X+ depth

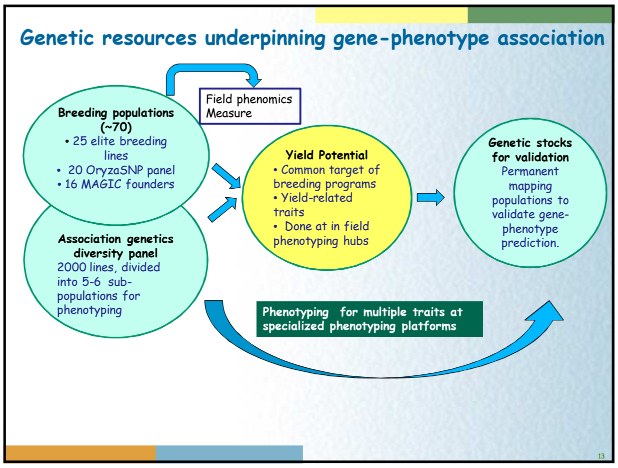
Apply low-cost sequencing by next generation technology

- 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 <u>7x coverage with a range of 3.6 to 15.5x</u>



Chinese Acad. Agric Sci (CAAS)

华大基因 BGI



Genetic resources underpinning gene-phenotype association

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

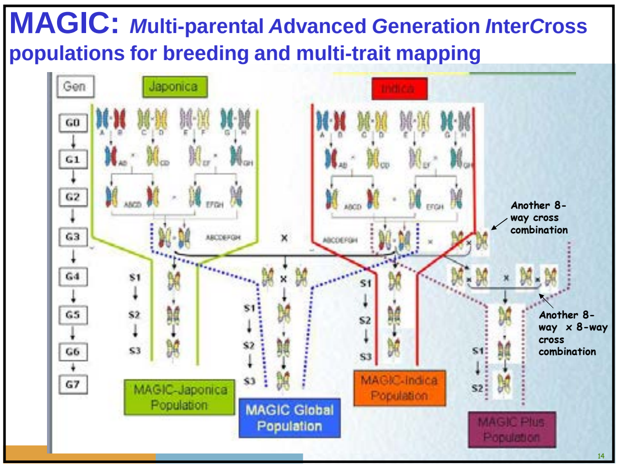
Association genetics diversity panel
2000 lines, divided into 5-6 sub-populations for phenotyping

Field phenomics Measure

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

Genetic stocks for validation
Permanent mapping populations to validate gene-phenotype prediction.

Phenotyping for multiple traits at specialized phenotyping platforms



MAGIC: *M*ulti-parental *A*dvanced *G*eneration *I*nter*C*ross populations for breeding and multi-trait mapping

Another 8-way cross combination

Another 8-way × 8-way cross combination

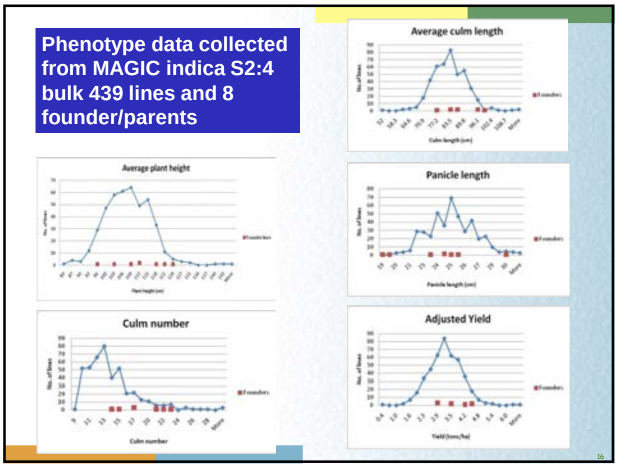
Indica MAGIC

Early maturity

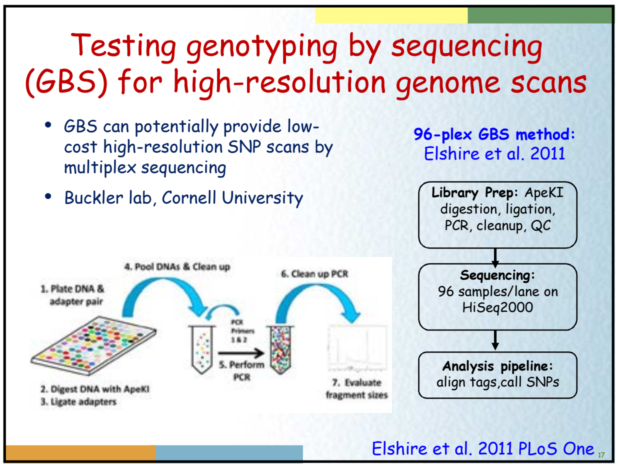
5-row plot

S2:4 Indica MAGIC: 5-row observation and seed production plots (Sept 17, 2011)

Global MAGIC: indica x japonica S1 lines



Phenotype data collected from MAGIC indica S2:4 bulk 439 lines and 8 founder/parents



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

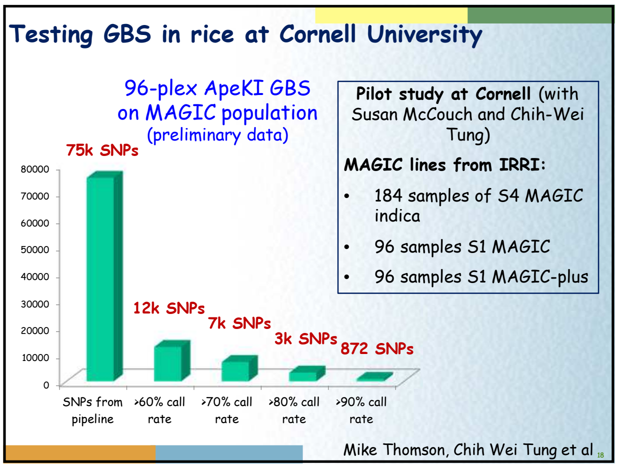
96-plex GBS method:
Elshire et al. 2011

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



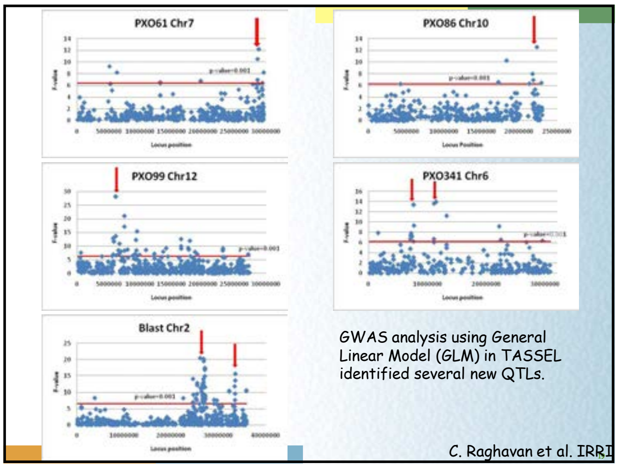
Testing GBS in rice at Cornell University

96-plex ApeKI GBS on MAGIC population (preliminary data)

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

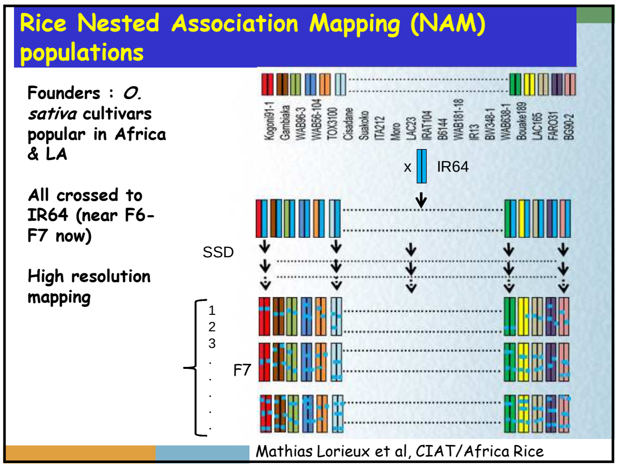
MAGIC lines from IRRI:
- 184 samples of S4 MAGIC indica
- 96 samples S1 MAGIC
- 96 samples S1 MAGIC-plus

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



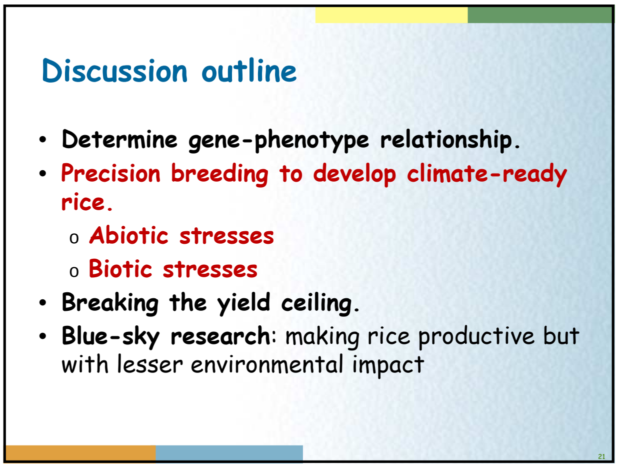
Rice Nested Association Mapping (NAM) populations

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.**
  - **Abiotic stresses**
  - **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

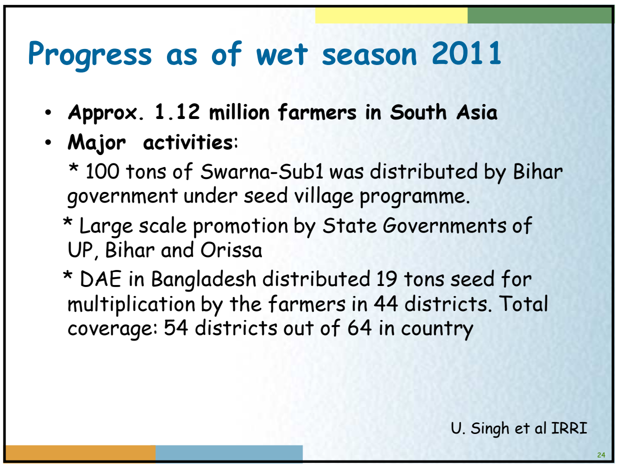
Consequences of Climate Change :

Rice systems will experience more…

- **Drought**
- **Submergence**
- **Salinity**
- **Heat waves**

Challenges of Climate Change = Challenges in unfavorable environments

Scaling up Sub1 varieties



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

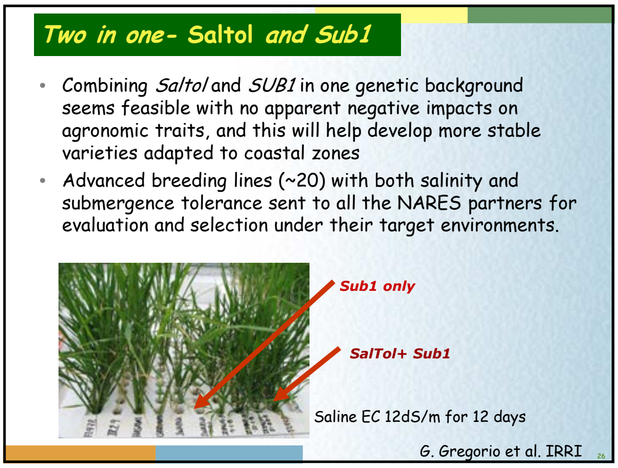
  * DAE in Bangladesh distributed 19 tons seed for multiplication by the farmers in 44 districts. Total coverage: 54 districts out of 64 in country

U. Singh et al IRRI

Mega varieties with SalTol

BR28-SalTol = Available and being tested in the field
IR64-SalTol = Breeder seeds
BR11-SalTol = Breeder seeds
BR29-SalTol= Almost completed

G. Gregorio et al. IRRI

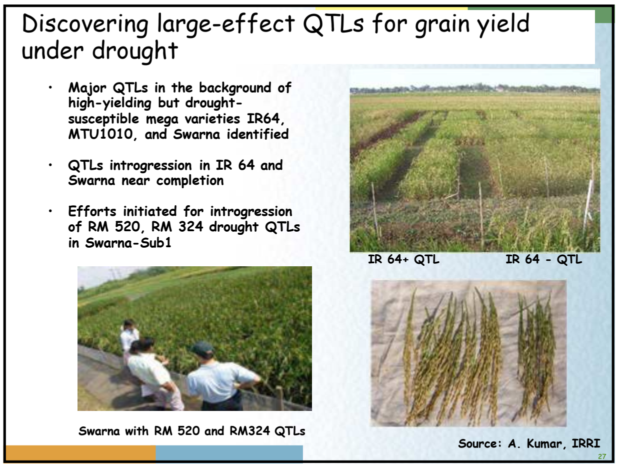


Two in one- Saltol and Sub1

- 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.

Saline EC 12dS/m for 12 days

G. Gregorio et al. IRRI



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**

IR 64+ QTL **IR 64 - QTL**

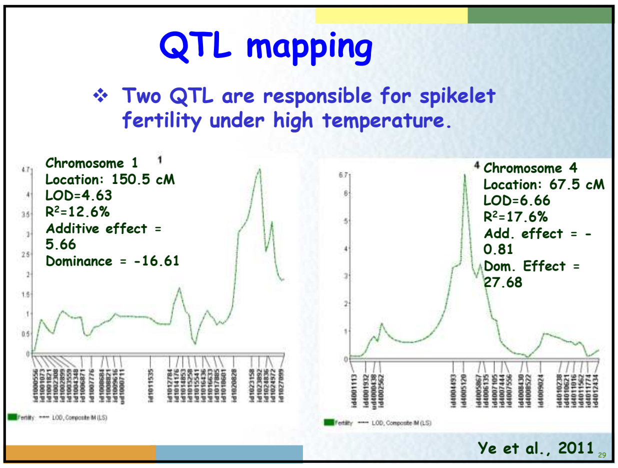
Swarna with RM 520 and RM324 QTLs

Source: A. Kumar, IRRI

Potential heat tolerance donors

Dular	IR28	PSBRc94
Ganga Rata	IR50	RJT174
Ganja Rangwala	IR72	Sadri
Ganjay	IR8866-30-3-1-4-2	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



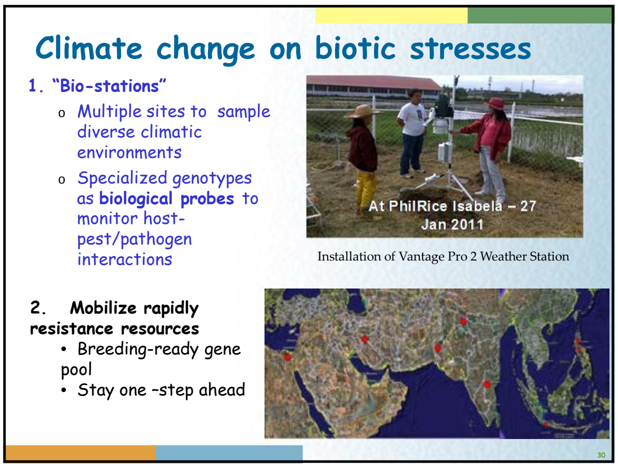
QTL mapping

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

Chromosome 1
Location: 150.5 cM
LOD=4.63
R^2=12.6%
Additive effect = 5.66
Dominance = -16.61

Chromosome 4
Location: 67.5 cM
LOD=6.66
R^2=17.6%
Add. effect = -0.81
Dom. Effect = 27.68

Ye et al., 2011



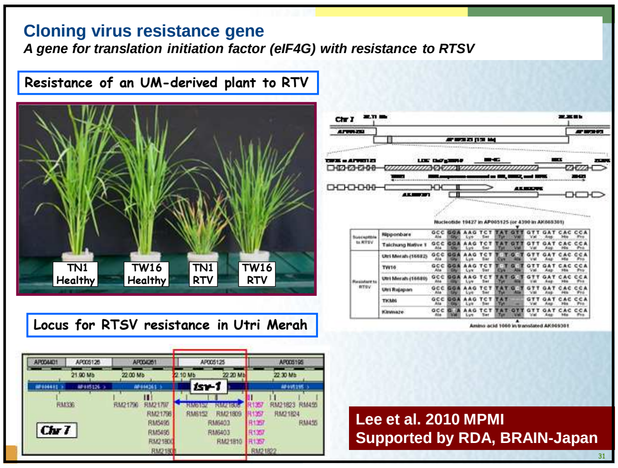
Climate change on biotic stresses

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

At PhilRice Isabela – 27 Jan 2011

Installation of Vantage Pro 2 Weather Station

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



Cloning virus resistance gene

A gene for translation initiation factor (eIF4G) with resistance to RTSV

Resistance of an UM-derived plant to RTV

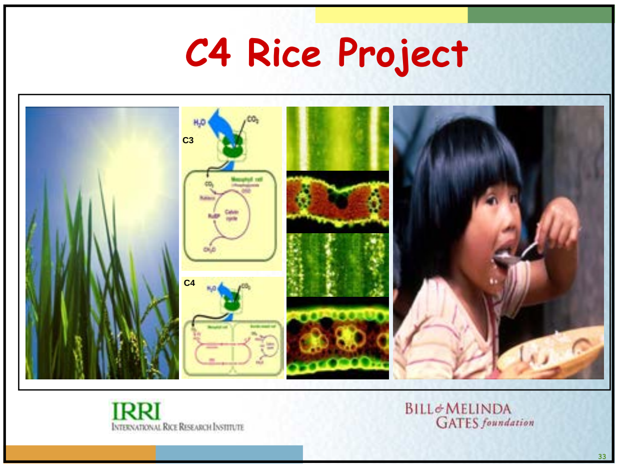
Locus for RTSV resistance in Utri Merah

Lee et al. 2010 MPMI
Supported by RDA, BRAIN-Japan

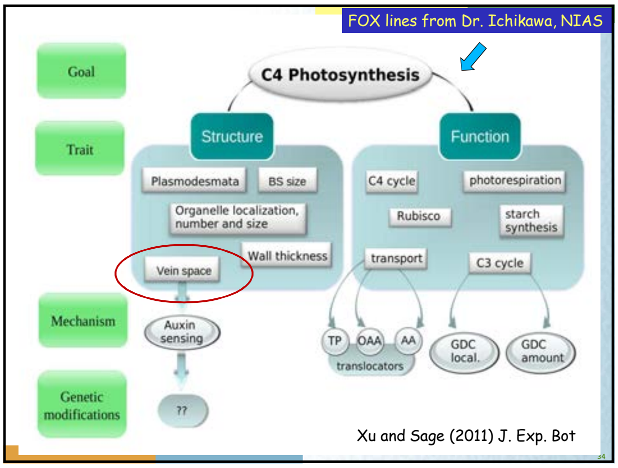
A GRiSP Workshop on Yield Potential

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



FOX lines from Dr. Ichikawa, NIAS

Goal
Trait
Mechanism
Genetic modifications

C4 Photosynthesis

Structure
Plasmodesmata
BS size
Organelle localization, number and size
Vein space
Wall thickness

Function
C4 cycle
photorespiration
Rubisco
starch synthesis
transport
C3 cycle

Auxin sensing
??
TP
OAA
AA
translocators
GDC local.
GDC amount

Xu and Sage (2011) J. Exp. Bot

Aligning research collaboration under GRiSP

GRiSP Product	Research activity	Japan collaboration
Genetic diversity and gene-phenotype relationship	Association genetics Phenotyping Network Specialized recombinant populations (MAGIC, NAM) Genebank sequencing	Root phenotyping (NIAS) Soil nutritent deficiency-P, Zn (JIRCAS) Genome assembly and annotation (?)
Climate-ready rice (gene discovery and breeding)	Submergence Salinity Drought Heat Biotic stresses	Deep root QTL (NIAS) Drought transcriptome (NIAS) Early-flowering (JIRCAS) Virus resistance (NIAS)
Yield Potential	Large-effect genes for yield components , Recurrent Selection and Genomic Selection, Wild species Heterosis Physiological traits.	WISH--large yield QTL (Nagoya Univ)
C4 Rice	Mutant screening	FOX lines (cDNA over-expression) (NIAS)

Acknowledgment

Abiotic stress and breeding
Glenn Gregorio, Ed Redona, Changrong Ye, RK Singh, Arvind Kumar

Biotic stress
Chitra Raghavan, Casiana Vera Cruz, IR Choi

Genetic resources
Ken McNally, Ruaraidh Sackville-Hamilton. R. Mauleon, C. Liang
Mike Thomson

Yield potential
Parminder Virk et al

C4 Rice
Paul Quick et al.

NAM population
Mathias Lorieux, IRD/CIAT

Financial Support
Generation Challenge Program
Rice SNP Consortium
Japan Rice Breeding Project
USAID Linkage Program
STRASA project (BMGF)
C4 Rice (BMGF)