

# Diversity analysis of the *Waxy* gene in *Oryza sativa* L. “Guizhou HE”

Zhongni Wang<sup>a</sup>, Qinhan Luo<sup>a</sup>, Xue Jiang<sup>a</sup>, Xian Wu<sup>b</sup>, Haifeng Xu<sup>a</sup>, Susong Zhu<sup>a,\*</sup>

<sup>a</sup> Guizhou Rice Research Institute, Guizhou Academy of Agricultural Sciences, Guiyang 550006 China

<sup>b</sup> College of Life Sciences, Guizhou University, Guiyang 550025 China

\*Corresponding author, e-mail: 13984033281@139.com

Received 20 Dec 2020

Accepted 22 Mar 2021

**ABSTRACT:** The *Waxy* (*Wx*) gene controls the synthesis of amylose, the key determinant of rice cooking and eating qualities. To date, several *Wx* alleles and their effects on rice quality have been extensively elucidated. However, studies on the diversity of *Wx* in *Oryza sativa* L. landrace “Guizhou HE”, which plays important roles in local people’s daily lives, have been limited. Here, we investigated polymorphisms in *Wx* alleles and their correlations with the apparent amylose content (AAC) and gel consistency (GC) in “Guizhou HE”. Genotype identification and sequencing data showed that all the glutinous “Guizhou HE” varieties contained a *wx* allele having a 23-bp duplication in exon 2. At the *Wx* locus, six (CT)<sub>n</sub> microsatellite alleles, having n = 17, 18, 19, 20, 21, and 24, were found. Among them, (CT)<sub>20</sub> was the most frequent, and (CT)<sub>21</sub> was reported for the first time. The AAC in glutinous “Guizhou HE” ranged from 0.14% to 5.15%, and the GC ranged from 82.5 mm to 113.5 mm. No significant differences in AAC and GC were found among the (CT)<sub>n</sub> alleles. In addition, *Wx*<sup>g1</sup>, *Wx*<sup>g4</sup>, and *Wx*<sup>t1</sup> alleles were also identified in non-glutinous “Guizhou HE”. In conclusion, our results confirmed the existence of various *Wx* alleles in “Guizhou HE”, and they may serve as excellent resources for genetic diversity determination and rice quality analysis.

**KEYWORDS:** apparent amylose content, (CT)<sub>n</sub> microsatellite, gel consistency, “Guizhou HE”, *Waxy*

## INTRODUCTION

Rice (*Oryza sativa* L.) is an important staple food worldwide [1]. Better quality rice, especially varieties with good eating and cooking quality traits, are preferred by consumers. There are many parameters that determine rice cooking, sensory, and processing properties; and apparent amylose content (AAC) is considered being the most important parameter [2, 3]. According to varied ranges of AAC, rice varieties are classified as glutinous (≤6% AAC) and non-glutinous (>6% AAC) [4].

Amylose is synthesized by granule-bound starch synthase (GBSS), which is encoded by the *Waxy* (*Wx*) gene [5]. In addition to AAC, the *Wx* gene is responsible for gel consistency (GC) regulation and amylose molecule size modulation [6, 7]. At present, numerous *Wx* alleles have been identified. *Wx*<sup>a</sup> is associated with high AAC, while *Wx*<sup>in</sup> and *Wx*<sup>b</sup> are associated with intermediated and low AACs, respectively; and the *wx* allele is associated with glutinous varieties [8]. There is no detectable GBSS or AAC in varieties harboring the *wx* allele, owing to the premature termination of transcription caused by a fragment insertion in exon 2, exon 7,

exon 9, or intron 10. A frameshift due to a 23-bp duplication in exon 2 is common [9–12]. The improvement of amylose content by manipulating the *Waxy* gene is of prime concern in increasing the quality of rice [13].

In addition to polymorphisms in the *Wx* gene’s coding region, (CT)<sub>n</sub> microsatellites in the 5′ untranslated region were also identified [14, 15]. At present, more than 10 (CT)<sub>n</sub> alleles have been identified in non-glutinous rice varieties, including n = 8, 10, 11, 12, 14, 16, 17, 18, 19, 20, 21, and 22 [8, 14, 15]. The association of (CT)<sub>n</sub> microsatellites with AAC has been studied extensively in several germplasm collections, and (CT)<sub>n</sub> and AAC appear to be inversely correlated. Thus, lower numbers of repeats (e.g., n = 8 or 10) have been found in varieties with high AACs (> 23%), and, conversely, greater numbers of repeats (n > 16) have been found in varieties with lower AACs (< 23%) [8]. However, few (CT)<sub>n</sub> microsatellites have been reported in glutinous varieties, and their correlations with physicochemical parameters, such as AAC and GC, have not been comprehensively studied.

“Guizhou HE” is a traditional rice landrace mainly cultivated in the Qiandongnan Autonomous

Prefecture of Guizhou Province. As an original ecological rice landrace, “Guizhou HE” has some characteristic of wild rice, such as taller plants, thicker stems, wider leaves, larger spikes, and more awns. Since “Guizhou HE” is susceptible to temperature and light, it should be planted in certain areas to grow normally [17]. Most conserved “Guizhou HE” is glutinous with excellent qualities. It is not only the staple food, but also an intricate part of the daily life, nation custom, and religious beliefs of the local people [17]. In recent years, studies on “Guizhou HE” mainly focused on genetic diversity analyses [18], relationship with ethnic traditional culture [19], and resource conservation and preservation [17]. However, limited works have investigated the regulation of starch quality in “Guizhou HE”. The objective of the present study was to analyze polymorphisms of the *Wx* gene in “Guizhou HE” and examine the correlations between these *Wx* alleles and starch physicochemical parameter (AAC and GC), which will help determine the molecular basis of “Guizhou HE” quality.

## MATERIALS AND METHODS

### Plant materials

A total of 123 “Guizhou HE” samples (named as GZH1-123) for this study were collected from Liping, Congjiang, and Rongjiang counties of the Qianlongnan District, Guizhou Province. All the samples were planted on the experimental farm of the Rice Research Institute of Guizhou Province in Guiyang during the 2018 rice growing season. Standard crop cultivation practices were followed. Fresh leaves of 30-day-old seedlings were harvested and stored at  $-80^{\circ}\text{C}$  until uses.

### DNA isolation, PCR amplification, and sequencing

Genomic DNA was extracted from leaves using the CTAB method [20]. Five pairs of primers used for amplifying the whole *Wx* sequence are listed in Table S1. The primer pair SSR-F (5'-TACAAATAGCCACCCACAC-3') and SSR-R (5'-TGGGTGTGTTTCTCTAGACT-3') was designed to amplify the fragment containing the  $(\text{CT})_n$  and  $(\text{AATT})_n$  microsatellites of the *Wx* gene. The primer pair We2-1-F/R used by Sun et al [21], was applied to distinguish glutinous and non-glutinous rice genotypes. For the former, the amplification product was 421 bp, whereas for the latter, no amplification product was obtained.

PCR amplification was performed using

KOD\_FX (Toyobo, Tokyo, Japan). The PCR products were cloned into the pEASY-T3 vector for sequencing. The full-length *Wx* gene was assembled from amplicon sequences using identical overlapping regions. The sequences with new single nucleotide polymorphisms (SNPs) or insertion-deletion (InDels) were resequenced from additional independent PCR products.

### Determination of the AAC and GC

Mature seeds were harvested and stored at room temperature for 3 months. The AAC was determined as described by Dobo et al [15]. Sample of 0.1 g was put into a 50 ml flask, then 1 ml 95% ethanol and 9 ml 1 N KOH were added. The flask was heated in boiling water for 10 min. When the sample solution was cool down, it was transferred into a 100 ml flask, with several water washings up to 100 ml, and mixed well. Then 5 ml sample solution, 1 ml 1 N acetic acid, and 2 ml iodine solution (0.2 g iodine and 2.0 g potassium iodide in 100 ml aqueous solution) were pipetted into a 100 ml volumetric flask, and distilled water was added to make up the solution to the volume, The solution was shaken and let stand for 20 min before measuring the solution absorbance at 620 nm with a spectrophotometer. AAC was determined by reference to a standard curve. The GC was measured in accordance with the method reported by Tran et al [6]. 100 mg flour was weighed in duplicate into 20 ml tubes. 200  $\mu\text{l}$  95% ethyl alcohol containing 0.025% thymol blue, and 2 ml 1 N KOH were added to each tube. Tubes were shaken and, then, placed in vigorously boiling water bath for 8 min, kept at room temperature for 5 min, and sit in ice water bath for 15 min. Finally, the tubes were laid horizontally on top of graphing papers for 1 h, and the distance of gel migrated in each tube was measured.

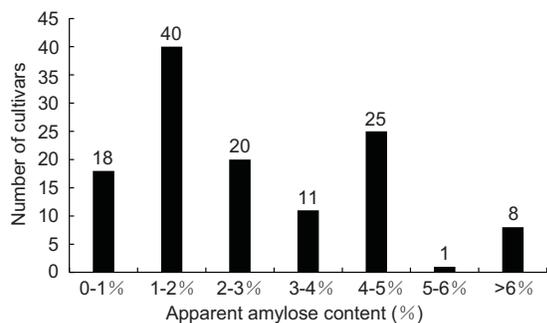
### Statistical analyses

Data were statistically analyzed using IBM SPSS statistic 22. A one-way analysis of variance and least significantly difference test, at a 0.05 probability level, were used to determine significant differences between alleles.

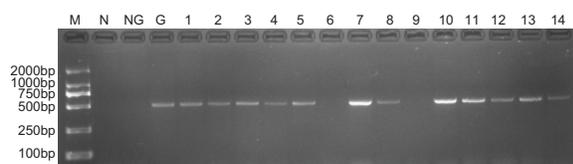
## RESULTS

### The AAC distribution in “Guizhou HE”

The AAC values in the 123 “Guizhou HE” samples ranged from 0.14% to 28.43%. Among them, sample GZH94 had the lowest AAC at 0.14%, and sample GZH102 had the highest AAC at 28.43%.



**Fig. 1** Apparent amylose content (AAC) distribution in “Guizhou HE”.



**Fig. 2** PCR products amplified with We2-1-F/R from representative “Guizhou HE” samples. M: 2000-bp marker; N: negative control; NG: non-glutinous variety Nipponbare; G: glutinous variety Suyunuo; lanes 1 and 2: GZH26, GZH71 (AAC<1%); lanes 3 and 4: GZH28, GZH96 (1%<AAC<2%); lane 5: GZH13 (AAC=5.15%); lane 6: GZH43 (AAC=11.91%); lanes 7 and 8: GZH58, GZH113 (2%<AAC<3%); lane 9: GZH102 (AAC=28.43%); lanes 10 and 11: GZH4, GZH5 (3%<AAC<4%); and lanes 12–14: GZH1, GZH2, GZH3 (4%<AAC<5%).

Using the AAC range, “Guizhou HE” was divided roughly into two groups: glutinous ( $\leq 6\%$  AAC) and non-glutinous ( $> 6\%$  AAC). In total, 93.5% of the cultivars (115 samples) belonged to the glutinous group, while the other 6.5% (8 samples) belonged to the non-glutinous group (Fig. 1). The result was in agreement with previous studies [19], which found that most “Guizhou HE” varieties were glutinous. Half of the glutinous samples (58 samples) had AACs less than 2%, while AACs of the other 57 accessions were between 2% and 6% (Fig. 1). Sample GZH13 had the highest AAC, at 5.15%, in the glutinous group.

#### Wx allele identification in glutinous “Guizhou HE”

The primer pair We2-1-F/R was used to detect the 23-bp duplication in exon 2. As expected, the 421-bp amplicon was present in the glutinous samples, but not in the non-glutinous control and

**Table 1** Samples sequenced corresponding to different apparent amylose contents (AAC).

AAC (%)	Sample no.	Sample name
0–1	4	GZH26, 71, 94, 104
1–2	2	GZH28, 96
2–3	2	GZH58, 113
3–4	8	GZH4, 5, 8, 9, 10, 11, 55, 103
4–5	13	GZH1, 2, 3, 6, 7, 12, 14, 15, 16, 17, 18, 19, 20
5–6	1	GZH13

non-glutinous “Guizhou HE” samples. PCR results of representative samples are shown in Fig. 2. To further verify the presence of the 23-bp insertion, the segment including exon 2 was amplified with primers WX2-F/R and sequenced in 30 selected “Guizhou HE” samples having different AAC values (Table 1). DNA sequencing results confirmed the existence of the 23-bp insertion (ACGGGTTCCAGGGCCTCAAGCCC) in glutinous “Guizhou HE” (Fig. 3). To further determine whether there were other mutations in the Wx gene, five overlapping primer pairs were used to amplify its genomic sequence in 10 “Guizhou HE” samples (GZH1–10). No mutation in the exon was found, except for (CT)<sub>n</sub> polymorphisms, and one single nucleotide mutation occurred in the promoter region. The G to A substitution at –619 in the promoter was observed in 8 samples, indicating that this is a prevailing point mutation.

#### Correlations between (CT)<sub>n</sub> microsatellite classes and the starch physicochemical parameters AAC and GC in glutinous “Guizhou HE”

The fragment, including the (CT)<sub>n</sub> and (AATT)<sub>n</sub> microsatellites in the 5′-untranslated region and the first intron splice site of the Wx gene, was amplified with primer pair SSR-F/R. Sequencing data allowed the identification of six classes of Wx (CT)<sub>n</sub> microsatellites, with n = 17, 18, 19, 20, 21, and 24, in the 115 glutinous samples. Among them, 64 samples had the (CT)<sub>20</sub> allele, and 33 samples had the (CT)<sub>18</sub> allele, indicating that these two alleles were the most frequent (Table 2). In addition, all the studied glutinous “Guizhou HE” possessed the AGTTATA sequence at the 5′-splice site of the first intron and (AATT)<sub>5</sub> in the first intron.

The correlations between the (CT)<sub>n</sub> microsatellite alleles and starch physicochemical parameters, AAC and GC, were analyzed (Table 2). Means of AACs in samples containing the six alleles ranged

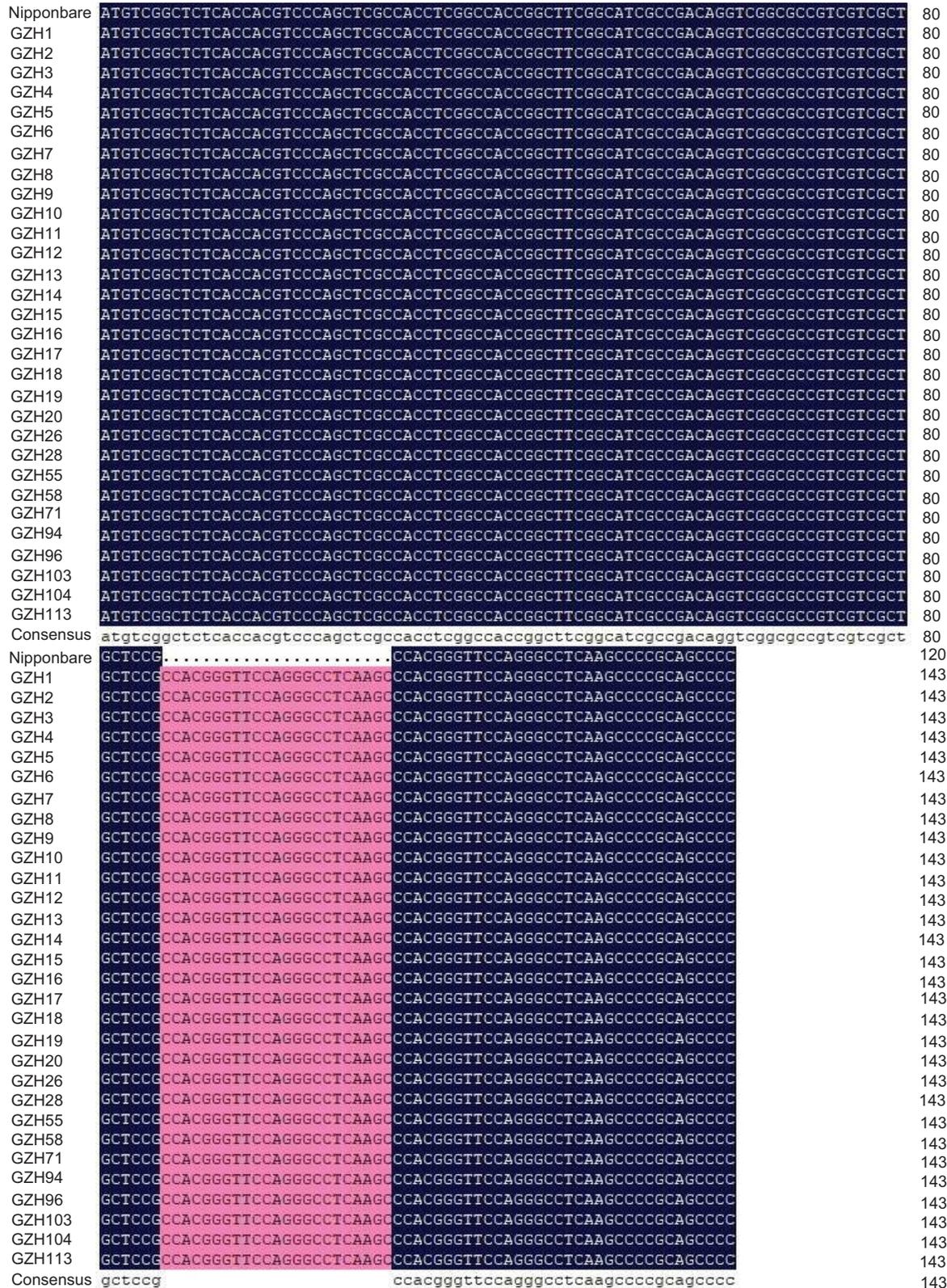


Fig. 3 Sequence alignment of 30 glutinous “Guizhou HE” samples and non-glutinous control Nipponbare of the region around the 23 bp duplication in exon 2.

**Table 2** Comparison of AAC and GC in six (CT)<sub>n</sub> classes of glutinous “Guizhou HE”.

Allele	Number	Frequency (%)	Range of AAC (%)	Mean of AAC (%)	Range of GC (mm)	Mean of GC (mm)
(CT) <sub>17</sub>	4	3.48	0.14–0.85	1.15	94.5–100.5	97.3
(CT) <sub>18</sub>	33	28.69	0.47–4.59	2.09	82.5–111.5	98.4
(CT) <sub>19</sub>	5	4.35	1.19–4.25	2.82	99.5–107.0	103.5
(CT) <sub>20</sub>	64	55.65	0.27–5.15	2.48	86.5–113.5	100.4
(CT) <sub>21</sub>	4	3.48	0.56–4.28	2.28	88.5–102.0	97.0
(CT) <sub>24</sub>	5	4.35	1.15–3.56	1.94	91.5–108.5	101.0

from 1.15% to 2.82%. All the samples in the (CT)<sub>17</sub> class had AACs of less than 2%, while samples in the other groups had higher AACs (2%–6%). However, there were no statistical differences in the AAC values among these allele classes. All the samples examined had high GC values, longer than 80 mm. There were no significant mean differences in GC values among the six (CT)<sub>n</sub> classes.

#### Wx alleles and AACs in non-glutinous “Guizhou HE”

To analyze the Wx alleles, the whole Wx sequence from the non-glutinous “Guizhou HE” was amplified and sequenced. The G/T polymorphism at the splicing donor site of the first intron and three SNPs in exons were identified (Table 3). The combination of the G/T polymorphism and SNPs in exon 6, 9, and 10 generated three allelic patterns: Wx<sup>g1</sup> (G-A-C-T), Wx<sup>g4</sup> (G-C-T-C), and Wx<sup>t1</sup> (T-A-T-C). The first letter in the allelic pattern corresponds to the G/T at the splicing donor site of the first intron; the second, third, and fourth letters to the A/C, C/T, and C/T polymorphisms, in exons 6, 9, and 10, respectively. The Wx<sup>g1</sup> pattern was found in three cultivars having intermediate (19%–23%) and high AACs (> 23%), whereas the Wx<sup>g4</sup> pattern was found in two cultivars having intermediate AACs. The Wx<sup>t1</sup> pattern was associated with the low AAC (6%–18%) class. For the (CT)<sub>n</sub> and (AATT)<sub>n</sub> microsatellites, (CT)<sub>11</sub>, (CT)<sub>12</sub>, and (AATT)<sub>6</sub> were associated with the Wx<sup>g1</sup> pattern; (CT)<sub>17</sub> and (AATT)<sub>5</sub> with the Wx<sup>g4</sup> pattern; and (CT)<sub>18</sub>, (CT)<sub>19</sub>, and (AATT)<sub>5</sub> with the Wx<sup>t1</sup> pattern (Table 3).

#### DISCUSSION

“Guizhou HE” is a traditional rice landrace mainly distributed in the Qiandongnan District of Guizhou Province, which is a birthplace of the “glutinous-food culture” [18]. Currently, > 90% of “Guizhou HE” is recognized as glutinous on the basis of their

opaque appearance [17]. However, the genetic and molecular bases of glutinous “Guizhou HE” are unknown. Glutinous rice has no detectable GBSS or amylose content because of a sequence mutation in the Wx gene. At present, several wx mutation types have been identified. Umeda et al [9] discovered a 139-bp short interspersed element in intron 10 of a wx mutant of *Oryza glaberrima*. Isshiki et al [10] identified two sequence mutations in exon 2 and exon 7. Wanchana et al [11] discovered a 23-bp duplicated sequence in exon 2 of tropical glutinous rice. A 7764-bp fragment inserted in exon 9 was detected by Hori et al [12]. According to the PCR amplification and DNA sequencing data, all of the glutinous “Guizhou HE” had the 23-bp duplication in exon 2, which is in agreement with previous studies in which the 23-bp duplication was common in glutinous rice [11]. In fact, the AAC in glutinous rice is < 2%. However, in our study, only 58 samples had AACs of < 2%, while the AACs in the other 57 samples were 2%–6%. Our result was in line with Hoai’s work in which the AAC values in more than 60% of the 165 tested glutinous cultivars were 2%–6% [4]. The higher AACs (> 2%) in glutinous samples might result from the presence of long chain amylopectin, which cannot be distinguished by iodine colorimetry [22]. Other detection methods, such as near-infrared reflectance spectroscopy and iodine voltage analysis, should be applied to determine the precise amylose content in “Guizhou HE”.

The (CT)<sub>n</sub> microsatellite in the Wx gene, first reported by Bligh et al [23], has been widely studied. To date, more than 10 (CT)<sub>n</sub> microsatellite alleles, with n = 8, 10, 11, 12, 14, 16, 17, 18, 19, 20, 21, and 22, have been detected in nonwaxy rice samples [14, 16, 24]. However, relatively few (CT)<sub>n</sub> alleles have been identified in glutinous rice. Bao et al [25] detected (CT)<sub>16</sub>, (CT)<sub>17</sub>, (CT)<sub>18</sub>, and (CT)<sub>19</sub> alleles from 56 waxy rice accessions, with (CT)<sub>17</sub> being the most frequent. (CT)<sub>17</sub> and

**Table 3** *Wx* alleles and apparent amylose contents (AAC) in non-glutinous “Guizhou HE”.

<i>Wx</i> alleles	Sample name	G/T	Ex6	Ex9	Ex10	(CT) <sub>n</sub>	(AATT) <sub>n</sub>	AAC (%)
<i>Wx<sup>g1</sup></i>	GZH42	G	TAT	CCC	TCT	12	6	22.69
	GZH41	G	TAT	CCC	TCT	11	6	22.18
	GZH102	G	TAT	CCC	TCT	11	6	28.43
<i>Wx<sup>g4</sup></i>	GZH36	G	TCT	CCT	CCT	17	5	20.26
	GZH40	G	TCT	CCT	CCT	17	5	19.61
<i>Wx<sup>t1</sup></i>	GZH43	T	TAT	CCT	CCT	18	5	11.91
	GZH84-2	T	TAT	CCT	CCT	19	5	11.33
	GZH65	T	TAT	CCT	CCT	18	5	15.03

(CT)<sub>18</sub> were also detected in glutinous landraces from Thailand [11]. The present study identified six (CT)<sub>n</sub> alleles, with n = 17, 18, 19, 20, 21, and 24, from 115 glutinous “Guizhou HE” samples, with (CT)<sub>20</sub> being the most frequent (Table 2). Sun et al [21] reported eight alleles, with n = 10, 15, 16, 17, 18, 19, 20, and 24, from 66 Guizhou local glutinous rice varieties, with (CT)<sub>20</sub> also being the most frequent. These results support the hypotheses that the genetic diversity of “Guizhou HE” is rich and the population structure of “Guizhou HE” is unique, making it different from glutinous rice in other regions [26]. Although (CT)<sub>10</sub>, (CT)<sub>15</sub>, and (CT)<sub>16</sub> were not found in our samples, which may be the result of different rice samples being used for genotyping, (CT)<sub>21</sub> was reported for the first time. The (CT)<sub>n</sub> microsatellite in the *Wx* gene is a valuable molecular marker to evaluate AAC values in nonwaxy rice [15,24]. However, the utility of this microsatellite in glutinous rice is limited. Our analysis revealed no correlations between (CT)<sub>n</sub> alleles and the AAC and GC parameters. Therefore, the utility of this microsatellite in glutinous rice needs to be further explored.

In addition to *wx* in glutinous “Guizhou HE”, *Wx* alleles in non-glutinous “Guizhou HE” were also analyzed. Three allelic patterns were identified, *Wx<sup>g1</sup>* pattern (G-A-C-T), *Wx<sup>g4</sup>* pattern (G-C-T-C), and *Wx<sup>t1</sup>* pattern (T-A-T-C), which corresponded to high, intermediate and low AACs, respectively. Additionally, (CT)<sub>11</sub>, (CT)<sub>12</sub>, (CT)<sub>17</sub>, (CT)<sub>18</sub>, and (CT)<sub>19</sub> were also identified in the above samples. Therefore, the results of our study support the hypothesis that the level of genetic diversity in “Guizhou HE” was high [18], and the overall study suggests that “Guizhou HE” is an advantageous resource valuable for genetic diversity and rice quality analyses.

### Appendix A. Supplementary data

Supplementary data associated with this article can be found at <http://dx.doi.org/10.2306/scienceasia1513-1874.2021.055>.

**Acknowledgements:** This project was financially supported by grants from the National Key Research and Development Project (2017YFD0100402), Science and Technology Platform and Talent Team Plan of Guizhou Province ([2017] 5719 and [2018] 5620), and Joint Fund of the National Natural Science Foundation of China and the Karst Science Research Center of Guizhou Province (Grant No. U1812401). We thank International Science Editing (<http://www.internationalscienceediting.com>) for editing this manuscript.

### REFERENCES

1. Tilman D, Balzer C, Hill J, Befort BL (2011) Global food demand and the sustainable intensification of agriculture. *Proc Natl Acad Sci USA* **108**, 20260–20264.
2. Li HY, Prakash S, Nicholson TM, Fitzgerald MA, Gilbert RG (2016) The importance of amylose and amylopectin fine structure for textural properties of cooked rice grains. *Food Chem* **196**, 702–711.
3. Nakwilai P, Cheabu S, Narumon P, Saensuk C, Arikrit S, Malumpong C (2020) Evaluation of japonica rice (*Oryza sativa* L.) varieties and their improvement in terms of stability, yield and cooking quality by pure-line selection in Thailand. *ScienceAsia* **46**, 157–168.
4. Hoai T, Matsusaka H, Toyosawa Y, Suu T, Satoh H, Kumamaru T (2014) Influence of single-nucleotide polymorphisms in the gene encoding granule-bound starch synthase I on amylose content in Vietnamese rice cultivars. *Breeding Sci* **64**, 142–148.
5. Hirano HY, Sano Y (1991) Molecular characterization of the waxy locus of rice (*Oryza sativa* L.). *Plant Cell Physiol* **32**, 989–997.
6. Tran N, Daygon V, Resurreccion A, Cuevas R, Corpuz H, Fitzgerald M (2011) A single nucleotide polymorphism in the *Waxy* gene explains a significant

- component of gel consistency. *Theor Appl Genet* **123**, 519–525.
7. Zhang C, Zhu J, Chen S, Fan X, Li Q, Lu Y, Wang M, Yu H, et al (2019)  $Wx^{dv}$ , the ancestral allele of rice *Waxy* gene. *Mol Plant* **12**, 1157–1166.
  8. Chen MH, Bergman C, Pinson S, Fjellstrom R (2008) *Waxy* gene haplotypes: Associations with apparent amylose content and the effect by environment in an international rice germplasm collection. *J Cereal Sci* **47**, 536–545.
  9. Umeda M, Ohtsubo H, Ohtsubo E (1991) Diversification of the rice *Waxy* gene by insertion of mobile DNA elements into introns. *Jpn J Genet* **66**, 569–586.
  10. Isshiki M, Yamamoto Y, Satoh H, Shimamoto K (2001) Nonsense-mediated decay of mutant *waxy* mRNA in rice. *Plant Physiol* **125**, 1388–1395.
  11. Wanchana S, Toojinda T, Tragoonrungs S, Vanavichit A (2003) Duplicated coding sequence in the *waxy* allele of tropical glutinous rice (*Oryza sativa* L.). *Plant Sci* **165**, 1193–1199.
  12. Hori Y, Fujimoto R, Sato Y, Nishio T (2007) A novel *wx* mutation caused by insertion of a retrotransposon-like sequence in a glutinous cultivar of rice (*Oryza sativa* L.). *Theor Appl Genet* **115**, 217–224.
  13. Cho KS, Kongsil P, Wangsawang T, Sreewongchai T (2020) Marker-assisted pseudo-backcross breeding for improvement of amylose content and aroma in Myanmar rice cultivar Sinthukha. *ScienceAsia* **46**, 539–547.
  14. Bao JS, 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.). *Theor Appl Genet* **113**, 1185–1196.
  15. Dobo M, Ayres N, Walker G, Park WD (2010) Polymorphism in the GBSS gene affects amylose content in US and European rice germplasm. *J Cereal Sci* **52**, 450–456.
  16. Ayres NM, McClung AM, Larkin PD, Blich HF, Jones CA, Park WD (1997) Microsatellites and a single-nucleotide polymorphism differentiate apparent amylose classes in an extended pedigree of US rice germ plasm. *Theor Appl Genet* **94**, 773–781.
  17. Wang Y, Jiao A, Chen H, Ma X, Cui D, Han B, Ruan R, Xue D, et al (2018) Status and factors influencing on-farm conservation of Kam Sweet Rice (*Oryza sativa* L.) genetic resources in southeast Guizhou Province. *China J Ethnobiol Ethnomed* **14**, ID 76.
  18. Lei QY, Xiong Y, Zhou JJ, Zhang WH, Sun J, Long CL (2017) Genetic diversity analysis and evaluation of Kam aromatic glutinous rice in Qiandongnan Prefecture based on AFLP molecular markers. *J Plant Genet Resour* **18**, 1023–1031. [in Chinese]
  19. Wang Y, Wang Y, Jiao A, Cai Z, Yang J, Ruan R, Xue D (2015) Influence of national traditional culture on crop genetic diversity—take an example of Kam Sweet Rice in Liping County of Guizhou Province. *J Nat Resour* **30**, 617–628.
  20. Fjellstrom R, Conaway-Bormans C, McClung AM, Marchetti MA, Shank AR, Park WD (2004) Development of DNA markers suitable for marker assisted selection of three *Pi* genes conferring resistance to multiple *Pyricularia grisea* pathotypes. *Crop Sci* **44**, 1790–1798.
  21. Sun Y, Ma J, Yang Y, Chen H, Jiao A, Tan J, Ruan R, Xu M (2019) Diversity analysis of *Wx* gene sequence of local rice varieties in Guizhou. *Chin Agr Sci Bull* **35**, 1–6. [in Chinese]
  22. Jiao MY, Gao H, Wang WN, Tian YL (2019) Comparison of four methods for the determination of amylose and amylopectin. *Sci Technol Food Ind* **40**, 259–264. [in Chinese]
  23. Blich HF, Till RI, Jones CA (1995) A microsatellite sequence closely linked to the *waxy* gene of *Oryza sativa*. *Euphytica* **86**, 83–85.
  24. Shu QY, Wu DX, Xia YW, Gao MW, Ayres NM, Larkin PD, Park WD (1999) Microsatellite polymorphisms on the *waxy* gene locus and their relationship to amylose content in *indica* and *japonica* rice, *Oryza sativa* L. *J Genet Genomics* **26**, 350–358.
  25. Bao JS, Corke H, Sun M (2002) Microsatellites in starch-synthesizing genes in relation to starch physicochemical properties in waxy rice (*Oryza sativa* L.). *Theor Appl Genet* **105**, 898–905.
  26. Wu X, Xu H, Zhang Z, Wang Q, Gong Y, Zhang X, Long S, Luo Q, et al (2020) Genetic diversity and population structure study of Guizhou He based on KASP marker. Available at: <https://kns.cnki.net/kcms/detail/46.1068.S.20201014.0958.006.html>.

**Appendix A. Supplementary data****Table S1** List of primers used for amplifying the whole sequence of *Wx* gene.

Primer name	Forward primer (5'-3')	Reverse primer (5'-3')
WX-2	ATCTGATCTGCTCAAAGCTCTGTGCATCTC	TCCACGCTTGTAGCAATGGAAAAACCTCAC
WX-3	GTTCTTGATCATCGCATTGG	ACTTGCCTTGCTAGGATCC
WX-4	GAAGATCAACTGGATGAAGG	GGCATGGTATAATATGGAAC
WX-5	TATTGACAAGTCAAGGGACC	AGCATGTATGAGACTACTTG
WX-6	CTCTTATTCAGATCGATCAC	CCATATGTGCTTATAAGGAC