



# Molecular markers for starch quality in rice

Harold Corke<sup>1</sup>, Mei Sun<sup>1</sup>, Jinsong Bao<sup>2</sup>

<sup>1</sup>School of Biological Sciences, The University of Hong Kong, Hong Kong SAR, China

<sup>2</sup>Institute of Nuclear Agricultural Sciences, Zhejiang University, Hua Jiachi Campus, Hangzhou 310029, China

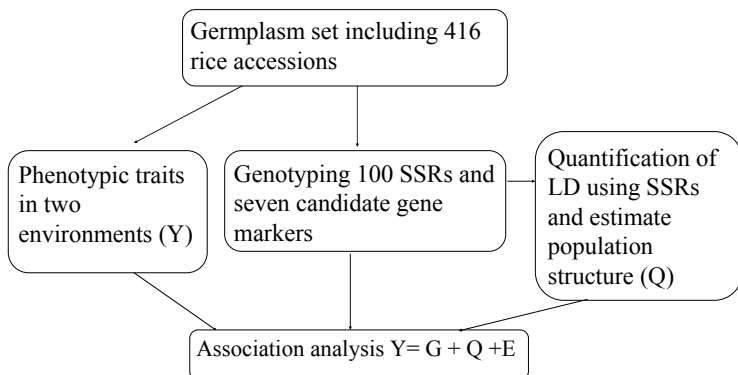
## Abstract

We investigated population structure of 416 rice accessions mostly collected in China. Substantial phenotypic and genotypic diversity in this germplasm set was found. Association analysis based on population structure using molecular markers for starch biosynthesis genes indicated that the *Wx* and *SSIIa* genes are strongly associated with apparent amylose content (AAC) and pasting temperature (PT) ( $P < 0.001$ ). In addition, genome-wide association mapping identified five and seven SSRs that were associated with AAC and PT, respectively.

## Introduction

- Rice starch, comprising ~90% of total grain dry weight of milled rice, strongly influences cooking and eating quality. DNA markers are widely accepted as potentially valuable tools for crop improvement in rice.
- Identifying genetic variants underlying rice quality is a challenging task facing plant breeders. Linkage mapping, and more recently, association or linkage disequilibrium (LD) mapping have been applied to elucidate the genetic basis of natural variation in important quantitative traits.
- In conventional linkage mapping studies, the LD required for mapping quantitative trait loci is generated as a result of the mating design, whereas association studies exploit the LD already present in the population of interest.
- We evaluated the population structure and genetic diversity of a set of 416 rice materials mainly collected in China, developed functional markers for starch biosynthesis genes, and conducted association mapping for starch quality based on the estimated population structure.

## Materials & Methods

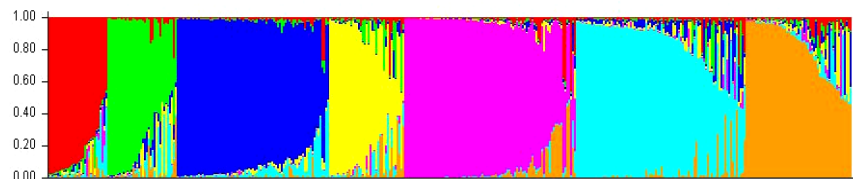


**Figure 1. The scheme and procedures of association mapping, including germplasm choice, estimation of population structure, trait evaluation, identification of candidate polymorphisms, and statistical analysis**

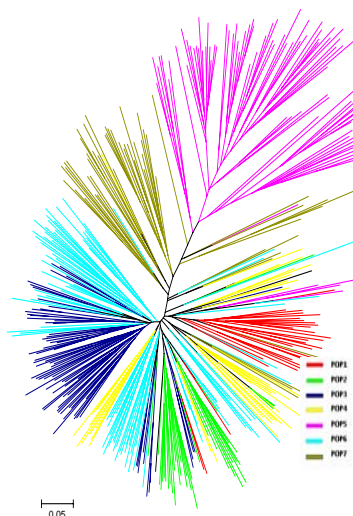
## Results

A total of 100 SSR markers, randomly distributed across the genome, were used to evaluate the genetic diversity of the rice population. All of the 100 SSR markers were polymorphic across the 416 rice accessions and a total of 390 alleles were detected. The average number of alleles per locus was 3.9, ranging from 2 to 9. The average gene diversity was 0.4736, ranging from 0.0473 to 0.7618. The average PIC value was 0.4214, ranging from 0.0466 to 0.7216.

The model-based simulation of population structure using SSRs showed that the likelihood was maximized when the number of populations was set at seven, suggesting that these rice accessions can be grouped into seven subpopulations, as inferred from the model, here denoted as POP1, POP2, POP3, POP4, POP5, POP6 and POP7, respectively (Fig. 2). All *japonica* rice clustered into POP5.



**Figure 2. Population structure of 416 rice accessions based on 100 SSRs ( $K=7$ )** Each accession is represented by a thin vertical line, which can be partitioned into seven colored segments that represent the probabilities of individual membership (Q) to the seven clusters, and then all accessions were sorted by Q.



**Figure 3. Unrooted neighbor-joining tree of 416 rice accessions based on Nei's genetic distance**

The neighbor-joining tree revealed genetic relationships reasonably consistent with the STRUCTURE model-based membership assignment. However, a few rice accessions were displaying as admixtures in different clusters. These "misplaced" rice accessions were breeding lines derived from *indica/japonica* crosses. With few exceptions, the neighbor-joining tree showed that the rice accessions could be well differentiated according to their subspecies affiliation, i.e. *indica* or *japonica*.

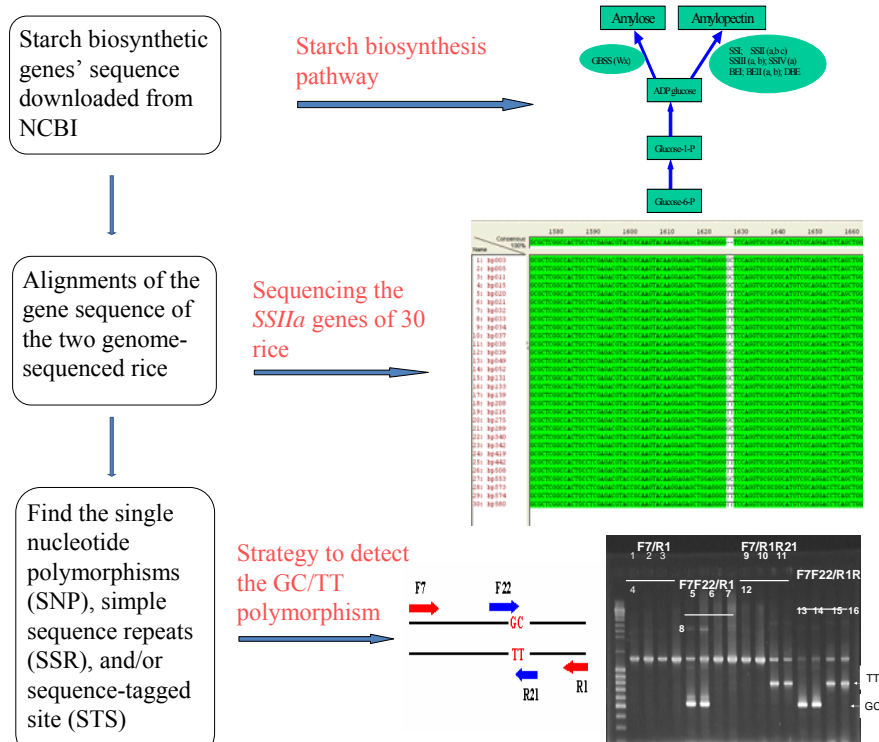
Substantial variation existed in nearly all the agronomic and quality traits measured in this diverse germplasm set of 416 rice accessions grown in 2006 and 2007 (Table 1).

Seven starch biosynthesis gene markers have been genotyped (Fig. 4) (Bao et al. 2006a,b). We performed structure-based association mapping for starch quality using both candidate gene mapping and genome-wide association mapping. In candidate gene mapping, the *Wx* and *SSIIa* genes could be identified as strongly associated with AAC and PT respectively ( $P < 0.001$ ). Genome-wide association mapping identified five and seven SSRs that were associated with AAC and PT, respectively. Among them, RM253, RM276 and RM484 have been detected by QTL mapping before. Two markers, RM253 and RM276, closely linked to the *SSIIa* gene, explained 5-10% of the variation of PT. Their associations with PT could be a result of LD with *SSIIa*.

**Table 1** Descriptive statistics and percentage of phenotypic variation explained by population structure for 25 agronomic traits

Trait	Year	Mean $\pm$ SD	Range	$H^2$	$R$
<b>Agronomic traits</b>					
Heating date (days)	2006	73.2 $\pm$ 12.5	50.0-126.0	0.976	34.4
	2007	73.4 $\pm$ 12.4	51.0-119.0		32.7
Plant height (cm)	2006	105.3 $\pm$ 27.0	54.3-182.4	0.908	46.6
	2007	110.8 $\pm$ 30.2	53.0-188.1		42.9
Flag leaf length (cm)	2006	36.2 $\pm$ 9.5	15.4-77.3	0.782	23.1
	2007	36.8 $\pm$ 9.0	15.9-66.7		23.8
Flag leaf width (cm)	2006	1.73 $\pm$ 0.27	0.90-2.68	0.802	9.3
	2007	1.70 $\pm$ 0.26	1.00-2.63		7.9
Grain length (mm)	2006	8.44 $\pm$ 0.90	6.76-11.18	0.686	31.9
	2007	8.52 $\pm$ 0.91	6.77-11.21		30.7
Grain width (mm)	2006	2.99 $\pm$ 0.35	2.19-4.07	0.711	39.1
	2007	3.03 $\pm$ 0.36	2.26-4.86		42.1
Grain length/width ratio	2006	2.87 $\pm$ 0.56	1.87-4.64	0.991	40.1
	2007	2.86 $\pm$ 0.55	1.83-4.47		40.9
Panicle length (cm)	2006	24.2 $\pm$ 4.0	13.1-38.4	0.918	18.6
	2007	23.3 $\pm$ 3.6	13.1-23.3		15.5
1000-grain weight (g)	2006	23.2 $\pm$ 3.3	12.9-33.4	0.906	6.2
	2007	22.8 $\pm$ 4.6	12.6-34.8		6.0
<b>Starch quality traits<sup>b</sup></b>					
Apparent amylose content, AAC (%)		23.1 $\pm$ 5.7	1.6-32.6	na	41.5
Peak viscosity (RVU)		242 $\pm$ 33.5	81.7-305.4	na	11.5
Hot paste viscosity (RVU)		173 $\pm$ 30.9	43.0-248.1	na	14.9
Cool paste viscosity (RVU)		308 $\pm$ 51.8	56.5-409.1	na	21.3
Breakdown (RVU)		69 $\pm$ 19.7	19.6-141.3	na	18.2
Setback viscosity (RVU)		69 $\pm$ 37.1	-82.6-193.2	na	27.0
Pasting temperature, PT ( $^{\circ}$ C)		74.9 $\pm$ 2.8	67.5-81.5	na	18.3
Hardness (g)		29.2 $\pm$ 11.1	0.8-61.1	na	43.6
<b>Grain color traits</b>					
$L^*$	2006	51.8 $\pm$ 9.5	14.1-67.4	na	8.5
$a^*$	2006	3.46 $\pm$ 2.90	-12.9-13.8	na	21.1
$b^*$	2006	14.3 $\pm$ 2.1	-1.8-18.7	na	6.7
$C$	2006	15.1 $\pm$ 1.4	9.48-19.7	na	9.5
$H^{\circ}$	2006	72.5 $\pm$ 31.0	-193.3-101.4	na	7.0
<b>Nutritional quality traits</b>					
Phenolics (mg GAE/l)	2006	208.1 $\pm$ 155.6	108.1-1244.9	na	8.0
Flavonoids (mg RE/l)	2006	134.6 $\pm$ 20.6	88.51-286.29	na	5.0
Antioxidant capacity (mMTAEC) 2006		0.474 $\pm$ 0.752	0.0120-5.533	na	8.4

<sup>a</sup>Percentage of phenotypic variation explained by population structure. <sup>b</sup>Data from Bao et al. (2006a, b). na, not available



**Figure 4.** Strategies to discover functional molecular markers for starch biosynthetic genes

**Table 2** Association of starch synthesizing gene markers (left) and SSRs (right) with apparent amylose content (AAC) and pasting temperature (PT) based on population structure

Gene markers	AAC		PT		Markers	Chromosome	Position	$P$	$R^2$
	$P$	$R^2$	$P$	$R^2$					
<b>AAC</b>									
<i>Wx</i> SNP	1.02E-105	0.3979	1.04E-04	0.0313	RM122	5	0	5.01E-04	0.0217
<i>Wx</i> SSR	2.64E-95	0.3968	0.008	0.0464	RM161	5	96.9	4.00E-04	0.0225
					RM346	7	47	2.70E-06	0.0444
					RM336	7	61	0.001	0.0328
<i>SSIIa</i> SSR	6.51E-11	0.0689	0.1584	0.0109	RM126	8	57	4.13E-04	0.0187
<b>PT</b>									
<i>SSIIa</i> SNP	3.73E-04	0.0188	4.43E-65	0.4312	RM253	6	37	3.88E-05	0.0520
<i>Sbe1</i> SSR	0.0217	0.0141	0.0162	0.0217	RM276	6	40.3	1.13E-10	0.1042
					RM346	7	47	2.48E-04	0.0449
<i>Sbe1</i> STS	0.0016	0.0145	0.0034	0.0179	RM278	9	77.5	1.38E-05	0.0523
<i>Sbe3</i> SNP	0.8005	9.39E-05	0.9824	1.04E-06	RM222	10	11.3	7.50E-04	0.0395
					RM484	10	97.3	9.76E-04	0.0295
					RM206	11	102.9	5.52E-04	0.0460

## References

- Bao JS, Corke H, Sun M (2006a) Nucleotide diversity in starch synthase IIa and validation of single nucleotide polymorphisms in relation to starch gelatinization temperature and other physicochemical properties in rice (*Oryza sativa* L.). *Theor Appl Genet* 113: 1171-1183
- Bao JS, Corke H, Sun M (2006b) Microsatellites, single nucleotide polymorphisms and a sequence tagged site in starch-synthesizing genes in relation to starch physicochemical properties in nonwaxy rice (*Oryza sativa* L.). *Theor Appl Genet* 113: 1185-1196