Carbohydrate Polymers 81 (2010) 120–127

Contents lists available at ScienceDirect

Carbohydrate Polymers

journal homepage: www.elsevier.com/locate/carbpol

Using chain-length distributions to diagnose genetic diversity in starch biosynthesis

Rosa Paula Cuevas [a,b], Venea Dara Daygon [a], Matthew K. Morell [c], Robert G. Gilbert [b], Melissa A. Fitzgerald [a,*]

[a] International Rice Research Institute (IRRI), DAPO 7777, Metro Manila, Philippines
[b] University of Queensland, Centre for Nutrition and Food Sciences, School of Land, Crop & Food Sciences, Brisbane, Qld 4072, Australia
[c] CSIRO Food Futures Flagship, P.O. Box 1600, Canberra, ACT 2601, Australia

ARTICLE INFO

Article history:
Received 16 December 2009
Received in revised form 29 January 2010
Accepted 2 February 2010
Available online 6 February 2010

Keywords:
Starch
Gelatinisation temperature
Chain-length distribution
ln*N* plot

ABSTRACT

Amylopectin synthesis is controlled by the coordinated action of several types of enzymes, including starch synthases, branching and debranching enzymes. The contributions of some of these enzymes to the building of starch molecules have been previously established. Changes to the activity of an enzyme can affect amylopectin structure, which is associated with diversity in functional properties. One such property, gelatinisation temperature, has been studied at the genetic, biochemical and phenotypic level. A technique, the '*log(number distribution)* approach', offers a means of collecting normalisation-free representations of the chain-length distributions of starch, with the potential to reveal new information about kinetics and processes of starch synthesis; this method of plotting the data can potentially reveal much more than that usually employed, viz., the simple number chain-length distribution. In this study, samples from genotypes with defined mutations in starch biosynthetic genes that specifically and differently alter the chain-length distribution of single-cluster chains, with a resultant effect on gelatinisation temperature, are used to show that *log(number distribution)* plots have sufficient discriminative capacity to diagnose the gene affected by mutations, provide new information, and determine those features of the plot which associate with genotype and phenotype.

1. Introduction

Molecules of amylopectin constitute the largest component of starch, and their structure contributes to many of the functional properties of the starch. Amylopectin molecules consist of linear chains of glucose linked by α1-4 bonds, and the chains are arranged in clusters with α1-6 linkages forming the branch points (Ball & Morell, 2003). The structure of whole molecules of amylopectin is highly complex and difficult both to specify and to interpret, mainly because of its branched structure. However, the debranched chain-length distribution (CLD) of the amylopectin can be obtained with relative ease by debranching with isoamylase, which specifically hydrolyses the α1-6 linkages, and separating with capillary electrophoresis (O'Shea et al., 1998). Information can be obtained from the CLD about the regulatory and synthetic processes leading to the various chain lengths of amylopectin (Ball & Morell, 2003; Castro, Dumas, Chiou, Fitzgerald, & Gilbert, 2005; Fujita et al., 2006, 2007; Konik-Rose et al., 2007; Morell et al., 2003; Ryoo et al., 2007).

The CLD of debranched starch can be plotted as the logarithm of the number distribution, $N(X)$, as a function of degree of polymerisation X (Castro et al., 2005) (to conform to IUPAC recommendations (Cohen et al., 2007; Jones et al., 2009) the notation N is used to refer to a generic number distribution, and X to refer to the degree of polymerisation (DP); in some earlier work (Castro et al., 2005), the symbol P has been used for number distribution and the symbol N for DP). The data can be plotted in this way because the simplest hypothesis for the formation of a given linear chain is that the only processes involved are random (chain length-independent) elongation and termination, and the distribution for such a process is a single exponential X. Moreover, this method of plotting the data is essentially assumption-free. Regions where this simple hypothesis is valid will show a linear $\ln N(X)$ (with a negative slope) (Eq. (1)):

$$N(X) = \exp\left(-\frac{\text{termination rate}}{\text{elongation rate}}X\right) \quad (1)$$

where the slope is the ratio of the rates of chain termination and elongation (Castro et al., 2005). Chain elongation is carried out by the family of starch synthases (SS) and chain termination occurs

Abbreviations: FACE, fluorophore-assisted capillary electrophoresis; DP, degree of polymerisation; SS, starch synthase; BE, branching enzyme; GBSS, granule-bound starch synthase.

* Corresponding author. Tel.: +63 2 580 5600; fax: +63 2 580 5699.
E-mail address: m.fitzgerald@cgiar.org (M.A. Fitzgerald).

0144-8617/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carbpol.2010.02.004

via a number of mechanisms. It can occur by cleavage of a portion of the chain by a branching enzyme (BE), by removal of a chain though debranching, or when affinity of the enzyme for the elongated chain rises to a point where turnover rates are zero. Regions where linear behaviour is seen can be reasonably ascribed to chain length-independent activity, whereby the collective processes of elongation and termination over that particular region, which may or may not be individually random, are independent of chain length (Castro et al., 2005).

Four regions of the ln*N* plots of amylopectin have been defined: Regions 1 and 3, where chain length-dependent behaviour is seen, and Regions 2 and 4, where the synthesis of a particular chain is independent of its length. Chains in Regions 1 and 2 fit into one lamella, while those in Regions 3 and 4 extend into subsequent lamellae (Castro et al., 2005). Any alteration to the kinetics of the synthesis of any group of chains should alter the slope of some region of the ln*N* plot (Castro et al., 2005); potentially, this provides an *in vivo* tool for extracting mechanistic information about starch synthesis, and a basis to generate biochemical and genetic hypotheses.

Our present knowledge of starch synthesis is incomplete, as is the potential capacity of each activity to alter the slope of the ln*N* plot. However, much is known in rice about the genetics and biochemistry of SSIIa (Bao, Corke, & Sun, 2006; Cao et al., 1999; Nakamura et al., 2005; Umemoto et al., 2004; Waters, Henry, Reinke, & Fitzgerald, 2006) and BEIIb (Jiang, Dian, & Wu, 2003; Nishi, Nakamura, Tanaka, & Satoh, 2001; Tanaka et al., 2004), and the effect of these enzymes on the CLD of the starch and the functional property of gelatinisation temperature. In rice, single nucleotide polymorphisms (SNPs) in exon 8 of the *SSIIa* gene classify rice into four haplotypes, two of which have fully functional forms of the SSIIa enzyme (haplotype 1, G/G/GC, and haplotype 2, A/G/GC) while the other two (haplotype 3, A/A/GC, and haplotype 4, A/G/TT) are associated with a decrease or loss of enzyme activity (Cuevas et al., 2010; Nakamura et al., 2005; Umemoto & Aoki, 2005). Loss of SSIIa activity leads to a decrease in gelatinisation temperature of rice (Cuevas et al., 2010; Nakamura et al., 2005; Umemoto & Aoki, 2005) and wheat (Konik-Rose et al., 2007) because of a decrease in the proportion of chains DP 12-24. By contrast, loss of BEIIb activity in rice leads to an increase in GT because of an increase in the proportion of long chains within an amylopectin cluster (Nishi et al., 2001; Tanaka et al., 2004). Such well characterised mutants can be used to test the utility of the ln*N* plot as a diagnostic tool for detecting genes affected by mutations.

An extension of the simple idea described above shows that Eq. (1) is in fact an approximation to a fuller treatment (Gray-Weale & Gilbert, 2009). If it is assumed, for these purposes, that the distributions of branches in a given region are controlled by only a single starch synthase, a single branching and a single debranching enzyme, and the branching enzyme randomly cleaves off a chain of any DP and joins it to any other chain to produce a cleaved and a new chain, both of which can undergo further propagation. Solution of the resulting evolution equations for starch synthesis then yields:

$$N(X) = X \exp\left(-\frac{\text{branching rate}}{\text{elongation rate}} X\right) \quad (2)$$

Another extension of this (Hasjim, Wu, Godwin, Gidley & Gilbert, unpublished results) is to take account of the fact that branching enzymes preferentially cleave off chains longer than a certain length denoted $X_{\min}$ (typically $X_{\min} \sim 8$). Eq. (2) then becomes:

$$N(X) = (X - X_{\min}) \exp\left(-\frac{\text{branching rate}}{\text{elongation rate}} (X - X_{\min})\right) \quad (3)$$

The derivation of this functional form means that it is only applicable to Regions 1 and 2 (extensions to take into account chains extending over more than one lamella will be given elsewhere (Hasjim, Gidley & Gilbert, unpublished results)). This derivation is only applicable to branching, and not to debranching or kinetic saturation of starch synthase activity.

While Eq. (3) is a mechanistically based means of data-fitting, the simple ln*N* plot suggested by Eq. (1) is always a good means of plotting the CLD. First, a simple change of ordinate enables data to be compared independent of normalisation, which otherwise introduces artefacts (Castro et al., 2005). Second, because the CLD typically varies by 3 orders of magnitude, it brings out features arising from significant variation in the distribution at higher DPs, which might not be apparent in a number (CLD) plot.

In the present paper, our objective is to test the applicability of Eq. (1), to data which obey Eq. (3), and then to test the diagnostic capacity and sensitivity of the ln*N* plot by using samples that differ in the ratio of chains DP 6-10 and DP 11-24. The first set of samples is a diverse collection of rices with allelic variability for *SSIIa* (Cuevas et al., 2010); the second is a doubled haploid population of mutants of wheat with progressive deletions of *SSIIa* (Konik-Rose et al., 2007) representing all the possible genotypes for *SSIIa*; and the third set consists of an amylose-extender mutant with functional SSIIa, and a sample that behaves like an amylose-extender and has reduced SSIIa functionality. Both enzymes are associated with gelatinisation temperature (Tanaka et al., 2004; Umemoto, Yano, Satoh, Shomura, & Nakamura, 2002).

2. Results

Fig. 1 compares the rate ratios obtained from the fits of Regions 1 and 2 to Eq. (3) for a wide variety of grains (six rice varieties, and several varieties of each of wheat, maize and barley (Castro et al., 2005)), with the slopes of Region 2 determined using Eq. (1). Results from the two methods correlate well (Fig. 1), so the simpler procedure of using the slope, rather than the more precise use of Eq. (3), was used for all further analyses.

Fig. 2a shows the CLDs of two varieties of rice obtained from fluorophore-assisted capillary electrophoresis (FACE), plotted as ln*N*(*X*). The plots can be divided into four regions, as previously reported (Castro et al., 2005). The difference in the ln*N*(*X*) plots of these varieties is the presence of an interruption to linearity at DP 18-22 in one and not the other. This shoulder is not observed in the slope of Region 4 of either variety.

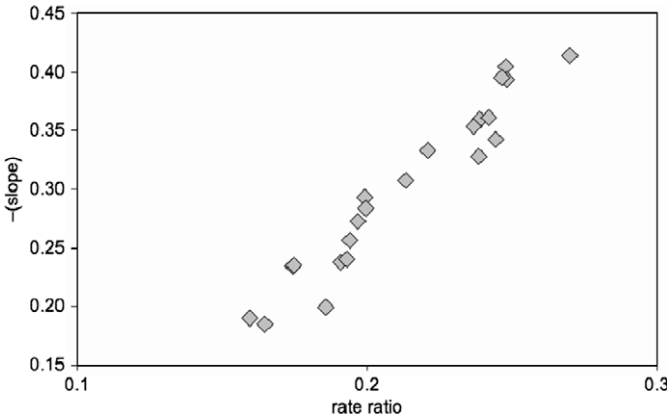

Fig. 1. Comparison of rate ratios obtained from the fits to Eq. (3) of ln*N* data for Regions 1 and 2, with the slopes of Region 2, using Eq. (1) for a wide number of species of various grains: barley, rice (ACCN 105, ACCN 186, ACCN 1799, ACCN 14487, IR 72 and Koshihikari), wheat and maize.

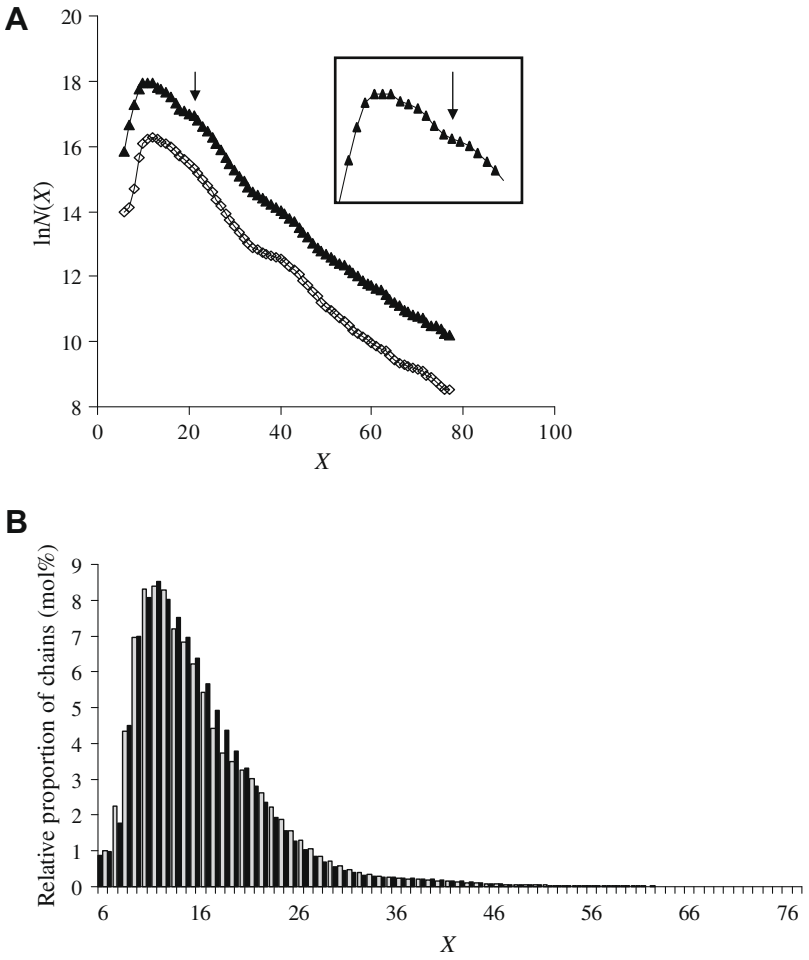

Fig. 2. (A) Chain-length distributions of two varieties plotted as ln$N(X)$: Khao Dawk Mali 105 (▲), which has the interruption to linearity at DP 18-22 (arrow and inset), and IR 64 (□) which does not have this feature. The plots are offset by 1.5. (B) Classical presentation of chain-length distributions for these two varieties. Khao Dawk Mali 105 is represented by grey columns, IR 64 by black columns.

Examination of the ln$N(X)$ plots obtained for a set of 93 rice varieties shows that the slope of Region 2 from DP 13-17 is significantly more negative in varieties with the shoulder at DP 18-22 than for varieties without, but no significant change was found between the varieties for the slope once linearity was resumed over the range of chain lengths DP 24-35 (Table 1). In this set of samples, the average gelatinisation temperature of varieties with the DP 18-22 shoulder (35.5%) clearly present was 66.7 ± 2.4 °C, while the average gelatinisation temperature of varieties showing a continuous Region 2 (64.5%) was 74.1 ± 2.3 °C (Table 2). Those varieties showing the DP 18-22 shoulder all carried haplotypes 3 or 4 of *SSIIa*, and those with uninterrupted slopes all carried haplotypes 1 or 2 of *SSIIa* (Table 2, Supplementary Table 1). Presenting the data as chain-length distributions (Figs. 2b and 3b) shows the shoulder at DP 18 in both rice and wheat samples, but it is more easily seen when the data is presented as ln$N(X)$ (Figs. 2a and 3b).

Table 1
Analysis of variance (ANOVA) on the mean values of the slopes of the ln$N$ plot at DP 13-17, and DP 24-35 for rice varieties with and without the interruption at DP 18-22 (n = 93).

Interruption at DP 18-22	Mean slope values of the ln$N$ plot[A]	
	DP 13-17	DP 24-35
Absent	−0.101[a]	−0.186[a]
Present	−0.119[b]	−0.194[a]
$LSD_{0.05}$	0.005	0.015

[A] In each column, means followed by a common letter are not significantly different at the 5% level.

Table 2
Average gelatinisation temperature of varieties with and without the shoulder at DP 18-22 using 93 diverse rice varieties.

Presence of the shoulder	Gelatinisation temperature (°C)	Proportion of the population (%)
(+) Shoulder	66.7 ± 2.4	35.5
(−) Shoulder	74.1 ± 2.3	64.5

Binary logistic regression analysis of the relationship between the slope of DP 13-17 in rice and observation of the DP 18-22 shoulder showed that, for every 0.01 increase in the slope of Region 2 above the shoulder at DP 18-22, the probability of the interruption appearing decreases by 71% (Table 3). The ln$N(X)$ plots of the wheat samples showed that linearity in Region 2 was interrupted between DP 18-22 for all genotypes (Fig. 3a). A decrease in dosage of *SSIIa* was accompanied by increasingly negative slopes of ln$N(X)$ at DP 13-17 (Table 4).

Analysis of the relative CLDs of two rice varieties, IR36 and Ilpumbyeo, carrying wildtype *BEIIb* shows that the proportion of chains DP 13-25 is greater for IR36 than Ilpumbyeo, while the proportion of chains DP 6-11 is greater for Ilpumbyeo than for IR36

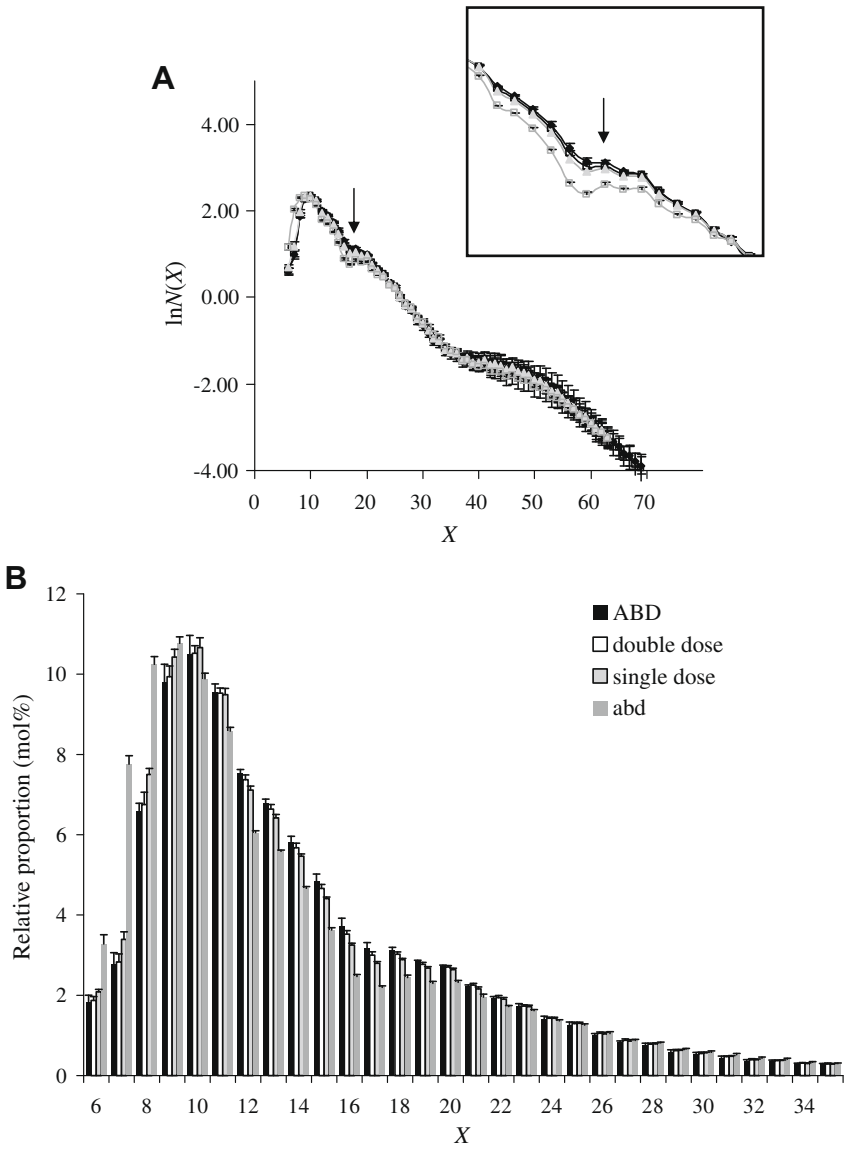

Fig. 3. (A) Chain-length distributions of different *SSIIa* genotypes of wheat plotted as ln$N(X)$ showing that linearity in Region 2 is interrupted: wildtype ABD (♦), single null (□), double null (▲), and triple null, abd (□). The inset shows that the interruption is more severe as dosage of *SSIIa* decreases. (B) Classical chain-length distributions of wheat representing these different *SSIIa* genotypes.

Table 3
Binary logistic regression coefficient statistics on the slope of DP 13-17 and on the shoulder at DP 18-22 in 93 varieties of rice, from CropStat. The results show that the probability of the bump appearing decreases by 71% for every 0.01 increase in the value of slope of DP 13-17.

	Coefficient statistics				
	Coefficient	Standard error	Asymptotic z-statistic	Asymptotic p-value	$1-\exp(s_{DP13\text{-}17})$
Intercept	−14.09	2.96	−4.75	0.00	
Slope DP 13-17	−1.23	0.27	−4.63	0.00	0.71

(Fig. 4a). IR36 carries haplotype 1 of *SSIIa* and Ilpumbyeo carries haplotype 4 of *SSIIa* (Supplementary Table 1). However, comparison of the CLDs of amylopectin of their respective *BEIIb* mutants, IR36*ae* and Goami 2, showed no significant differences at any chain length (Fig. 4b). When the CLDs were replotted as ln$N(X)$ plots (Fig. 5), it was observed that Ilpumbyeo, Goami 2, and IR36*ae* all showed the shoulder at DP 18-22, and IR36 did not.

The mean values of the slopes of the different regions of the ln$N(X)$ plots for the mutants and wildtypes are shown in Table 5. The ln$N(X)$ plots of IR36*ae* and Goami 2 were similar for all features (Fig. 5a and b), and the slope at DP 13-17 was significantly more positive than either wildtype. In addition, wildtypes Ilpumbyeo and IR36 had significantly higher proportions of chains DP $\leqslant$ 12 than their mutants (Fig. 5c and d). There were also differences in proportion of chains DP 13-24 between the wildtype and the mutant.

3. Discussion

3.1. Fitting of the data

As shown above, a simple straight-line fitting of Region 2, as in Eq. (1), rather than fitting to the more complex form of Eq. (3),

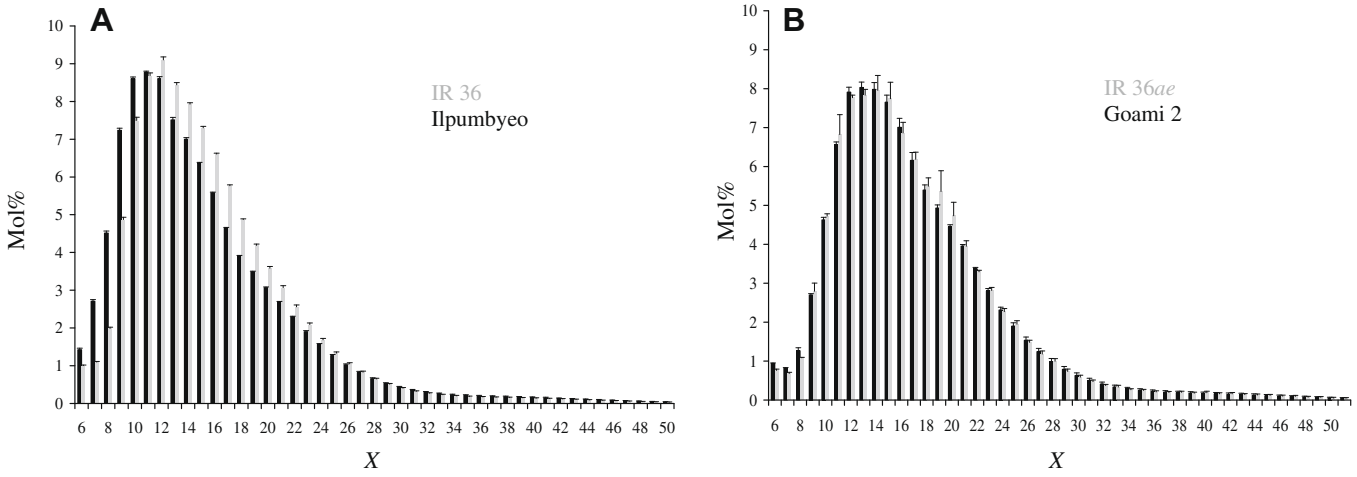

Fig. 4. Comparison of chain-length distributions (in mol%) between (a) wildtypes IR36 and Ilpumbyeo; and (b) mutants IR36*ae* and Goami 2.

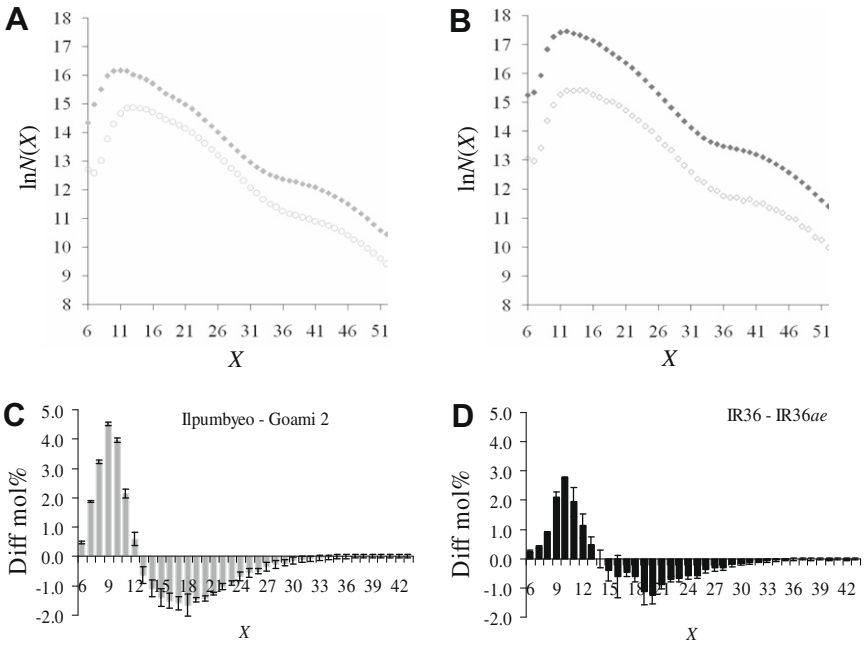

Fig. 5. Comparison of chain-length distributions presented as ln*N(X)* plots of (A) Ilpumbyeo (filled) and Goami 2 (unfilled), and (B) IR36 (filled) and IR36ae (unfilled), and presented as difference plots of relative amounts per chain length between (C) Ilpumbyeo and Goami 2 and (D) IR36 and IR36*ae*.

gives a value of the slope which has the same trends as the rate ratio of the latter equation. This is because, except close to the maximum, the more complex form has only a slight curvature when plotted logarithmically, and the average slope obtained by using Eq. (1) will then vary monotonically with the rate ratio of the more complex form. Thus use of Eq. (1) is adequate for the semi-quantitative needs of the present paper.

3.2. Information from lnN plots

In the present study, the previously reported four-region behaviour (Castro et al., 2005) was seen for a wide range of varieties (exemplified in Fig. 3a). A shoulder in the ln*N*(*X*) plot was observed for some varieties at DP 18-22, in Region 2 (Fig. 2a). This can be distinguished as a deviation in the decreasing linear function, to a greater or lesser extent, on close examination of previously published data plotted as ln*N*(*X*) (Castro et al., 2005). Further, a shoulder at DP 18 was noted previously in the chain-length distribution of a number of starches from different botanical sources (Hanashiro, Abe, & Hizukuri, 1996; Jane et al., 1999), which presumably would correspond to the DP 18-22 shoulder if those data were replotted as ln*N*.

It would seem that the shoulder in Region 2 arises from two separate types of processes, each of which has mechanisms similar to those originally proposed for Region 2. SSIIa and BEIIb are both clearly involved in the first component, since the absence of either one of them alters the slope of Regions 1 and 2 as far as the shoulder. The slope following the DP 18-22 shoulder in Region 2 did not differ between a mutant and its respective wildtype (Tables 4 and 5), so we conclude that neither SSIIa nor BEIIb are essential for the synthesis of chains to DP 24- 35 in the mutants; if they are involved in the wildtypes, other SSs and BEs must be able to carry out required activity in the mutants.

3.3. SSIIa and the interruption at DP 18-22

It is well established that mutations in *SSIIa* alter the chain-length distributions of amylopectin in different rice varieties

Table 4
Analysis of variance (ANOVA) on the mean values of the slopes of the ln*N* plots at DP 6-10, DP 13-17, and DP 24-35 in 40 wheat lines with different genotypes for *SSIIa*. Wildtype (*ABD*), single null (*AbD*, *ABd*, *aBD*), double null (*Abd*, *aBd*, *abD*) and triple null (*abd*).

Genotype[A]	Mean values of the slopes from the ln*N* plot (%)[A]		
	DP6-10	DP 13-17	DP 24-35
Wildtype	0.48[b]	−0.20[c]	−0.16[a]
Single null	0.48[b]	−0.21[b,c]	−0.15[a]
Double null	0.44[b]	−0.22[b]	−0.15[a]
Triple null	0.25[a]	−0.25[a]	−0.14[a]
$LSD_{0.05}$	0.02	0.01	0.01

[A] In each column, means followed by a common letter are not significantly different at the 5% significance level.

Table 5
Analysis of variance (ANOVA) on the mean values of the slopes of DP 6-10, DP 13-17, and DP 24-35 of the ln*N* plot of IR36, IR36*ae*, Ilpumbyeo, and Goami 2.

Varieties	Mean values of the slopes from the ln*N* plot[A]			Gelatinisation temperature (°C)
	DP 6-10	DP 13-17	DP 24-35	
IR36	0.55[a]	−0.09[b]	−0.21[a]	74
IR36*ae*	0.51[ab]	−0.06[a]	−0.21[a]	77
Ilpumbyeo	0.46[bc]	−0.12[c]	−0.19[a]	64
Goami 2	0.44[c]	−0.07[a]	−0.21[a]	77
$LSD_{0.05}$	0.02	0.01	0.02	

[A] Average of three replications; in each column, means followed by a common letter (a,b,c) are not significantly different at the 5% level.

(Nakamura et al., 2005; Umemoto & Aoki, 2005; Umemoto et al., 2002, 2004). These mutations increase the proportion of chains DP 6-10 and decrease the proportion of chains DP 12-24, which lead to a decrease in gelatinisation temperature (Konik-Rose et al., 2007; Nakamura et al., 2005; Umemoto & Aoki, 2005). Haplotypes 3 and 4 of *SSIIa* are defined by SNPs in the binding and catalytic region of the gene (Waters et al., 2006), which could explain the decrease of SSIIa activity. SSIIa activity is an elongation event in starch synthesis. The interruption to linearity, observed between DP 18-22, correlates with gelatinisation temperature and with the decrease in SSIIa activity (Table 2). Table 2 also shows that each variety with the DP 18-22 shoulder carries a low-activity haplotype of *SSIIa* and varieties without the DP 18-22 shoulder carry the active haplotypes, as previously defined (Nakamura et al., 2005). This indicates that the DP 18-22 shoulder arises from the decrease of SSIIa activity and that the ln*N* technique is capable of detecting that loss.

In the doubled haploid wheat population (Konik-Rose et al., 2007), an obvious interruption to Region 2 for every genotype (Fig. 3a) was noted when the data were plotted as ln*N*. The decreasing dosage of *SSIIa* in these lines led to increasingly negative slopes in the ln*N* plot of chains DP 13-17 compared with the wildtype (Table 4), in the same way as was seen in rice. The changes in the slope at DP 13-17 indicate changes in the kinetics of amylopectin chain synthesis as affected by changes in SSIIa functionality. However, the absence of SSIIa activity in wheat and in rice does not lead to a complete loss of chains DP 12-24, indicating that other starch synthases either contribute to the elongation of those chains, or are able to compensate when SSIIa activity is lost. Furthermore a shoulder at DP 18-22 was also seen in the wheat lines carrying full SSIIa activity (Fig. 3a), which is in contrast to rice that carries an active *SSIIa* haplotype. This suggests that there are significant differences between processes or regulation of those processes in wheat and rice for the synthesis of chains DP 13-17.

3.4. BEIIb and the interruption at DP 18-22

Another enzyme that affects gelatinisation temperature in rice by altering the relative ratios of single-cluster chains is BEIIb (Jiang et al., 2003; Nishi et al., 2001; Tanaka et al., 2004; Yamakawa, Hirose, Kuroda, & Yamaguchi, 2007). BEIIb cleaves a portion of an elongating chain to create a new short chain via an α1-6 linkage (Nakamura, 2002). The new chain and the remaining stub are preferentially elongated by SSI (Fujita et al., 2006) and then by SSIIa (Umemoto et al., 2004). BEIIb activity is therefore a termination event in the process of the synthesis of chains DP 11-24.

IR36*ae*, a mutant of IR36, carries a mutation in *BEIIb* (Asaoka et al., 1986; Juliano, Perez, Kaushik, & Khush, 1990; Nishi et al., 2001; Vandeputte & Delcour, 2004; Yano, Okuno, Kawakami, Satoh, & Omura, 1985), and Goami 2 is considered to be a *BEIIb* mutant of Ilpumbyeo (Kang, Hwang, Kim, & Choi, 2003). IR36 carries a functional haplotype of *SSIIa* and Ilpumbyeo does not (Supplementary Table 1). In the BEIIb wildtypes, there are proportionately more chains of DP 6-12 and fewer of DP 12-24 in Ilpumbyeo compared with IR36 (Fig. 4a), which is consistent with carrying a non-functional haplotype of *SSIIa*. The similarity of the chain-length distribution of Goami 2 and IR36*ae* (Fig. 4b) suggests that a decrease in BEIIb activity over-rides any effect of the difference in *SSIIa* status of the two genotypes. In wheat and maize, enzymes involved in starch biosynthesis are reported to form complexes in the amyloplast, with altered functionality in any one of the components leading to pleiotropic effects (Hennen-Bierwagen et al., 2008; Tetlow, Morell, & Emes, 2004; Tetlow et al., 2008). The effect of BEIIb activity on SSIIa activity in Goami and IR36*ae* chain-length distributions (Table 5), as compared to those of their respective wildtypes, suggests that these two enzymes might also be in complex in rice.

The ln*N* of Region 2 in IR 36 is smooth, but a shoulder, or disruption of linearity, at DP 18-22 is observed in IR36*ae* (Fig. 5b). This is also observed in Ilpumbyeo, which has altered elongation mechanisms due to the decreased activity of SSIIa, and in Goami 2, which has non-functional alleles of both *SSIIa* and *BEIIb* (Fig. 5a). The presence of the DP 18-22 shoulder in IR36*ae* (Fig. 5b) suggests that BEIIb is part of the process that synthesises chains DP 13-17. However, the similarity of the slopes for chains between IR36 and IR36*ae* and between Ilpumbyeo and Goami 2 (Table 5) suggest either that BEIIb is not directly involved in the synthesis of chains of DP 24-35, or that in its absence other BEs (BEIIa and/or BEI) compensate for any loss of BEIIb activity in this region of the profile. Further, the significantly higher proportion of short chains in Ilpumbyeo and IR36 as compared to their mutants (Fig. 5c and d) indicate a decrease in the synthesis of chains DP $\leqslant$ 12 in the mutants, consistent with the effects of losing BEIIb activity; this is particularly noticeable in the temperate japonica pair.

Mutations in *SSIIa* increase the negativity of the slope at DP 13-17, but inactivation of *BEIIb* is associated with an increase in slope (the value becomes more positive) of DP 13-17 (Table 5). This difference is consistent with the gene affected by mutation. SSIIa activity is an elongation process and BEIIb activity is a termination process. The direction of change seen in the slopes is consistent with Eq. (1), and it indicates that the ln*N* plot is sufficiently sensitive to discriminate between interruptions to elongation and termination events, at least in the processes that are involved in creating single-cluster chains.

4. Experimental procedures

4.1. Plant materials

A set of rice varieties (*Oryza sativa* L.) was selected from the collection at the International Rice Research Institute (IRRI) based on

known diversity for traits of grain quality. The set (93 accessions) was obtained from the TT Chang Genetic Resources Centre (GRC) at IRRI, Philippines. A table of the list of varieties is available in the Supplementary Material. Samples of milled rice of Ilpumbyeo and Goami 2 were a gift from Dr Park of the Korean Rural Development Administration and IR36 and its mutant IR36*ae* were obtained from the collection at IRRI. A doubled haploid wheat population was established, genotyped, and phenotyped previously (Konik-Rose et al., 2007).

For rice obtained as paddy, grains were dehulled (Satake Rice Machine, Tokyo, Japan), milled (Grainman 60-230-60-2AT, Grain Machinery Mfg. Corp., Miami, FL), and ground (Udy Cyclone Sample Mill 3010-030, Fort Collins, CO) to pass through a 0.5 mm sieve. Reagent-grade chemicals were used. Reverse osmosis water, filtered through a 0.22 μm Millipore (Billerica, MA) filter, was used throughout the study.

4.2. Gelatinisation temperature

Gelatinisation temperature was measured by differential scanning calorimetry (DSC) (model Q100 TA Instruments, New Castle, DE). Flour (4 mg) was mixed with water (8 μL) in an aluminium hermetic pan which was then hermetically sealed. The temperature was raised from 35 to 140 °C at 4 °C/min, with modulation of ±0.5 °C every 40 s. Thermal transitions were recorded and analysed using Universal Analysis 2000 software. The gelatinisation temperature is reported as the peak of the gelatinisation endotherm.

4.3. Chain-length distributions

Starch was gelatinised and debranched according to a previously published method (Lisle, Martin, & Fitzgerald, 2000). Debranched chains were labelled following an established method (O'Shea et al., 1998) with one modification: samples were labelled at 50 °C for 24 h, which resulted in more labelled chains than when labelling was carried out at 37 °C (data not shown). After the labelled samples were filtered (Freeze and Squeeze spin columns, Bio-Rad), chain-length distributions were obtained using a P/ACE MDQ capillary electrophoresis unit (Beckman-Coulter, Fullerton) with the cathode at the injection side (reverse polarity). A laser-induced fluorescence detector was used with an argon laser as the excitation source. The N-CHO coated capillary (50 μm I.D., 50.2 cm total length) and the carbohydrate separation gel buffer were obtained from Beckman-Coulter. Separations were conducted at 23.5 kV at 25 °C (O'Shea et al., 1998) for 80 min. Separation conditions were controlled, and data was recorded and analysed using 32-Karat 7.0 software.

The degree of polymerisation, X, was assigned to peaks based on the migration time of maltose. The area of each peak was converted to velocity-area (Demorest & Dubrow, 1991), to obtain the number distribution of chain lengths, $N(X)$. Note that the normalisation of this distribution is arbitrary. The chain-length distribution is presented as $\ln N(X)$; metrics of the distribution were calculated as previously reported (Castro et al., 2005; Chiou, Fellows, Gilbert, & Fitzgerald, 2005).

4.4. Genotyping for SSIIa

DNA was extracted from the leaf samples using an established protocol (Fulton, Chunwongse, & Tanksley, 1995). The polymerase chain reaction (PCR) was performed using a G-Storm Thermal Cycler (GS1, Gene Technologies Ltd, Essex, UK). The amplification conditions and PCR primers used were as previously reported (Cuevas et al., 2010). Amplified products were resolved in 2% agarose gel stained with SybrSafe nucleic acid stain (Invitrogen, Carlsbad, CA, USA), and visualised with Dark Reader non-UV transilluminator (DR 195M, Clare Chemicals, Dolores, CO, USA).

4.5. Statistical analysis

Statistical analysis was carried out using binary logistic regression (Kleinbaum, Kupper, Muller, & Nizam, 1998), in CropStat (IRRI) for Windows (version 6.1.2007.1). This is a technique for predicting the mean value of a binary response variable as a function of one or more covariates. Statistical analyses of the metrics of the chain-length distributions, and of the number of chains, were conducted by balanced analysis of variance (ANOVA) in CropStat (IRRI) for Windows (version 6.1.2007.1). Comparison of means was done using least significant difference (LSD) at a 5% level of significance.

Acknowledgements

Rosa Paula Cuevas and Melissa Fitzgerald gratefully acknowledge the support of a RIRDC Grant (Project No. DAN 212A), and Robert G. Gilbert gratefully acknowledges the support of a grant from the Australia Research Council (DP0986043). We thank Ms. Ruby Jimenez and Ms. Violeta Bartolome for assisting in the statistical analysis of the data, Ms. Fatemeh Habibi and Mr. Arvin Tuaño for assistance in conducting FACE experiments, and Ms. Christina Virrey for assistance with differential scanning calorimetry. We thank the TT Chang Genetic Resources Centre of IRRI for providing the rice samples. The authors thank Christine Konik-Rose and Zhongyi Li for the provision of FACE data for wheat *SSIIa* genotypes.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.carbpol.2010.02.004.

References

Asaoka, M., Okuno, K., Sugimoto, K., Yano, M., Omura, T., & Fuwa, H. (1986). Characterization of endosperm starch from high-amylose mutants of rice (*Oryza sativa* L.). *Starch – Stärke, 38*(4), 114.

Ball, S. G., & Morell, M. K. (2003). From bacterial glycogen to starch: Understanding the biogenesis of the plant starch granule. *Annual Review of Plant Biology, 54*, 207–233.

Bao, J.-S., Corke, H., & Sun, M. (2006). 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.). *Theoretical and Applied Genetics, 113*(7), 1185–1196.

Cao, H., Imparl-Radosevich, J., Guan, H.-P., Keeling, P. L., James, M. G., & Myers, A. M. (1999). Identification of the soluble starch synthase activities of maize endosperm. *Plant Physiology, 120*(1), 205–216.

Castro, J. V., Dumas, C., Chiou, H., Fitzgerald, M. A., & Gilbert, R. G. (2005). Mechanistic information from analysis of molecular weight distributions of starch. *Biomacromolecules, 6*(4), 2248–2259.

Chiou, H., Fellows, C. M., Gilbert, R. G., & Fitzgerald, M. A. (2005). Study of rice starch structure by dynamic light scattering in aqueous solution. *Carbohydrate Polymers, 61*, 61–71.

Cohen, E. R., Cvitas, T., Frey, J. G., Holmstrom, B., Kuchitsu, K., Marquardt, R., et al. (2007). *Quantities, units, and symbols in physical chemistry (The IUPAC Green Book)*. Cambridge: Royal Society of Chemistry.

Cuevas, R. P., Daygon, V. D., Corpuz, H. M., Reinke, R. F., Waters, D. L. E., & Fitzgerald, M. A. (2010). Melting the secrets of gelatinisation temperature. *Functional Plant Biology*, in press.

Demorest, D., & Dubrow, R. (1991). Factors influencing the resolution and quantitation of oligonucleotides separated by capillary electrophoresis on a gel-filled capillary. *Journal of Chromatography, 559*(1–2), 43–56.

Fujita, N., Yoshida, M., Asakura, N., Ohdan, T., Miyao, A., Hirochika, H., et al. (2006). Function and characterization of starch synthase I using mutants in rice. *Plant Physiology, 140*(3), 1070–1084.

Fujita, N., Yoshida, M., Kondo, T., Saito, K., Utsumi, Y., Tokunaga, T., et al. (2007). Characterization of SSIIIa-deficient mutants of rice: The function of SSIIIa and pleiotropic effects by SSIIIa deficiency in the rice endosperm. *Plant Physiology, 144*, 2009–2023.

Fulton, T. M., Chunwongse, J., & Tanksley, S. D. (1995). Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Molecular Biology Reporter, 13*, 207–209.

Gray-Weale, A., & Gilbert, R. G. (2009). General description of the structure of branched polymers. *Journal of Polymer Science Part A: Polymer Chemistry, 47*(15), 3914–3930.

Hanashiro, I., Abe, J.-i., & Hizukuri, S. (1996). A periodic distribution of the chain length of amylopectin as revealed by high performance anion exchange chromatography. *Carbohydrate Research, 283*, 151–159.

Hennen-Bierwagen, T. A., Liu, F., Marsh, R. S., Kim, S., Gan, Q., Tetlow, I. J., et al. (2008). Starch biosynthetic enzymes from developing maize endosperm associate in multisubunit complexes. *Plant Physiology, 146*(4), 1892–1908.

Jane, J.-L., Chen, Y. Y., Lee, L. F., McPherson, A. E., Wong, K.-S., Radosavljevic, M., et al. (1999). Effects of amylopectin branch chain length and amylose content on the gelatinization and pasting properties of starch. *Cereal Chemistry, 76*(5), 629–637.

Jiang, H., Dian, W., & Wu, P. (2003). Effect of high temperature on fine structure of amylopectin in rice endosperm by reducing the activity of the starch branching enzyme. *Phytochemistry, 63*(1), 53–59.

Jones, R. G., Kahovec, J., Stepto, R., Wilks, E. S., Hess, M., & Kitayama, T., et al. (Eds.). . *Compendium of polymer technology and nomenclature: IUPAC recommendations 2008*. Cambridge: Royal Society of Chemistry.

Juliano, B. O., Perez, C. M., Kaushik, R. P., & Khush, G. S. (1990). Some grain properties of IR36-based starch mutants. *Starch – Stärke, 42*, 256–260.

Kang, H. J., Hwang, I. K., Kim, K. S., & Choi, H. C. (2003). Comparative structure and physicochemical properties of Ilpumbyeo, a high-quality japonica rice, and its mutant, Suweon 464. *Journal of Agricultural and Food Chemistry, 51*(22), 6598–6603.

Kleinbaum, D. G., Kupper, L. L., Muller, K. E., & Nizam, A. (1998). *Applied regression analysis and multivariable methods*. Pacific Grove, CA: Duxbury Press.

Konik-Rose, C., Thistleton, J., Chanvrier, H., Tan, I., Halley, P. J., Gidley, M. J., et al. (2007). Effects of starch synthase IIa gene dosage on grain, protein and starch in endosperm of wheat. *Theoretical and Applied Genetics, 115*(8), 1053–1065.

Lisle, A. J., Martin, M., & Fitzgerald, M. A. (2000). Chalky and translucent rice grains differ in starch composition and structure and cooking properties. *Cereal Chemistry, 77*(5), 627–632.

Morell, M. K., Kosar-Hashemi, B., Cmiel, M., Samuel, M. S., Chandler, P., Rahman, S., et al. (2003). Barley *sex6* mutants lack starch synthase IIa activity and contain a starch with novel properties. *Plant Journal, 34*(2), 172–184.

Nakamura, Y. (2002). Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: Rice endosperm as a model tissue. *Plant and Cell Physiology, 43*(7), 718–725.

Nakamura, Y., Francisco, P. B., Hosaka, Y., Sato, A., Sawada, T., Kubo, A., et al. (2005). Essential amino acids of starch synthase IIa differentiate amylopectin structure and starch quality between japonica and indica rice varieties. *Plant Molecular Biology, 58*(2), 213–227.

Nishi, A., Nakamura, Y., Tanaka, N., & Satoh, H. (2001). Biochemical and genetic analysis of the effects of amylose-extender mutation in rice endosperm. *Plant Physiology, 127*(2), 459–472.

O'Shea, M. G., Samuel, M. S., Konik, C. M., & Morell, M. K. (1998). Fluorophore-assisted carbohydrate electrophoresis (FACE) of oligosaccharides: Efficiency of labelling and high-resolution separation. *Carbohydrate Research, 307*(1–2), 1–12.

Ryoo, N., Yu, C., Cheon-Seok, P., Baik, M.-Y., Park, I.-M., Cho, M.-H., et al. (2007). Knockout of a starch synthase gene *OsSSIIIa*/*Flo5* causes white-core floury endosperm in rice (*Oryza sativa* L.). *Plant Cell Reports, 26*, 1083–1095.

Tanaka, N., Fujita, N., Nishi, A., Satoh, H., Hosaka, Y., Ugaki, M., et al. (2004). The structure of starch can be manipulated by changing the expression levels of starch branching enzyme IIb in rice endosperm. *Plant Biotechnology Journal, 2*(6), 507–516.

Tetlow, I. J., Beisel, K. G., Cameron, S., Makhmoudova, A., Liu, F., Bresolin, N. S., et al. (2008). Analysis of protein complexes in wheat amyloplasts reveals functional interactions among starch biosynthetic enzymes. *Plant Physiology, 146*(4), 1878–1891.

Tetlow, I. J., Morell, M. K., & Emes, M. J. (2004). Recent developments in the understanding the regulation of starch metabolism in higher plants. *Journal of Experimental Botany, 55*(406), 2131–2145.

Umemoto, T., & Aoki, N. (2005). Single-nucleotide polymorphisms in rice *starch synthase IIa* that alter starch gelatinisation and starch association of the enzyme. *Functional Plant Biology, 32*(9), 763–768.

Umemoto, T., Aoki, N., Lin, H., Nakamura, Y., Inouchi, N., Sato, Y., et al. (2004). Natural variation in rice *starch synthase IIa* affects enzyme and starch properties. *Functional Plant Biology, 31*, 671–684.

Umemoto, T., Yano, M., Satoh, H., Shomura, A., & Nakamura, Y. (2002). Mapping of a gene responsible for the difference in amylopectin structure between japonica-type and indica-type rice varieties. *Theoretical and Applied Genetics, 104*(1), 1–8.

Vandeputte, G. E., & Delcour, J. A. (2004). From sucrose to starch granule to starch physical behaviour: A focus on rice starch. *Carbohydrate Polymers, 58*(3), 245–266.

Waters, D. L. E., Henry, R. J., Reinke, R. F., & Fitzgerald, M. A. (2006). Gelatinization temperature of rice explained by polymorphisms in *starch synthase*. *Plant Biotechnology Journal, 4*(1), 115–122.

Yamakawa, H., Hirose, T., Kuroda, M., & Yamaguchi, T. (2007). Comprehensive expression profiling of rice grain filling-related genes under high temperature using DNA microarray. *Plant Physiology, 144*(1), 258–277.

Yano, M., Okuno, K., Kawakami, J., Satoh, H., & Omura, T. (1985). High amylose mutants of rice, *Oryza sativa* L.. *Theoretical and Applied Genetics, 69*(3), 253–257.