Mol Breeding (2008) 21:127–137
DOI 10.1007/s11032-007-9141-7

The QTL controlling amino acid content in grains of rice (*Oryza sativa*) are co-localized with the regions involved in the amino acid metabolism pathway

Lingqiang Wang · Ming Zhong · Xianghua Li · Dejun Yuan · Yunbi Xu · Huifang Liu · Yuqing He · Lijun Luo · Qifa Zhang

Received: 14 December 2006 / Accepted: 8 October 2007 / Published online: 31 October 2007

Abstract The improvement of grain quality, such as protein content (PC) and amino acid composition, has been a major concern of rice breeders. We constructed a population of 190 recombinant inbred lines (RILs) from a cross between Zhenshan 97 and Nanyangzhan to map the quantitative trait locus or loci (QTL) for amino acid content (AAC) as characterized by each of the AACs, total essential AAC, and all AAC. Using the data collected from milled rice in 2002 and 2004, we identified 18 chromosomal regions for 19 components of AAC. For 13 of all the loci, the Zhenshan 97 allele increased the trait values. Most QTL were co-localized, forming ten QTL clusters in 2002 and six in 2004. The QTL clusters varied in both effects and locations, and the mean values of variation explained by individual QTL in the clusters ranged from 4.3% to 28.82%. A relatively strong QTL cluster, consisting of up to 19 individual QTL, was found at the bottom of chromosome 1. The major QTL clusters identified for two different years were coincident. A wide coincidence was found between the QTL we detected and the loci involved in amino acid metabolism pathways, including N assimilation and transfer, and amino acid or protein biosynthesis. The results will be useful for candidate gene identification and marker-assisted favorable allele transfer in rice breeding programs.

Keywords Amino acid content · Grain quality · Lys content · QTL · Rice (*Oryza sativa* L.)

L. Wang · M. Zhong · X. Li · D. Yuan · H. Liu · Y. He (✉) · Q. Zhang
National Key Laboratory of Crop Genetic Improvement, National Center of Plant Gene Research and National Center of Crop Molecular Breeding, Huazhong Agricultural University, Wuhan 430070, China
e-mail: yqhe@mail.hzau.edu.cn

Y. Xu
Genetic Resources Program, International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, 06600 Mexico, DF, Mexico

L. Luo
Shanghai Agrobiological Gene Center, Shanghai 201106, China

Introduction

Rice (*Oryza sativa* L.) is one of the most widely grown crops and a main staple food for about half the world's population. The nutritional quality of rice grain is important to all rice consumers, especially where it is the population's main staple. The contents of protein and amino acids are the major factors of nutritional quality, and their regulation has increasingly become a major breeding objective. Lin et al. (1993) reported that rice protein content (PC) varied from 4.9% to 19.3% in *indica* and from 5.9% to 16.5% in *japonica*. The genetic variation in PC provides a basis for breeding PC. However, manipulating this trait in traditional breeding is difficult

because such substantial variation is quantitatively inherited (Kambayshi et al. 1984; Sood and Siddiq 1986; Gupta et al. 1988; Shenoy et al. 1991).

With the advent of molecular markers such as restriction fragment length polymorphisms (RFLPs) and microsatellites and their maps, quantitative trait locus or loci (QTL) that control quality traits can be dissected. Several recent studies have been undertaken to decipher the genetic basis of PC in rice by QTL mapping (Tan et al. 2001; Yoshida et al. 2002; Aluko et al. 2004; Hu et al. 2004), providing useful information for improving the nutritional quality of rice. However, some problems still remain. In previous studies, PC was calculated from Kjeldahl nitrogen multiplied by a factor of 5.95, which is based on the nitrogen content (16.8%) of the major rice protein glutelin, so it can only indicate the relative quantity of protein. In fact, the amino acids in total protein of grain can be unbalanced because certain essential amino acids are extremely low. Lysine (Lys) and threonine (Thr) were identified as the first and second limiting essential amino acids of milled rice protein, based on human requirements as estimated by the World Health Organization (1973). Therefore, special attention should be paid to improving the quality of rice protein, which is usually represented by essential amino acid content (AAC) and essential amino acid index (FAO 1970). Unfortunately, little is known about the genetic basis of the AAC in rice, particularly the genes associated with their metabolism pathways.

The present study uses a recombinant inbred line (RIL) population planted in two different years to dissect the genetic basis of the AAC in milled rice as illustrated by each amino acid, the total contents of essential amino acids and all amino acids. The results will help rice breeders develop strategies for improving the quality of rice grain.

Materials and methods

The mapping population and the field experiment

The mapping population consisting of 190 RILs was developed from a cross between Zhenshan 97 and Nanyangzhan. The former is the female parent for a number of widely cultivated hybrids in China while the latter is a local red grain variety. The field experiment was conducted during the rice-growing seasons of 2002 and 2004 on the experimental farm at Huazhong Agricultural University, Wuhan, China.

About 20 plants per line were transplanted in two rows with 16.5 cm between plants and 26.4 cm between rows. Field management essentially followed normal agricultural practice, with fertilizer applied (per hectare) as follows: 48.75 kg nitrogen, 58.5 kg phosphorous, and 93.75 kg potassium as the basal fertilizer; 86.25 kg nitrogen at the tilling stage; and 27.6 kg nitrogen at the booting stage.

Determination of the amino acid content

Preparation of head rice and flour followed the methods described by Tan et al. (1999), and samples were refrigerated until analysis. About 100 mg of each sample was hydrolyzed with 6 ml of 6 M hydrochloric acid (HCl) in a sealed air-evacuated tube at 110°C for 22 h. The hydrolysate was diluted to 50 ml, then 1 ml dilution was transferred to a centrifugal tube. The centrifugal tube was placed in a rotary evaporator to remove the HCl and water from the hydrolysate. The residue was shaken vigorously and completely dissolved in 1 ml of 0.02 M HCl, then centrifuged at 14,000 rpm for 15 min. About 0.8 ml of supernatant was placed in an autosampler bottle and analyzed by an amino acid autoanalyzer (model L-8800 Hitachi). The amount of each amino acid in the sample was calculated with reference to the standard sample in units of milligram/gram. The assay for each line was conducted with two replicates. The glutamic and aspartic acids detected here included the fractions derived from glutamine and asparagine. Tryptopan (an essential amino acid) was destroyed by HCl hydrolysis and was not detected by the method used here, so the total content of essential amino acids was calculated as the sum of the nine detectable essential amino acids: Thr, valine (Val), methionine (Met), isoleucine (Ile), leucine (Leu), phenylalanine (Phe), Lys, histadine (His), and arginine (Arg).

DNA markers and assays

About 190 polymorphic simple sequence repeat (SSR) markers covering all 12 chromosomes were

used to genotype the population. The primers of the rice microsatellite (RM) series were designed according to Temnykh et al. (2000, 2001), and those of the monsanto rice genome (MRG) series were designed according to the rice genome sequences of the Monsanto Company (McCouch et al. 2002). The SSR assay was performed essentially as described by Jiang et al. (2004).

Data analysis

Mapmaker 3.0 (Lincoln et al. 1992) was used to construct a genetic linkage map. The average of the measurements for each line was used for QTL analysis. The means of the traits were used to identify QTL by composite interval mapping (Zeng 1994) methodologies with QTLCartographer 2.0 software (Basten et al. 1997). The permutation method was used to obtain the thresholds (1.89–2.38 for these traits) of the experiment based on 1,000 runs of randomly shuffling the trait values ($P = 0.05$). The peak points of the likelihood rate (LR test statistic) in the linkage map were taken as the putative positions of the effects, and additive effects were taken from the points showing the largest effects. The relative contribution of a genetic component was calculated as the proportion of phenotypic variation explained by that component. The statistic analysis, which included a *t*-test and a correlation analysis, was conducted with the statistical package Statistica (StatSoft 1991).

The determination of the map location of genes involved in amino acid metabolism pathways

The DNA sequences of available genes associated with seed storage proteins (glutelin, prolamin, globulin, and albumin) and the amino acids of the aspartate family were mainly identified through the literature (Duan and Sun 2005). The chromosome locations of the some loci for nitrogen assimilation and transfer, which were also deduced on the basis of genomic sequences, were taken from a previous study by Lian et al. (2005). The map locations of the other loci and of the genes cited from literature were determined or verified by searching the genes on the gramene website (http://www.gramene.org).

Results

Phenotypic variation and correlation analysis

In both years, significant phenotypic differences between the parents for the all 19 components of AAC were detected using a *t*-test ($P < 0.01$), in which Zhenshan 97 had higher values for all components (Table 1). The RIL population showed transgressive segregations in both directions for all components.

Significant positive correlations between the various components, except the contents of Cys, Met, and Tyr, were found in milled rice (r values 0.8568–0.9934), while these three components (the contents of Cys, Met, and Tyr) showed weaker or no correlations with the contents of the other amino acids, especially of Ser, Ile, His, and praline (Pro) (data not shown).

Linkage map

A linkage map was constructed consisting of 190 SSR markers spanning a total of 1,362.1 cM, with an average interval of 7.77 cM between adjacent markers (Fig. 1). The marker orders in the map were in good agreement with those of Temnykh et al. (2000, 2001).

QTL analysis

QTL for amino acid content in milled rice in 2002

About 80 individual QTL were identified for the 19 components of AAC in milled rice in 2002, ranging from one to six QTL for each component, with an average of 4.21. The total variation explained by the QTL for each component varied from 7.4% for Met content to 59.1% for Thr content, with an average of 38.64%. For Lys content, five QTL distributed on chromosome 1, 7, and 8 were detected, explaining 50% of the variation. Almost all QTL, except the one flanked by MRG4302-RM323 on chromosome 1, were co-mapped with the QTL for other components (Table 2 and Fig. 1). For the two QTL located in MRG4302-RM323 (1–2) and in RM137-RM556 (8–10), the Nanyangzhan alleles increased their values, whereas for the other three QTL located in RM472-RM104 (1–19), RM125-RM542 (7–4, 5), and MRG2702-RM38 (8–3), the Zhenshan 97 alleles

Table 1 Descriptive statistics of the 19 components of AAC in the parents and RIL population observed in rice grain

Traits	Year 2002			Year 2004		
	Parents		Population	Parents		Population
	Zhenshan 97	Nanyangzhan	Range	Zhenshan 97	Nanyangzhan	Range
Asp	12.95 ± 0.10	9.44 ± 0.14	8.08–15.08	12.81 ± 0.15	7.93 ± 0.07	7.52–15.05
Thr	4.86 ± 0.03	3.70 ± 0.05	3.13–5.38	4.77 ± 0.05	3.20 ± 0.04	3.01–5.73
Ser	6.54 ± 0.06	5.14 ± 0.08	4.28–7.80	6.48 ± 0.05	4.41 ± 0.07	4.06–8.48
Glu	26.90 ± 0.20	20.25 ± 0.29	16.55–30.93	27.21 ± 0.24	16.91 ± 0.17	15.84–34.05
Gly	6.01 ± 0.05	4.50 ± 0.06	3.88–6.63	6.12 ± 0.05	3.85 ± 0.02	3.63–7.16
Ala	7.71 ± 0.06	5.64 ± 0.08	4.83–9.23	7.61 ± 0.06	4.77 ± 0.03	4.56–8.85
Cys	3.03 ± 0.10	2.60 ± 0.05	2.40–3.93	2.87 ± 0.29	2.39 ± 0.03	1.83–3.61
Val	7.59 ± 0.07	5.74 ± 0.05	4.95–8.53	8.19 ± 0.16	5.04 ± 0.04	4.86–9.66
Met	2.81 ± 0.02	2.70 ± 0.01	2.03–3.55	2.26 ± 0.07	2.06 ± 0.04	1.42–3.49
Ile	5.37 ± 0.06	4.05 ± 0.09	3.40–6.15	6.17 ± 0.13	3.52 ± 0.03	3.42–7.28
Leu	11.39 ± 0.10	8.51 ± 0.13	7.20–12.53	11.76 ± 0.10	7.43 ± 0.04	6.82–13.94
Tyr	5.35 ± 0.17	3.91 ± 0.13	3.55–6.78	5.15 ± 0.12	3.95 ± 0.08	2.48–6.90
Phe	7.77 ± 0.04	5.76 ± 0.06	4.85–8.70	8.11 ± 0.06	4.81 ± 0.05	4.60–9.04
Lys	4.73 ± 0.03	3.42 ± 0.04	3.10–5.60	5.09 ± 0.04	3.16 ± 0.02	3.08–5.71
His	3.20 ± 0.03	2.42 ± 0.03	2.05–3.78	3.86 ± 0.03	2.16 ± 0.09	2.10–4.33
Arg	11.52 ± 0.10	8.33 ± 0.09	7.03–12.68	11.00 ± 0.12	7.13 ± 0.07	5.95–13.66
Pro	5.87 ± 0.07	4.38 ± 0.07	2.93–6.38	5.73 ± 0.05	3.50 ± 0.02	3.36–6.99
Eaa	59.23 ± 0.57	44.62 ± 0.48	37.93–65.03	61.21 ± 0.59	38.43 ± 0.27	35.26–71.50
Total	133.59 ± 0.74	100.48 ± 1.04	85.80–154.13	135.20 ± 0.87	86.14 ± 0.74	78.85–160.20

Eaa, the total amount of essential amino acids; Total, content of all amino acid (17 kinds here)

increased their values. For the total content of essential amino acids, four QTL located on chromosomes 1, 7, and 8 were identified, jointly explaining 37% of the variation; the positive alleles all came from Zhenshan 97.

Most of the QTL detected tended to be co-localized within the genome, and thus 13 chromosomal regions were revealed, including 10 QTL clusters and three single QTL (Table 2 and Fig. 1). The allele effects of the QTL in the same QTL cluster were all in the same direction. The QTL cluster on chromosome 7 was the largest in number, consisting of 14 individual QTL, followed by the QTL clusters on chromosomes 1, 8, and 10 (each containing 11 individual QTL). However, considering the average variation explained by the individual QTL in one cluster, the QTL cluster in interval RM322-RM521 on chromosome 2 was the largest (12.1 ± 2.9%), followed by the QTL clusters on chromosome 1 (10.2 ± 3.7%) and on chromosome 8 (9.3 ± 2%). The relative effect of the QTL for Lys content in a QTL cluster, calculated as the additive effect of the QTL for Lys content divided by the additive effect of the QTL for total AAC, differed among the QTL clusters. The QTL on chromosome 1 had the largest relative effect (5.42%), followed by the QTL in RM125-RM542 (7–4, 5) (4.34%), the QTL in MRG2702-RM38 (8–3) (4.11%), and the QTL in RM137-RM556 (8–10) (3.82%), and a similar tendency could be observed in the relative variation explained (data not shown). In most loci, except those in RM520-RM468 (3–15), RM348-RM131 (4–13), and RM137-RM556 (8–10), the positive alleles originated from Zhenshan 97. This provided a ready explanation for the higher values in Zhenshan 97 and the transgressive segregation in the population in 2002.

QTL for amino acid content in milled rice in 2004

About 75 individual QTL were identified for the 19 components of AAC in milled rice in 2004, ranging

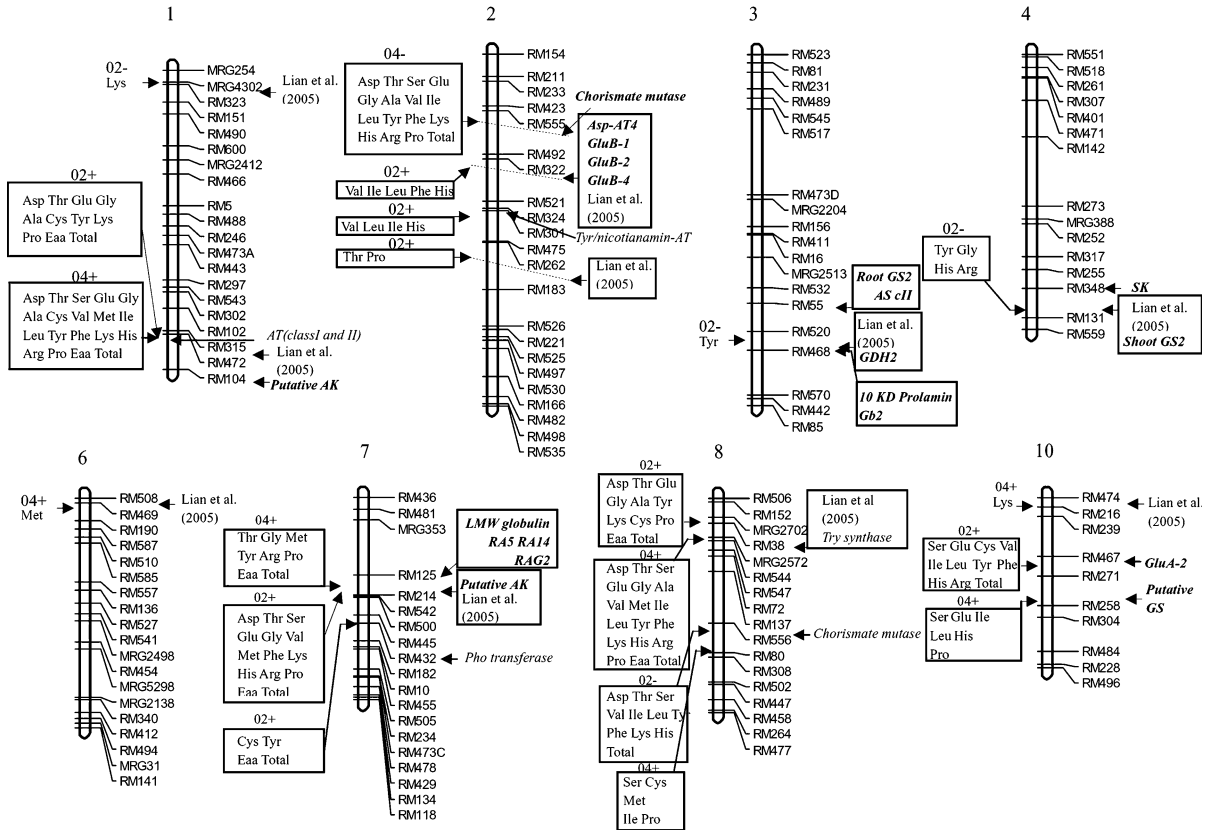

Fig. 1 Distribution of the QTL for 19 components of AAC on the linkage map. The marker name is shown on the right of the chromosome. The numbers on the pane (QTL cluster) or a single QTL indicate the years in which the loci were detected; and the plus and minus signs indicate that the positive alleles came from *Zhenshan 97* and *Nanyangzhan*, respectively

from two to six QTL for each component, with an average of 3.95. The total variation explained by the QTL for each component varied from 19.92% for Cys content to 63.86% for Pro content, with an average of 52.7%. Four QTL were detected for the Lys content, with the largest located in the interval RM472-RM104 on chromosome 1. All QTL for Lys content, except the one in 10-1, were co-mapped with the QTL for other components (Table 3 and Fig. 1). Four QTL were identified for the content of essential amino acids, located in RM472-RM104 (1–19), RM555-RM4922 (2–4, 5), RM125-RM214 (7–4), and MRG2572-RM544 (8–5). The positive alleles in all these loci, except the locus on the short arm of chromosome 2, came from Zhenshan 97.

Comparison of the location of QTL revealed eight chromosomal regions, including two single QTL and six QTL clusters that were distributed on five chromosomes (Table 3 and Fig. 1). The allele effects of the QTL in the same QTL cluster were all in the same direction. For most QTL clusters, except the one on chromosomes 2, the positive alleles originated from Zhenshan 97. Among the six QTL clusters, the largest one in number was in the interval RM472-RM104 on chromosome 1, consisting of 19 individual QTL, followed by the second one on the short arm of chromosome 8 (containing 18 individual QTL), and the other two on chromosome 2 (containing 17 individual QTL) and on chromosome 7 (containing eight individual QTL). A similar tendency could be observed in the average variation explained by the individual QTL in these clusters. The largest one was also in the interval RM472-RM104 on chromosome 1, with mean values of 22.82 ± 5.36%, followed by the second one on chromosome 2 (11.70 ± 1.85%). The relative effect of QTL for Lys content compared with total AAC differed among the QTL clusters. The

Table 2 QTL identified by QTLCartographer for the 19 components of AAC in milled rice in 2002

Chr interval	Flanking marker	Traits	Peak (cM)	LOD	Additive effects	Var (%)
1–2	MRG4302-RM323	Lys	6.2	2.2	−0.12	5.6
1–19[a]	RM472-RM104	Asp/Thr/Glu/Gly/Ala/Cys/Tyr/Pro/Eaa/Total	150.2	2.0–4.5	0.40/0.17/0.74/0.23/0.38/0.08/0.22/0.18/2.4/3.8	5.7–14.5
		Lys	152.2	5.0	0.21	15.9
2–7	RM322-RM521	Val/Ile/Leu/His	73.5	2.8–4.5	0.28/0.23/0.37/0.11	10.0–16.7
		Phe	71.5	2.6	0.25	10.1
2–9(10)	RM324-RM301	His	81.6	3.1	0.10	7.9
	RM301-RM475	Val/Ile/Leu	84.9	2.3–3.4	0.19–0.32	7.2–11.0
2–12	RM262-RM183	Thr	114.0	2.3	0.15	9.2
		Pro	120.0	2.2	0.18	8.7
3–15	RM520-RM468	Tyr	146.7	2.5	−0.18	7.1
4–13	RM348-RM131	Thr	130.2	2.3	−0.15	8.6
		Gly/His	132.2	2.0–2.0	−0.17/−0.09	6.6–7.12
		Arg	134.2	2.4	−0.35	8.2
7–4(5)[a]	RM125-RM214	Pro	47.3	2.8	0.19	8.6
		Gly/Met/Arg	49.3	2.1–3.7	0.19/0.09/0.41	7.4–11.6
		Thr/Glu/Val/Phe	51.3	2.2–3.5	0.16/0.79/0.19/0.23	6.2–10.0
	RM214-RM542	Asp/Ser/Lys/His	52.0	2.3–4.1	0.45/0.18/0.16/0.08	5.9–10.4
		Eaa/Total	52.3	2.3–2.7	1.6/3.7	7.0–7.7
7–8(9)	RM445-RM432	Total	65.4	2.1	3.46	6.5
	RM432-RM182	Eaa	71.1	2.6	1.82	9.3
		Cys	73.1	2.5	0.09	8.3
		Tyr	75.1	2.4	0.16	6.9
8–3	MRG2702-RM38	Asp/Thr/Glu/Gly/Ala/Tyr/Lys/Pro/Eaa/Total	19.9	2.4–4.3	0.44/0.17/0.77/0.18/0.43/0.18/0.17/0.18/2.2/4.1	6.9–12.9
8–9(10)	RM137-RM556	Tyr	80.9	2.1	−0.19	8.5
	RM556-RM80	Asp/Ser/Ile/Leu/Phe/His/Total	86.2	2.1–4.4	−0.37/−0.18/−0.17/−0.39/−0.27/−0.12/−3.5	5.5–11.7
		Val/Lys	88.2	2.2–4.5	−0.32/−0.13	6.6–16.0
		Thr	90.2	2.7	−0.17	10.1
8–14	RM447-RM458	Cys	120.7	3.3	−0.10	11.0
10–4	RM467-RM271	Cys/Leu	37.5	2.7–2.9	0.10/0.39	9.4–9.8
		Ile/Phe	39.5	2.1–3.7	0.21/0.23	6.5–11.7
		Ser/Glu/Val/Tyr/His/Arg/Total	41.5	2.1–4.4	0.22/0.80/0.25/0.19/0.13/0.37/3.7	5.9–12.3

Chr interval, the interval where the QTL was located, with Chr being an abbreviation for chromosome; Flanking marker, with the most significant marker underlined; Peak, the QTL points showing the largest effects; Additive effects: a positive value implies that the Zhenshan 97 allele had a positive effect on that trait; Var, the percentage of the phenotypic variations explained by QTL

[a] Locus also detected in another year

one in interval 2–5 had the largest relative additive effect (3.87%), followed by the QTL in RM472-RM104 (1–19) (3.27%) and the QTL in MRG2572-RM544 (8–5) (2.7%). The same tendency could be achieved in the relative variations explained by these QTL.

Co-localization of the QTL for amino acid content with the loci involved in amino acid metabolism pathways

In total, we identified 18 chromosomal regions for 19 components of AAC in rice grain, with ten specific to

Table 3 QTL identified by QTLCartographer for the 19 components of AAC in milled rice in 2004

Chr interval	Flanked marker	Traits	Peak (cM)	LOD	Additive effects	Var (%)
1–19[a]	RM472-RM104	Ser/Glu/Pro	150.2	12.2–16.0	0.48/1.76/0.37	27.07–33.37
		Asp/The/Gly/Ala/Val/Met/Ile/Leu/Tyr/Phe/His/Arg/Eaa/Total	152.2	6.5–13.0	0.82/0.31/0.37/0.49/0.48/0.17/0.42/0.78/0.41/0.50/0.23/0.78/3.95/8.89	17.53–33.3
		Cys/Lys	154.2	3.5–9.6	0.14/0.29	12.01–29.17
2–4(5)	RM423-RM555	Gly/Ala	27.8	4.4–4.7	−0.21/−0.26	9.22–9.73
		Tyr	29.8	3.6	−0.19	7.45
	RM555-RM492	Asp/Thr/Ser/Val/Ile/Leu/His/Pro	31.9	4.2–5.8	−0.48/−0.19/−0.28/−0.31/−0.26/−0.46/−0.13/−0.23	9.8–12.9
		Glu/Phe/Lys/Arg/Eaa/Total	33.9	4.7–5.3	−1.17/−0.33/−0.20/−0.49/−2.55/−5.59	12.97–14.45
6–2	RM469-RM190	Met	10.3	2.5	−0.10	6.22
7–4[a]	RM125-RM214	Met/Tyr/Pro	41.3	2.3–3.5	0.12/0.16/0.15	4.68–7.69
		Thr/Gly/Arg/Total	43.3	2.0–3.5	0.11/0.17/0.37/3.15	4.05–7.08
		Eaa	47.3	1.8	1.41	3.98
8–4(5)	RM38-MRG2572	Met	34.9	5.0	0.18	18.27
	MRG2572-RM544	Asp/Thr/Ser/Glu/Gly/Ala/Val/Ile/Leu/Tyr/Phe/Lys/His/Arg/Pro/Eaa/Total	40.5	2.3–4.9	0.38/0.16/0.22/0.91/0.19/0.25/0.29/0.23/0.36/0.20/0.23/0.13/0.13/0.38/0.22/2.03/4.6	4.68–9.46
8–11	RM80-RM308	Ser/Cys/Met/Ile/Pro	101.8	2.0–3.6	0.18/0.11/0.14/0.15/0.14	3.48–12.05
10–1	RM474-RM216	Lys	0.0	2.0	0.10	3.7
10–5(6)	RM271-RM258	Ile/His	52.0	2.6–2.9	0.21/0.11	6.53–7.16
	RM258-RM304	Leu/Pro	54.0	2.0–2.1	0.32/0.15	4.57–4.59
		Ser	56.0	3.2	0.21	6.06
		Glu	57.3	2.3	0.73	4.27

See footnotes in Table 2

2002 and six specific to 2004. At least two major QTL clusters in RM472-RM104 (1–19) and RM125-RM542 (7–4,5) were detected consistently in two years. We noticed that the variation explained by these two QTL was large, with values of about 30% in 2002 and about 40% in 2004. The location relationship of the other QTL in the 2 years needs further investigated.

Of all the 18 QTL (clusters), 12 were found to correspond to the loci involved in amino acid metabolism pathways (Table 4; Fig. 1). Some of the QTL were co-mapped with more than one locus in amino acid metabolism pathways. For example, the QTL clusters in RM322-RM521 (2–7) were localized to a region where Asp-AT4 and three members of a glutelin subfamily (GluB-1, GluB-2, and GluB-4) were located; the QTL cluster in RM125-RM542(7–4,5) was co-mapped to a region corresponding to several storage proteins including low-molecular-weight (LMW) globulin and the subfamily of albumin (RA5, RA14, and RAG2); the strong QTL located on the long arm of chromosome 1 (RM315-RM104) was coincident with the gene encoding a putative aspartate aminotransferase that belongs to the aminotransferase classes I and II and the gene encoding a putative Asp kinase, an enzyme that controls the first important step in metabolic pathways for aspartate family amino acid biosynthesis. Altogether, 25 of these loci were found to be involved in amino acid metabolism pathways, with six in amino acid assimilation and transfer pathways, nine in amino acid biosynthesis pathways, and 10 in storage protein pathways (Table 4). Seventeen of these were precisely consistent with the QTL for AAC as indicated by the peak LOD of QTL, and eight of them were located in nearby regions in the same interval. Some genes appeared more than once, such as Asp kinase, GS, and chorismate mutase. We

Table 4 Coincidence of the loci involved in amino acid metabolism with the QTL for 19 components of AAC

Chr interval	Code	Description	Genomic	Location (Mb)	Markers
Amino acid assimilation and transfer					
1–19	AT (classes I and II)	Aspartate aminotransferase (Aminotransferase, classes I and II), putative	AP003235	38.11	RM472, RM431
2–7	Asp AT4	Aspartate aminotransferase, mitochondrial precursor, putative	AP003991	7.71	RM322, RM145
3–14	Root GS2	Glutamine synthetase root isozyme 2, putative	AC082645	28.78	RM55
3–15	GDH2	Glutamate dehydrogenase, putative	AC090871	32.98	RM571, RM468
4–14	Shoot GS2	Glutamine synthetase, chloroplast precursor, putative	AL662953	33.41	RM349, RM349
10–6	Putative GS	Glutamine synthetase, putative	AC025905	16.36	RM184, RM258
Amino acid biosynthesis					
1–19	Putative AK	Aspartate kinase family protein	AP004332	41.04	RM414, RM104
2–5	Chorismate mutase	Chorismate mutase family protein	AP004087	4.49	RM555, RM53
2–9	Tyr/nicotianamine AT	Tyrosine/nicotianamine aminotransferases	AP005532	11.72	RM324, RM301
3–14	AS cII	Anthranilate synthase component II, putative	AC091532	29.02	RM55
4–14	SK	Shikimate kinase, chloroplast precursor, putative	AL606649	32.36	RM349, RM348
7–4	Putative AK	Aspartate kinase family protein	AP006343	11.87	RM542, RM214
7–9	Pho transferase	Phosphoribosylanthranilate transferase, 3′ partial	AP005186	19.00	RN432, RM11
8–3	Try synthase	Tryptophan synthase beta chain 2, chloroplast precursor, putative	AP005620	2.00	RM38
8–10	Chorismate mutase	Chorismate mutase family protein	AP004703	21.39	RM284, RM556
Storage protein					
2–7	GluB-1	Storage protein glutelin B subfamily	AP004018	8.50	RM71, RM145
2–7	GluB-2	Storage protein glutelin B subfamily	AP005511	8.45	RM322, RM145
2–7	GluB-4	Storage protein glutelin B subfamily	AP005428	9.60	RM6911, RM452
3–15	10 KD prolamin	10 KD prolamin	AC099043	31.65	RM293, RM468
3–15	Gb2	Globulin precursor	AC090871	32.90	RM468, RM571
7–4	LMW globulin	Low-molecular-weight globulin	AP004002	6.26	RM125, RM3917
7–4	RA5	Albumin	AP003963	6.33	RM125, RM5673
7–4	RA14	Albumin	AP004002	6.26	RM125, RM5672
7–4	RAG2(RA17)	Albumin	AP004002	6.25	RM125, RM3917
10–4	GluA-2(Gt1)	Glutelin	AC021891	13.15	RM467, RM5689

noticed that these genes were all in key steps of amino acid metabolism pathways.

Discussion

Genetic improvement of nutritional quality is an important objective in crop breeding programs, and is based on better understanding of the genetic basis for nutritional quality characteristics. To our knowledge, this study is the first to partition the AAC in rice grain on a genetic basis. We identified 18 chromosomal regions for 19 components of AAC, with ten specific to 2002 and six specific to 2004. In most loci, the positive alleles came from Zhenshan 97.

Some loci detected here were co-mapped to the loci for PC detected in previous studies. The QTL cluster (flanked by RM315-RM104) on chromosome 1 was consistent with the QTL (flanked by RM315-F16722k) for PC in milled rice reported by Yoshida et al. (2002) and was also near the QTL for PC detected by Hu et al. (2004). The QTL clusters in RM322-RM521 (2–7) and RM556-RM80 (8–10) were also co-mapped with the QTL for PC in brown rice reported by Yoshida et al. (2002). The reason for co-localization might be that the amino acids analyzed here mostly came from the hydrolysate of protein, as free amino acids of brown rice constitute 0.69% of brown rice protein and the free amino acids fraction of milled rice protein is 0.1–0.2% of total protein (Juliano 1985) However, the possibility that some of these QTL also control the content of free amino acid, the sources of protein synthesis, cannot be excluded. Cagampang et al. (1971) reported that the free Lys content was higher than the low-protein lines in an analysis of seven high-protein rice lines, and Wang et al. (2001) found that a higher level of both PC and free AAC were observed in the opaque-2 mutation in maize. Compared with previous studies that focused on PC, the measurement of AAC in rice grains gave our study at least two advantages. First, in this study, some QTL for the contents of Lys and other essential amino acids were mapped and their effects were estimated, which provides direct information for breeders. The QTL with higher additive effects for the contents of Lys or total essential amino acids, such as the QTL on chromosome 1, will be very valuable. The second advantage is that the QTL analyses for the different components of AAC can be regarded as replications. Thus, the precision of QTL location could be improved by integrating the positions of the individual QTL in a QTL cluster.

Our results have significant implications for rice quality improvement programs. By using the QTL for the AAC, the protein quality and content for both Zhenshan 97 and Nanyangzhan via marker-assisted selection could be improved. Since the chemical method using an amino acid analyzer is operationally complex, time consuming, and expensive, the SSR markers closely linked to the main QTL could be a convenient alternative for indirect selection for protein composition. Zhenshan 97 had the highest AAC among all 13 varieties evaluated in our laboratory (data not shown). So, if Zhenshan 97 is consumed only as cooked rice with no special use, attention turns to improvement of essential AAC or stability, as a higher PC may be detrimental to the physicochemical properties of cooked rice, as reported by Martin and Fitzgerald (2002). For this purpose, QTL alleles from Zhenshan 97 itself that increase the contents of Lys and/or other essential amino acids should be maintained in the regions of RM472-RM104 (1–19), RM322-RM521 (2–7), RM301-RM475 (2–10), and RM474-RM216 (10–1). In addition, some favorable QTL alleles such as those in RM555-RM492 (2–5) and RM556-RM80 (8–10) could be introduced from Nanyangzhan.

The final objective of a breeding program is to obtain overall better performance of the variety. Thus, we need to ensure that improving some traits is not accompanied by negative impacts on other important traits. This require informations on the QTL that control all of these traits in the genetic material under study to determine which regions and alleles are to be transferred with markers. Thus, the QTL for AAC could also be used as a background control to improve the grain quality or some important agronomic traits. As we know, Zhenshan 97 is the female parent of a number of widely used hybrids for rice production in China. However, this line is of poor quality because of its high amylose content (AC), hard gel consistency (GC), and high gelatinization temperature (GT), together with a chalky endosperm. Recent results from QTL analyses of AC, GC, and GT have revealed that these traits are mainly controlled by the *Wx* and *Alk* loci, both of which are on the short arm of chromosome 6 (Tan et al. 1999; Lanceras et al. 2000; Septiningsih et al. 2003; Aluko et al. 2004; Tian et al. 2005; Fan et al. 2005; He et al. 2006). Although several studies have reported that the *Wx* locus is associated with PC (Tan et al. 2001; Aluko et al. 2004; Hu et al. 2004), the only QTL for Met content was found at the *Wx* locus, indicating that AAC is relatively independent of the *Wx* and *Alk* loci or their nearby regions in this population. Thus, we can introduce the *Wx* and *Alk* loci from Nanyangzhan, a variety with low GT, long GC, and medium AC, to Zhenshan 97 with no negative effect on nutritional quality.

Our results revealed the co-localization of the QTL for AAC with the loci involved in amino acid metabolism pathways. It was shown that four QTL might be co-localized with the some putative genes

associated with seed storage proteins (Table 4; Fig. 1). Takahashi et al. (2003) found that double null mutations of 11S globulin and 17S globulin caused seed protein deficiency and simultaneously caused high levels of free amino acids to accumulate in soybean seeds. However, we cannot deduce that the AAC was predominantly influenced by the biosynthesis pathways of stored protein. Wang and Larkins (2001) found in maize that a strong QTL for AAC occurs in proximity to genes encoding a monofunctional Asp kinase 2 (Ask2) and further study indicated that it was the best candidate gene (Wang et al. 2001). Evidence from transgenic research has indicated that manipulating the content of a single amino acid influences the whole free AAC in soybean (Simon-sarkadi et al. 2006) and both the free AAC and PC in rice (Wang et al. 2005). Our study showed that the locations of nine QTL correspond to the loci involved in amino acid biosynthesis pathways, such as the loci for putative Asp kinases on chromosome 7 (Table 4 and Fig. 1). We could thus safely assume that amino acid biosynthesis pathways, which provide sufficient amino acids for protein synthesis and N-transport amino acids and serve as a backbone to assimilate and transport nitrogen, also affect the AAC in rice grain.

More interestingly, 10 of the 18 QTL (about 60%) we identified were co-mapped with more than half the QTL for low nitrogen tolerance detected by Lian et al. (2005) who used a population derived from the cross between Zhenshan 97 and Minghui 63 (Fig. 1). Furthermore, five loci correspond to the loci for nitrogen assimilation and transfer, such as *ASP-AT*4 on chromosome 2 and shoot *GS2* on chromosome 4 (Fig. 1). Lohaus et al. (1998) concluded that Illinois high-protein maize differs from Illinois low-protein maize in that it has a high capacity for delivering asparagine as a product of root metabolism. In rice, Obara et al. (2004) found that the QTL on chromosome 2 for the activities of GS1 was co-mapped with the QTL for the soluble PC in senescing leaf blades. Lohaus and Moellers (2000) concluded that the phloem translocation of amino-N and the phloem loading process of amino acids are decisive factors for PC in the seeds of *Brassica* species. A recent study of wheat reported that protein composition was primarily influenced by the nitrogen accumulation rate, and two of the five QTL associated with the kinetics of dry matter and nitrogen accumulation influenced protein composition (Charmet et al. 2005). So, nitrogen assimilation and transfer may also play a major role in controlling AAC in rice grain.

Many studies are needed to clarify the relationship between the loci and physiological processes in nitrogen metabolism and the AAC in rice grain. For example, the environment factors that influence QTL stability remain to be investigated and we are also undertaking a further study on the free AAC in rice grains to distinguish it from the total AAC. Thus, the present work is only a starting point for cloning and characterizing the QTL that underlie the AAC in rice grains. The association of these genes in nitrogen metabolism with the QTL in the present study may be helpful for gene identification through a candidate gene approach.

Acknowledgments This work was supported in part by a grant from the National Program on the Development of Basic Research, the National Program of 863 High Technology Development, and a grant from the National Natural Science Foundation of China.

References

Aluko G, Martinez C, Tohme J, Castano C, Bergman C, Oard JH (2004) QTL mapping of grain quality traits from the interspecific cross *Oryza sativa* × *O. glaberrima*. Theor Appl Genet 109:630–639

Basten CJ, Weir BS, Zeng ZB (1997) QTLCartographer: a reference manual and tutorial for QTL mapping. Department of Statistics, North Carolina State University, Raleigh

Cagampang GB, Cruz LJ, Juliano BO (1971) Free amino acids in the bleeding sap and developing grain of the rice plant. Cereal Chem 48:533–539

Charmet G, Robert N, Branlard G, Linossier L, Martre P, Triboï E (2005) Genetic analysis of dry matter and nitrogen accumulation and protein composition in wheat kernels. Theor Appl Genet 111:540–550

Duan MJ, Sun SSM (2005) Profiling the expression of genes controlling rice grain quality. Plant Mol Biol 59:165–178

Fan CC, Yu XQ, Xing YZ, Xu CG, Luo LJ, Zhang QF (2005) The main effects, epistatic effects and environmental interactions of QTLs on the cooking and eating quality of rice in a doubled-haploid line population. Theor Appl Genet 110:1445–1452

FAO (Food, Agriculture Organization of The United Nations) (1970) Amino acid content of foods and biological proteins. FAO, Roma

Gupta MP, Gupta PK, Singh IB, Singh P (1988) Genetic analysis for quality characters in rice. Genetika 20: 141–146

He Y, Han YP, Jiang L, Xu CW, Lu JF, Xu ML (2006) Functional analysis of starch-synthesis genes in

determining rice eating and cooking qualities. Mol Breed 18:277–290

Hu ZL, Li P, Zhou MQ, Zhang ZH, Wang LX, Zhu LH, Zhu LG (2004) Mapping of quantitative trait loci (QTLs) for rice protein and fat content using doubled haploid lines. Euphytica 135:47–54

Jiang GH, He YQ, Xu CG, Li XH, Zhang Q (2004) The genetic basis of stay-green in rice analyzed in a population of doubled haploid lines from an *indica* by *japonica* cross. Theor Appl Genet 108:688–698

Juliano BO (1985) Rice chemistry and technology, 2nd edn. American Association of Cereal Chemists, Incorporated Saint Paul, Minnesota

Kambayshi M, Tsurumi I, Sasahara T (1984) Genetic studies on improvement of protein content in rice grain. Jpn J Breed 34:356–363

Lanceras JC, Huang ZL, Naivikul O, Vanavichit A, Ruanjaichon V, Tragoonrung S (2000) Mapping of genes for cooking and eating qualities in Thai jasmine rice (KDML105). DNA Res 7:93–101

Lian XM, Xing YZ, Yan H, Xu CG, Li XH, Zhang QF (2005) QTLs for low nitrogen tolerance at seedling stage identified using a recombinant inbred line population derived from an elite rice hybrid. Theor Appl Genet 112:85–96

Lin R, Luo Y, Liu D, Huang C (1993) Determination and analysis on principal qualitative characters of rice germplasm. In: Ying C (ed) Rice germplasm resources in China. Agricultural Science and Technology Publisher of China, Beijing, pp 83–93

Lincoln S, Daly M, Lander E (1992) Constructing genetics maps with MAPMAKER/EXP 3.0. Whitehead Institute Technical Report, Whitehead Institute, Cambridge, Massachusetts

Lohaus G, Moellers C (2000) Phloem transport of amino acids in two *Brassica napus* L. genotypes and one *B. carinata* genotype in relation to their seed protein content. Planta 211:833–840

Lohaus G, Büker M, Hußmann M, Soave C, Heldt HW (1998) Transport of amino acids with special emphasis on the synthesis and transport of asparagine in the Illinois low protein and Illinois high protein strains of maize. Planta 205:181–188

Martin M, Fitzgerald MA (2002) Proteins in rice grain influence cooking properties. J Cereal Sci 36:285–294

McCouch SR, Teytelman L, Xu Y, Lobos KB, Clare K, Walton M, Fu B, Maghirang R, Li Z, Xing Y, Zhang Q, Kono I, Yano M, Fjellstrom R, DeClerck G, Schneider D, Cartinhour S, Ware D, Stein L (2002) Development and mapping of 2,240 new SSR markers for rice (*Oryza sativa* L.). DNA Res 9:199–207

Obara M, Sato T, Sasaki S, Kashiba K, Nagano A, Nakamura I, Ebitani T, Yano M, Yamaya T (2004) Identification and characterization of a QTL on chromosome 2 for cytosolic glutamine synthetase content and panicle number in rice. Theor Appl Genet 110:1–11

Septiningsih EM, Trijatmiko KR, Moeljopawiro S, McCouch SR (2003) Identification of quantitative trait loci for grain quality in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. Theor Appl Genet 107:1433–1441

Shenoy VV, Seshu DV, Sachan JKS (1991) Inheritance of protein per grain in rice. Indian J Genet 51:214–220

Simon-sarkadi L, Kocsy G, Várhegyi Á, Galiba G, De ronde JA (2006) Stress-induced changes in the free amino acid composition in transgenic soybean plants having increased proline content. Biol Plant 50:793–796

Sood BC, Siddiq EA (1986) Genetic analysis of crude protein content in rice. Indian J Agric Sci 56:796–797

StatSoft (1991) Statistica. StatSoft, Tulsa, Oklahoma

Takahashi M, Uematsu Y, Kashiwaba K, Yagasaki K, Hajika M, Matsunaga R, Komatsu K, Ishimoto M (2003) Accumulation of high levels of free amino acids in soybean seeds through integration of mutations conferring seed protein deficiency. Planta 217:577–586

Tan YF, Li JX, Yu SB, Xing YZ, Xu CG, Zhang QF (1999) The three important traits for cooking and eating quality of rice grains are controlled by a single locus in an elite rice hybrid, Shanyou 63. Theor Appl Genet 99:642–648

Tan YF, Sun M, Xing YZ, Hua JP, Sun XL, Zhang QF, Corke H (2001) Mapping quantitative trait loci for milling quality, protein content and color characteristics of rice using a recombinant inbred line population derived from an elite rice hybrid. Theor Appl Genet 103:1037–1045

Temnykh S, Park WD, Ayres N, Cartihour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR (2000) Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:697–712

Temnykh S, Declerck G, Luashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11:1441–1452

Tian R, Jiang GH, Shen LH, Wang LQ, He YQ (2005) Mapping quantitative trait loci underlying the cooking and eating quality of rice using a DH population. Mol Breed 15:117–124

Wang XL, Larkins BA (2001) Genetic analysis of amino acid accumulation in opaque-2 maize endosperm. Plant Physiol 125:1766–1777

Wang XL, Stumpf DK, Larkins BA (2001) Aspartate kinase 2. A candidate gene of a quantitative trait locus influencing free amino acid content in maize endosperm. Plant Physiol 125:1778–1787

Wang WM, Zhao Q, Yu JJ, Zhu DY, Ao GM (2005) Transfer of high lysine gene sb401 into rice and analysis for protein and amino acid content in transgenic rice seeds. Acta Agronomica Sinica 31:603–607

WHO (1973) Energy and protein requirements. WHO Tech. Rep. Ser. 522. World Health Organization, Geneva

Yoshida S, Ikegami M, Kuze J, Sawada K, Hashimoto Z, Ishii T, Nakamura C, Kamijima O (2002) QTL analysis for plant and grain characters of sake-brewing rice using a doubled haploid population. Breed Sci 52:309–317

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136:1457–1468