

OPEN

# The *Zea mays* mutants *opaque2* and *opaque16* disclose lysine change in waxy maize as revealed by RNA-Seq

Wei Wang<sup>1,2</sup>, Suzhen Niu<sup>1</sup>, Yi Dai<sup>1</sup>, Mingchun Wang<sup>2</sup>, Yan Li<sup>1</sup>, Wenpeng Yang<sup>2</sup> & Degang Zhao<sup>1,2</sup>

In maize, *opaque2* (*o2*) and *opaque16* (*o16*) alleles can increase lysine content, while the *waxy* (*wx*) gene can enhance the amylopectin content of grains. In our study, *o2* and *o16* alleles were backcrossed into waxy maize line (*wxwx*). The *o2o2o16o16wxwx* lines had amylopectin contents similar to those of waxy line. Their nutritional value was better than waxy line, but the mechanism by which the *o2* and *o16* alleles increased the lysine content of waxy maize remained unclear. The *o2o2o16o16wxwx* lines and their parents on kernels (18th day after pollination) were subjected to RNA sequencing (RNA-Seq). The RNA-Seq analysis revealed 272 differentially expressed genes (DEGs). Functional analyses revealed that these DEGs were mainly related to biomass metabolism. Among them, in *o2o2o16o16wxwx* lines, 15 genes encoding  $\alpha$ -zein were down-regulated, which resulted in the reduction of  $\alpha$ -zein synthesis and increased lysine content; *lkr/sdh1* and *Zm00001d020984.1* genes involved in the lysine degradation pathway were down-regulated, thereby inhibited lysine degradation; *sh2*, *bt2* and *ae1* genes involved in starch metabolism were upregulated, led to wrinkling kernel and farinaceous endosperm. Our transcriptional-level identification of key genes responsible for increased grain lysine content and farinaceous endosperm formation following introgression of *o2* and *o16* alleles should promote molecular breeding for maize quality.

Maize (*Zea mays* L.), a very important food and feed crop, has a low protein nutritional value because it lacks lysine, a necessary amino acid in humans and monogastric animals. Maize line *opaque2* (*o2*) is a high-lysine mutant. The *O2* gene, which is located on the short arm of maize chromosome 7, encodes a leucine zipper family transcription factor containing a basic domain that activates the expression of 22 kDa  $\alpha$ -zein and 15 kDa  $\beta$ -zein genes<sup>1</sup> and can also directly or indirectly regulate other non-storage protein genes, such as *b-32* and *b-70*<sup>2</sup>. The *o2* mutant, which is widely used in genetics and breeding studies, has a grain lysine content of approximately 0.4%. Afterwards, several other mutations such as *floury1* (*(fl1, (opaque8, o8; opaque4, o4))*<sup>3–5</sup>, *opaque5* (*o5*)<sup>6,7</sup>, *proline1* (*pro1, (opaque6, o6)*<sup>8–10</sup>, *opaque7* (*o7*)<sup>11–14</sup>, *shrunk4* (*sh4, (opaque9, o9)*<sup>15,16</sup>, *opaque15* (*o15*)<sup>17</sup>, *De\*-B30*<sup>18</sup>, *Mc*<sup>18</sup>, *floury2* (*fl2*)<sup>19,20</sup> and *floury3* (*fl3*)<sup>16,21,22</sup> were discovered. These mutants have been experimentally tried singly or in combinations, but resulted in severe yield losses due to negative effects of the individual mutation<sup>23,24</sup>. To enhance the germplasm resources of high lysine maize, we previously isolated a novel high lysine mutation from Robertson's *Mutator* (*Mu*) stocks, which had opaque endosperm and was named *opaque16* (*o16*). The mutant line had a grain lysine content of over 0.36%. The corresponding gene was located on chromosome 8L between molecular markers *umc1141* and *umc1121* within a 3 cM distance of *umc1141*<sup>25</sup>. The *o16* mutant is useful for germplasm improvement and quality breeding<sup>26–28</sup>, and its identification and study have contributed to high-lysine maize germplasm resources.

Despite the above improvements, high-lysine maize with a single gene mutation cannot meet the nutritional quality needs of food and food processing nor the lysine content requirements of livestock and poultry feed<sup>29</sup>. To further improve the lysine content of maize by marker-assisted selection (MAS), *o16* and *o2* alleles were pyramided. This approach yielded grains with a lysine content higher than 0.5%<sup>25,30</sup>, sufficient for the needs of human consumption. Konsam *et al.*<sup>28</sup> then backcrossed *o16* alleles into *o2*-based parental inbreds (HKI161,

<sup>1</sup>The State Key Laboratory Breeding Base of Green Pesticide and Agricultural Bioengineering, The Key Laboratory of Plant Resources Conservation and Germplasm Innovation in Mountainous Region (Ministry of Education), Guizhou University, Guiyang, 550025, China. <sup>2</sup>Guizhou Institute of Upland Food Crops, Guiyang Station for DUS Testing Center of New Plant Varieties (MOA), Guizhou Academy of Agricultural Sciences, Guiyang, 550006, China. Correspondence and requests for materials should be addressed to W.Y. (email: [ywpmaize@126.com](mailto:ywpmaize@126.com)) or D.Z. (email: [dgzhaogzu.edu.cn](mailto:dgzhaogzu.edu.cn))

Received: 28 January 2019

Accepted: 31 July 2019

Published online: 22 August 2019

HKI193-1, HKI193-2 and HKI163) of four quality protein maize (QPM) hybrids (HQPM-1, HQPM-4, HQPM-5 and HQPM-7) using marker-assisted backcross breeding (MABB). Compared with the recurrent parents, the contents of lysine and tryptophan increased by 76% and 91%, respectively. Average lysine and tryptophan contents of the hybrids increased by 49% and 60%, respectively, with maximum increases of 64% and 86%, respectively. Consequently, the pyramiding of high-lysine mutant genes can increase lysine and tryptophan contents to improve the nutritional quality of maize grains.

Waxy maize endosperm contains 95–100% amylopectin<sup>31</sup>. The *Waxy1* (*Wx*) gene is located on the short arm of chromosome 9, encodes granule-bound starch synthase I (GBSS I). The activity of GBSS I is significantly decreased in the *wx1* mutant, leading to the low level of amylose but high level of amylopectin in maize endosperm and pollen<sup>32,33</sup>. Waxy maize has excellent taste, texture and other culinary qualities, but its nutritional value is relatively low, and has a lysine content of only 0.24–0.34%<sup>34</sup>. To boost the lysine content of waxy maize, Zhang *et al.*<sup>35</sup> used MAS to generate 18 inbred quality-protein maize lines containing *wx* and *o2* alleles. These lines had lysine contents of 0.36–0.54%, which was 1.15–27.06% higher than those of the original parents. Using MAS, Zhang *et al.*<sup>36</sup> backcrossed the *o2* allele into waxy maize and increased lysine contents by 48.5–61.9%. Yang *et al.*<sup>37</sup> incorporated the *o16* allele into waxy maize by crossing and backcrossing; they obtained four and three families, respectively, with lysine content increases of 18–28% compared with the original waxy maize lines. Also using MABB, Zhang *et al.*<sup>38</sup> introgressed *o2* and *o16* alleles into *wxwx* lines and obtained three waxy maize lines containing *o2* and *o16* alleles, the mean grain lysine content was 0.62%, and the waxy property was equal to that of the recurrent parents, consequently meeting the food and feed requirements of humans, livestock and poultry. The pyramiding of high-lysine genes (*o2* and *o16*) and the waxy gene (*wx*) can therefore improve grain lysine content while retaining a constant waxy property<sup>37–39</sup>, thus yielding maize germplasm with good nutrition and flavour.

Single gene mutation and multiple-mutation pyramiding techniques can not only improve the nutritional quality of maize, but also affect body metabolism. Introgression of the recessive *o2* gene into common maize materials with different genetic backgrounds can cause different changes in transcription patterns<sup>40</sup>. The *o2* allele can also trigger physiological, biochemical, and proteomic changes in waxy maize<sup>41</sup>, *o2* introgression not only decreased the accumulation of various zein proteins, but also affected other endosperm proteins related to amino acid biosynthesis, starch-protein balance, stress response and signal transduction. In addition, the pyramiding of *o2* and *o7* alleles may affect amino acid metabolism, carbon metabolism, storage protein synthesis, transcription and translation, and signal transduction<sup>42</sup>. However, which physiological metabolic changes will be induced by the simultaneous introgression of *o2* and *o16* alleles into waxy maize is unknown, the molecular underlying mechanism of the *o2* and *o16* alleles introgression into waxy corn to increase the lysine content is still unknown.

To clarify the above questions, *o2* and *o16* alleles were backcrossed to waxy maize using MABB. Through target-gene foreground selection and genome background selection, mutant maize lines QCL8006\_1 and QCL8006\_2 containing three recessive *o2*, *o16* and *wx* gene mutations were obtained. To identify the molecular mechanism underlying the increased lysine content of waxy maize following introgression of *o2* and *o16* alleles, we carried out a transcriptome analysis of kernels on 18th day after pollination (18DAP) to examine transcriptional expression differences between *o2o2o16o16wxwx* mutants and their recurrent parent.

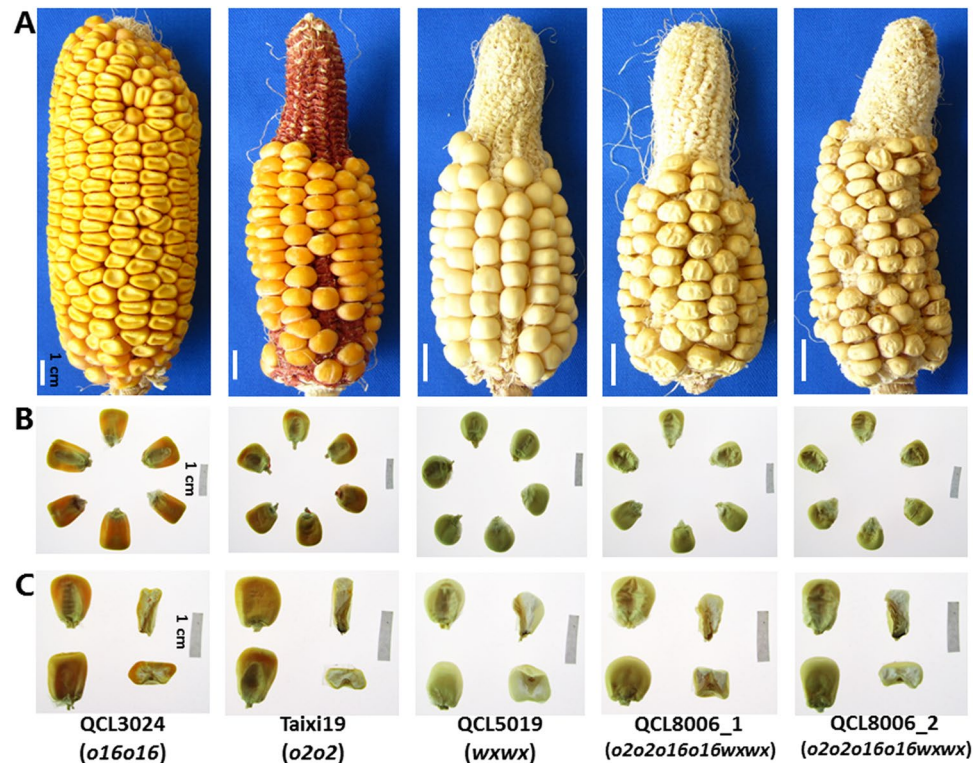
## Results

**Kernel characteristics and submicroscopic structure.** Grain phenotypes and endosperm cross-sections of *o2o2o16o16wxwx* mutant lines and their parent were observed under natural light, and the grain transparency was examined under projected light. Grain coats of *o2o2o16o16wxwx* mutant lines were non-glossy and displayed different degrees of wrinkling (Fig. 1A); the grains were completely opaque (Fig. 1B), with farinaceous endosperm and no full kernels (Fig. 1C). Grain coats of wild type (WT, CML535) were smooth and lustrous (Supplementary Fig. S1A), and the grains were vitreous (Supplementary Fig. S1B). Grains of *o16* mutant (QCL3024) were opaque with modifier, and grains of *o2* mutant (Taixi19) were opaque (Supplementary Fig. S1C). Grain coats of recurrent parent QCL5019 were smooth and lustrous, the grains were opaque and full with waxy endosperm.

Meanwhile, scanning electron microscopy revealed that the grain-endosperm starch granules of the *o2o2o16o16wxwx* mutants had an irregular shape and arrangement and an uneven volume and size, with a high density of matrix proteins that dispersed in the gap between starch granules. The grain-endosperm starch granules of recurrent parent QCL5019 were mostly ellipsoid or spherical; the matrix-protein density was low, and the starch granules were closely encapsulated (Fig. 2A,B).

**Changes of protein, starch and FAAs composition in *o2o2o16o16wxwx* endosperm.** To know biochemical differences between *o2o2o16o16wxwx* lines and their parent QCL5019, we measured protein, starch and free amino acids (FAAs) contents for their mature kernels. The total protein contents of QCL8006\_1 and QCL8006\_2 were 11.73% and 11.71%, respectively, which reduced by 4.52% and 4.65%, respectively, in comparison to QCL5019 (Fig. 3A). The total starch contents of QCL8006\_1 and QCL8006\_2 were 67.10% and 67.48%, respectively, and not significantly different from QCL5019 (Fig. 3B). The contents of 17 FAAs differed by varying degrees between recurrent parent QCL5019 and the *o2o2o16o16wxwx* mutants. Figure 3C shows that aspartate (Asp), glycine (Gly), cysteine (Cys), valine (Val), methionine (Met), histidine (His), lysine (Lys), and arginine (Arg) contents were higher in the mutants—significantly so in the case of Lys, Cys, Arg and Gly, which increased by 75.10%, 64.85%, 54.45%, and 40.33%, respectively. However, the contents of the remaining nine amino acids were lower in the mutants, and the total amino acid content also decreased slightly.

**The quality of RNA-seq and biological replicates.** To dissect the regulator network of the *o2o2o16o16wxwx*, kernels (18DAP) of the *o2o2o16o16wxwx* and their parent lines were used for RNA-seq



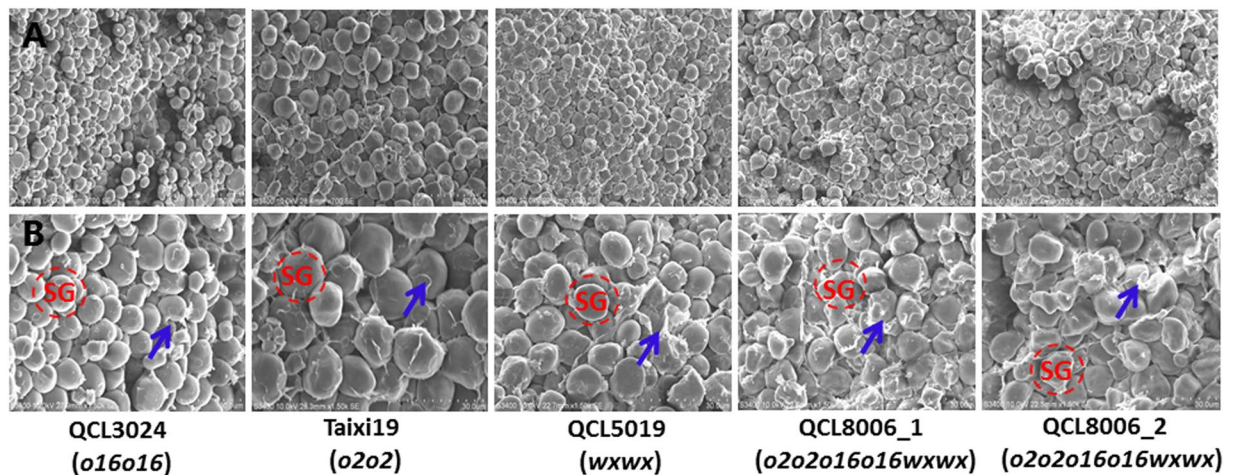
**Figure 1.** Phenotypic features of *o2o2o16o16wxwx* lines and their parents. (A) Photographs of intact ears taken under normal light. (B) Light transmission of mature kernels on a light box. (C) Cross-sections of mature kernels on a light box, Bars = 1 cm.

analysis. Each sample generated on average 23,648,912 clean reads (Supplementary Table S1), the average mapping ratio to reference gene and genome were 82.47% (Supplementary Table S2) and 91.12% (Supplementary Table S3). The average Q20 and Q30 (The percentage of the number of bases with Sequencing base mass value greater than 20 or 30 in the total number of bases in the original data) of all samples were 95.65% and 86.18%, respectively (Supplementary Table S4), which indicates that the sequencing data were of good quality.

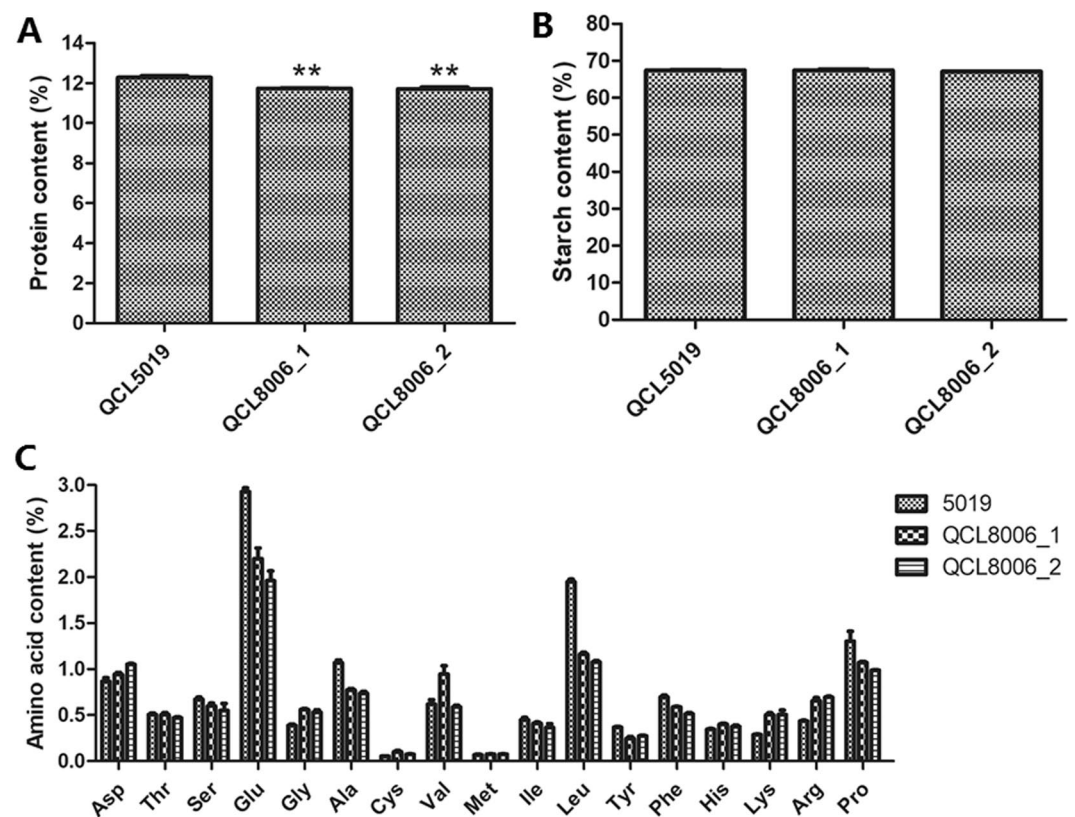
According to our quantitative analysis (Supplementary Table S5), 29,137, 28,875 and 28,589 genes were respectively detected in QCL8006\_1 and QCL8006\_2 and their recurrent parent QCL5019; these numbers corresponded to more than 70% of the total number of genes (Supplementary Fig. S2), thus indicating that the sequencing saturation was in line with expected levels. Correlation coefficients among three sequencing repeats of QCL8006\_1, QCL8006\_2 and QCL5019 were 0.9516, 0.9948 and 0.9494, respectively (Supplementary Fig. S3); all of these values were higher than 0.92, which indicates that the transcriptome sequencing data of all three biological repeats in all samples were highly reliable and thus suitable for analysing differentially expressed genes.

**Identification of differentially expressed genes (DEGs).** Gene expression differences between QCL8006\_1 and QCL8006\_2 and their recurrent parent QCL5019 were analysed using NOISeq software, and hierarchical clustering was also carried out. Compared with the recurrent parent QCL5019, 390 genes were differentially expressed in QCL8006\_1, among which 299 were up-regulated and 91 were down-regulated. Meanwhile, in QCL8006\_2, 917 genes were differentially expressed relative to QCL5019, and among them, 747 were up-regulated and 170 were down-regulated (Fig. 4A). In addition, to minimize differences due to different genetic recombination and genetic background, we examined the intersection of differentially expressed genes (DEGs) between QCL8006\_1 and QCL8006\_2 by COUNTIF function in Excel 2010 and the results show that there were 272 DEGs in common, 185 up-regulated, 79 down-regulated in both lines of QCL8006-1 and QCL8006-2, and other 8 up-regulated in QCL8006\_1 but down-regulated in QCL8006\_2 (Fig. 4B and Supplementary Table S6). It is thus evident that the expression trends of those 272 DEGs were almost identical, although there were obvious differences in DEGs detected between QCL8006\_1 and QCL8006\_2.

**GO annotation and KEGG analysis.** WEGO software<sup>43</sup> was used to classify the differentially expressed genes into Gene Ontology (GO) functional categories related to biological process (BP), cellular component (CC), and molecular function (MF) (Supplementary Table S7). A total of 96 genes were classified into BP, mainly metabolic process (GO: 0008152), cellular process (GO: 0009987), single-organism process (GO: 0044699), localization (GO: 0051179), and response to stimulus (GO: 0050896). In terms of CC, 120 genes were classified mostly into categories related to cell (GO: 0005623), cell part (GO: 0044464), and organelle (GO: 0043226). A total of 117 genes were classified into MF categories, primarily those involving binding (GO: 0005488), catalytic activity (GO: 0003824) and molecular function regulation (GO: 0098772) (Fig. 5).

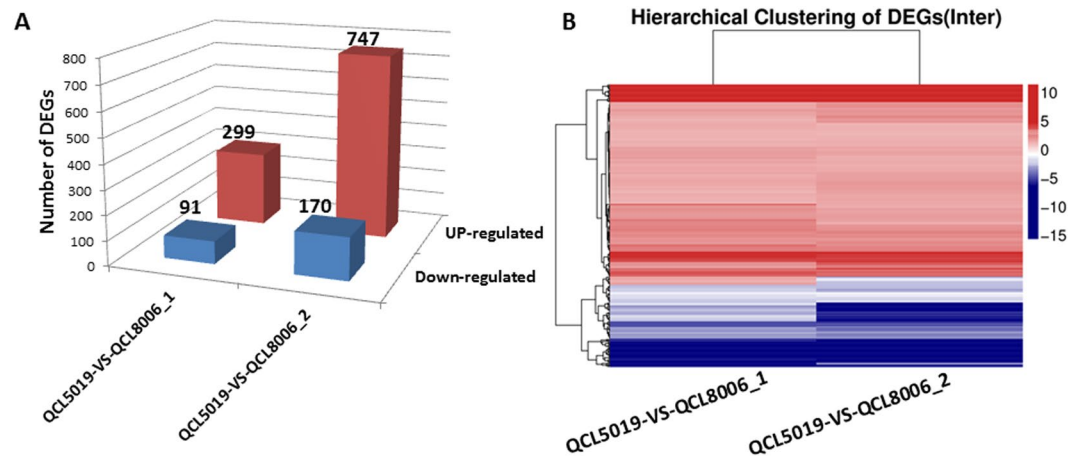


**Figure 2.** (A) Scanning electron micrograph for endosperms of QCL3024, Taixi19, QCL5019, QCL8006\_1 and QCL8006\_2 at 700 $\times$  magnification, respectively. (B) Scanning electron micrograph for endosperms of QCL3024, Taixi19, QCL5019, QCL8006\_1 and QCL8006\_2 at 1500 $\times$  magnification. SG, starch granules; Blue arrows, matrix protein.

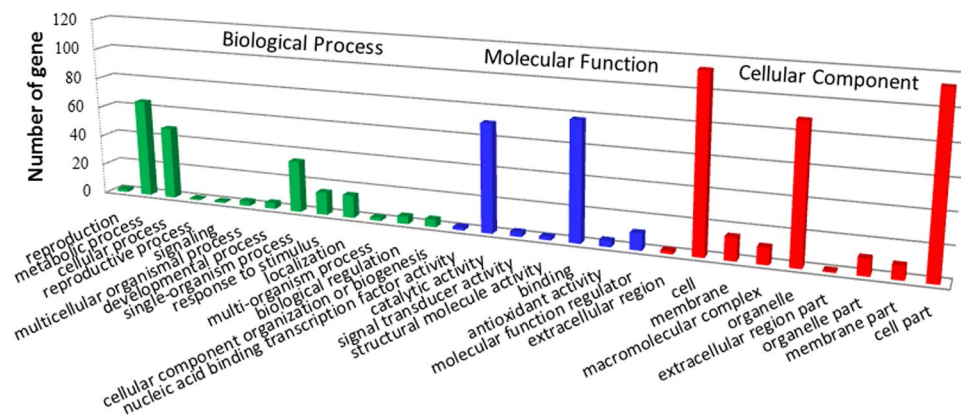


**Figure 3.** The contents of protein (A), starch (B) and 17 FAAs (C) in mature kernels of QCL5019, QCL8006\_1 and QCL8006\_2. \* $P < 0.05$  and \*\* $P < 0.01$ .

To further functional characterization of DEGs, pathway analysis of DEGs based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database assigned 79 DEGs to 64 KEGG pathways (Supplementary Table S8 and Fig. 6). Seventy-five of these genes were annotated into metabolic pathways. Among genes related to metabolic pathways, 21 were related to carbohydrate metabolism, 13 to global and overview maps, 13 to energy metabolism, and 12 to amino acid metabolism. Of the 12 differentially expressed genes related to amino acid metabolism, seven were involved in amino acid synthesis and five participated in amino acid degradation.



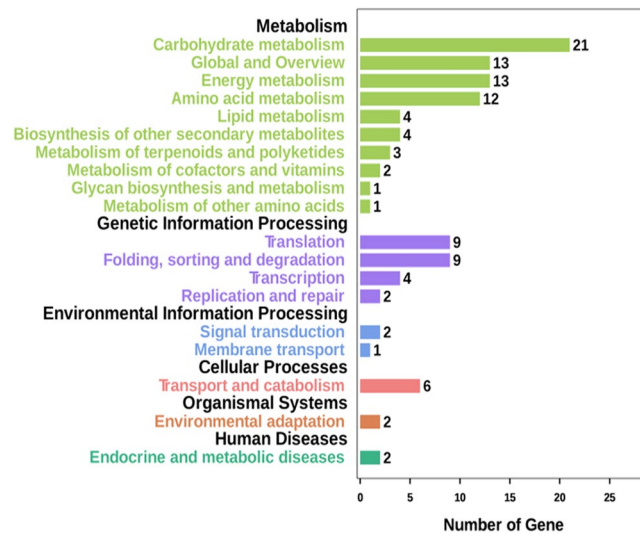
**Figure 4.** (A) The column diagram of DEGs for QCL5019 vs. QCL8006\_1 and QCL5019 vs. QCL8006\_2. X axis represents pairwise and Y axis means number of screened DEGs. Blue bar denotes down-regulated genes and red bar for the up-regulated. (B) The intersection heatmap of DEGs for QCL5019 vs. QCL8006\_1 and QCL5019 vs. QCL8006\_2. Gradient color barcode at the right top indicates log<sub>2</sub>(FC) value (FC, Fold change of expression in triple recessive mutant vs waxy parent). Each row represents a DEG and each column represents one condition pairwise. DEGs with similar fold change value are clustered both at row and column level.



**Figure 5.** The GO analysis of differentially expressed genes. X axis represents GO terms. Y axis means number of DEGs (the number is presented by its square root value). All GO terms are grouped into three ontologies: brown is for biological process, orange is for cellular component and blue is for molecular function.

**Amino acid metabolism.** In regards to pathways related to amino acid synthesis, *Zm00001d014258.1* and *Zm00001d030557.1*, which encode glutamate-pyruvate aminotransferase and participate in the conversion of pyruvic acid to alanine in the alanine synthesis pathway, were up-regulated in *o2o2o16o16wxwx*. Other up-regulated genes were *Zm00001d016198.1* encoding glutamate-oxaloacetic transaminase, which converts oxaloacetic acid to aspartic acid in the aspartic acid synthesis pathway; *Zm00001d035443.1* encoding acetylornithine transferase in the conversion of *N*-acetyl-L-glutamate to *N*-acetyl-L-ornithine in the arginine synthesis pathway; and *Zm00001d027536.1* encoding serine acetyltransferase to transform L-serine to *O*-acetyl-L-serine in the cysteine synthesis pathway. Two genes were down-regulated: *Zm00001d025862.1* encoding threonine aldolase in the conversion of threonine into glycine and *Zm00001d027861.1* encoding serine-pyruvate aminotransferase in the conversion of glyoxalic acid to glycine (Table 1).

In regards to amino acid degradation pathways, one gene was up-regulated: *Zm00001d037498.1* (*tar*) that encodes tryptophan aminotransferase and participates in the conversion of tryptophan to indole-3 pyruvate in tryptophan metabolism. The remaining four genes were down-regulated. *Zm00001d047124.1* encodes proline dehydrogenase for the conversion of proline to (*S*)-1-pyrroline-5-carboxylate in the proline degradation pathway. *Zm00001d049380.1* encodes 4-aminobutyrate transaminase and participates in the conversion of 4-aminobutanoate to succinate semialdehyde in the glutamate degradation pathway. *Zm00001d052079.1* encodes lysine-ketoglutarate reductase/saccharopine dehydrogenase (LKR/SDH1) (EC1.5.1.8) and thereby transforms DL-saccharopine to L-2-aminoglycolic acid-6-hemialdehydes in the lysine degradation metabolism pathway. Finally, *Zm00001d020984.1* encodes sarcosine oxidase (EC1.5.3.7/EC1.5.3.1) and participates in lysine degradation, which inhibits lysine degradation and thus increases lysine content.



**Figure 6.** Pathway analysis of differentially expressed genes. X axis means number of DEGs. Y axis represents second KEGG pathway terms. All second pathway terms are grouped in top pathway terms indicated in different color.

**Carbohydrate metabolism.** The 21 significantly differentially expressed genes involved in carbohydrate metabolism were mainly involved in starch/sucrose metabolism (ko00500), amino sugar/nucleotide sugar (ko00520), glycolysis/gluconeogenesis (ko00010), fructose/mannose metabolism (ko00051), and pentose/glucuronate interconversions (ko00040). Under starch and sucrose metabolism, genes *Zm00001d044129.1* (*shrunk2, sh2*) and *Zm00001d050032.1* (*brittle2, bt2*) encoding ADP-glucose pyrophosphorylase (EC2.7.7.27) in the conversion of  $\alpha$ -D-glucopyranose 1-phosphate to ADP- $\alpha$ -D-glucose in the starch synthesis pathway were up-regulated. Other up-regulated genes included *Zm00001d016684.1* encoding starch branching enzyme II (EC2.4.1.18) in the conversion of (1,4- $\alpha$ -D-glucosyl)<sub>(n+1)</sub> to  $\alpha$ (1,6)- $\alpha$ -D-glucosyl-(1,4)- $\alpha$ -glucan; *Zm00001d010801.1* encoding sucrose synthase (EC2.4.1.13) in the conversion of UDP-glucose to sucrose 6-phosphate in the sucrose synthesis pathway; *Zm00001d012433.1* encoding UDP-glucuronate decarboxylase (EC4.1.1.35) in the conversion of UDP-D-glucuronate to UDP-D-xylose in the xylose synthesis pathway; *Zm00001d015129.1* encoding galacturan 1,4- $\alpha$ -galacturonidase (EC3.2.1.67) in the conversion of pectate to D-galacturonate in the D-galacturonate synthesis pathway; and *Zm00001d021421.1* encoding UDP-glucose pyrophosphorylase (EC2.7.7.9/EC2.7.7.64) in the conversion of  $\alpha$ -D-glucose-1P to UDP-glucose in the sucrose synthesis pathway. The gene *Zm00001d048099.1* encoding  $\beta$ -glucosidase (EC3.2.1.21) in the conversion of glucoside to  $\alpha$ -D-glucose as well as the conversion of 1,4- $\beta$ -D-glucan and cellobiose to  $\beta$ -D-glucose in the  $\alpha$ -D-glucose synthesis pathway was down-regulated (Table 2).

**qRT-PCR validation.** The DEGs identified by RNA-seq were further validated by qRT-PCR. In this study, qRT-PCR (quantitative real-time polymerase chain reaction) verification was conducted for 17 DGEs, including 9 up-regulated and 8 down-regulated genes, involved in amino acid and carbohydrate metabolism (Supplementary Table S9 and Fig. 7A,B). The results showed that the expression patterns of these 17 genes were similar to those measured by transcriptome sequencing; the RNA-seq results showed high relevance to the qRT-PCR results (Person's  $r = 0.8409, 0.8377$ ) (Fig. 7C,D), indicating the reliability of the RNA-seq data.

## Discussion

In this study, we introgressed the *o2* and *o16* alleles into waxy maize line QCL5019 by using MABB technique, and two mutants (QCL8006\_1 and QCL8006\_2) (*o2o2o16o16wxwx*) were acquired. SNP (single nucleotide polymorphism) microarray analysis revealed that the recovery rate of the genetic background was over 95%; this level was higher than the theoretical genetic background recovery rate (87.50%), and nearly identical to that of the recurrent parent, indicated that the generated material was suitable for transcriptome analysis. On this basis, we analysed the transcription of the mutants (*o2o2o16o16wxwx*) and their recurrent parent at 18DAP by RNA-seq. We detected 272 DEGs in *o2o2o16o16wxwx* that were significantly differentially expressed relative to recurrent parent. These genes were mainly involved in cell, cell part, organelle, binding, catalytic activity, metabolic process, and cellular process. Among them, 75 genes were involved in metabolism pathways, including 21 participating in carbohydrate metabolism and 12 associated with amino acid metabolism.

Endosperm of *o2o2* and *o16o16* mutants is usually soft, fragile<sup>1,2,25</sup> and that of waxy mutant was dark, smooth and waxy. In our study, the *o2o2o16o16wxwx* mutants, QCL8006\_1 and QCL8006\_2, were found to share similar ear phenotype with the recurrent parent QCL5019, but their kernel coats were non-glossy and wrinkled. In addition, their kernels were completely opaque, farinaceous and not full, suggesting that the *o2* and *o16* alleles may play a crucial role in the waxy genetic background<sup>38</sup>. SEM revealed that starch granules of the *o2o2o16o16wxwx* mutants had an irregular shape and arrangement and an uneven volume and size, with a high density of matrix

Gene ID	log <sub>2</sub> Ratio	Description	Pathway
Zm00001d047124.1	-4.59	proline dehydrogenase EC1.5.5.2	proline metabolism
Zm00001d027861.1	-2.36	alanine-glyoxylate transaminase EC2.6.1.44	alanine biosynthesis
Zm00001d025862.1	-1.41	L-threonine aldolase EC4.1.2.5	alanine biosynthesis
Zm00001d014258.1	1.43	alanine transaminase EC2.6.1.2	alanine biosynthesis
Zm00001d030557.1	1.41	alanine transaminase EC2.6.1.2	alanine biosynthesis
Zm00001d049380.1	-1.13	4-aminobutyrate transaminase EC2.6.1.19	glutamate metabolism
Zm00001d035443.1	1.13	acetylornithine transaminase EC2.6.1.11	L-arginine biosynthesis
Zm00001d016198.1	1.41	glutamate-oxaloacetic transaminase EC2.6.1.1	L-aspartate biosynthesis
Zm00001d027536.1	1.98	serine O-acetyltransferase EC2.3.1.30	L-cysteine biosynthesis I
Zm00001d052079.1	-2.88	lysine-ketoglutarate reductase EC1.5.1.8	Lysine degradation
Zm00001d020984.1	-2.28	sarcosine oxidase EC1.5.3.7/EC1.5.3.1	Lysine degradation
Zm00001d037498.1	1.73	tryptophan aminotransferase related EC2.6.1.99	Tryptophan metabolism

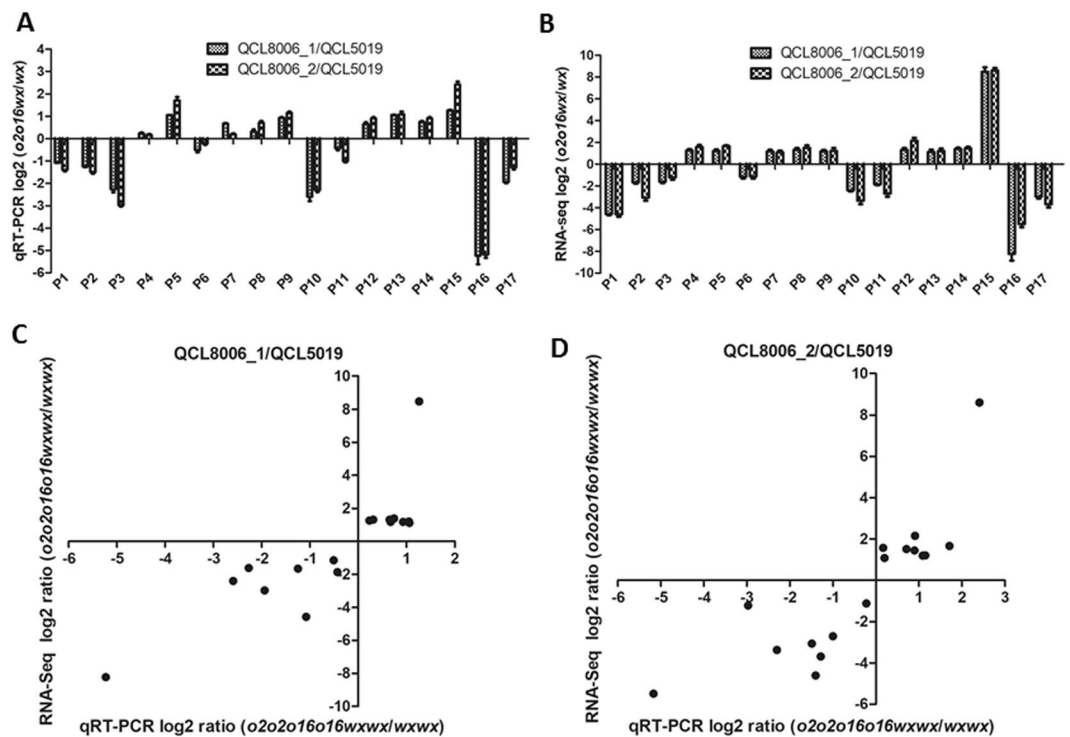
**Table 1.** Twelve DEGs involved in amino acid metabolism.

Gene ID	log <sub>2</sub> Ratio	Description	Pathway
Zm00001d018966.1	-3.23	chitinase EC3.2.1.14	Amino sugar and nucleotide sugar metabolism
Zm00001d025753.1	1.55	chitinase EC3.2.1.14	Amino sugar and nucleotide sugar metabolism
Zm00001d049380.1	-1.13	4-aminobutyrate-pyruvate transaminase EC2.6.1.96	Butanoate metabolism
Zm00001d031727.1	1.56	L-iditol 2-dehydrogenase EC1.1.1.14	Fructose and mannose metabolism
Zm00001d037278.1	1.36	diphosphate-fructose-6-phosphate 1-phosphotransferase EC2.7.1.90	Fructose and mannose metabolism
Zm00001d044754.1	1.27	diphosphate-fructose-6-phosphate 1-phosphotransferase EC2.7.1.90	Fructose and mannose metabolism
Zm00001d001999.1	1.26	glucose-6-phosphate 1-epimerase EC5.1.3.15	Glycolysis/Gluconeogenesis
Zm00001d024575.1	-6.86	glyceraldehyde-3-phosphate dehydrogenase EC1.2.1.12	Glycolysis/Gluconeogenesis
Zm00001d051001.1	2.51	glyceraldehyde-3-phosphate dehydrogenase EC1.2.1.12	Glycolysis/Gluconeogenesis
Zm00001d047893.1	1.34	phosphoenolpyruvate arboxykinase EC4.1.1.49	Glycolysis/Gluconeogenesis
Zm00001d008318.1	1.49	phosphoenolpyruvate arboxykinase EC4.1.1.49	Glycolysis/Gluconeogenesis
Zm00001d045431.1	1.67	phosphopyruvate hydratase EC4.2.1.11	Glycolysis/Gluconeogenesis
Zm00001d046234.1	-3.33	inositol oxygenase EC1.13.99.1	Inositol phosphate metabolism
Zm00001d044129.1 ( <i>sh2-shrunken2</i> )	1.42	adp glucose pyrophosphorylase, EC2.7.7.27	Starch and sucrose metabolism
Zm00001d050032.1 ( <i>bt2-brittle endosperm2</i> )	1.66	adp glucose pyrophosphorylase, EC2.7.7.27	Starch and sucrose metabolism
Zm00001d048099.1	-1.7	beta-glucosidase EC3.2.1.21	Starch and sucrose metabolism
Zm00001d015129.1	1.59	galacturan 1,4-alpha-galacturonidase EC3.2.1.67	Starch and sucrose metabolism
Zm00001d016684.1 ( <i>ae1-amylose extender1</i> )	1.16	starch branching enzyme II EC 2.4.1.18	Starch and sucrose metabolism
Zm00001d010801.1	8.55	sucrose synthase EC2.4.1.13	Starch and sucrose metabolism
Zm00001d012433.1	1.74	UDP-glucuronate decarboxylase EC4.1.1.35	Starch and sucrose metabolism
Zm00001d021421.1 ( <i>ugp2</i> )	1.31	ugp2-UDP-glucose pyrophosphorylase EC2.7.7.9, 2.7.7.64	Starch and sucrose metabolism

**Table 2.** Twenty-one DEGs involved in carbohydrate metabolism.

proteins that dispersed in the gap between starch granules. Normally, maize starch granules were mostly ellipsoid or spherical, and tightly packed with protein bodies. However, in *o2o2o16o16wxwx* mutants, the shape and arrangement of starch granules were irregular and uneven, which may account for the wrinkling and fragile phenotype. Meanwhile, the protein matrix accumulates together rather than tightly surrounding the starch granules, may contribute to farinaceous endosperm.

In the present study, the total protein contents of QCL8006\_1 and QCL8006\_2 were reduced by over 4%, compared to QCL5019, which is consistent with previous studies in *o2o2o7o7* mutant<sup>42</sup> and *o2o2wxwx* mutant<sup>41</sup>. However, the total starch content was not significantly changed between QCL8006\_1, QCL8006\_2 and QCL5019, which is inconsistent with the previous study<sup>41</sup>. In addition, to varying extents, the contents of free amino acids Glu, Leu, Ser, Ala, Ile, Tyr and Phe were lower in the *o2o2o16o16wxwx* lines than in the recurrent parent (*wxwx*), a result consistent with changes in amino acids due to introgression of the *o2* allele into *wxwx* maize observed by Zhou *et al.*<sup>41</sup>. In contrast, Lys, Cys, Arg and Gly contents increased significantly, especially Lys and Gly, this result was different from the observations of Zhou *et al.*<sup>41</sup>, who analysed *o2o2wxwx* mutants containing two recessive genes. The above differences may be related to the introgression of the high-lysine gene *o16* in our study and the different genetic backgrounds. RNA-seq analysis of DEGs uncovered 527 differentially expressed genes (448 up-regulated and 79 down-regulated) presented in QCL8006\_2 but not in QCL8006\_1. This disparity may be due to different genetic recombination and genetic background recovery rates. For example, the transcriptional profiles of *o2* and wild-type lines with different genetic backgrounds are different<sup>40</sup>. In the present study, we



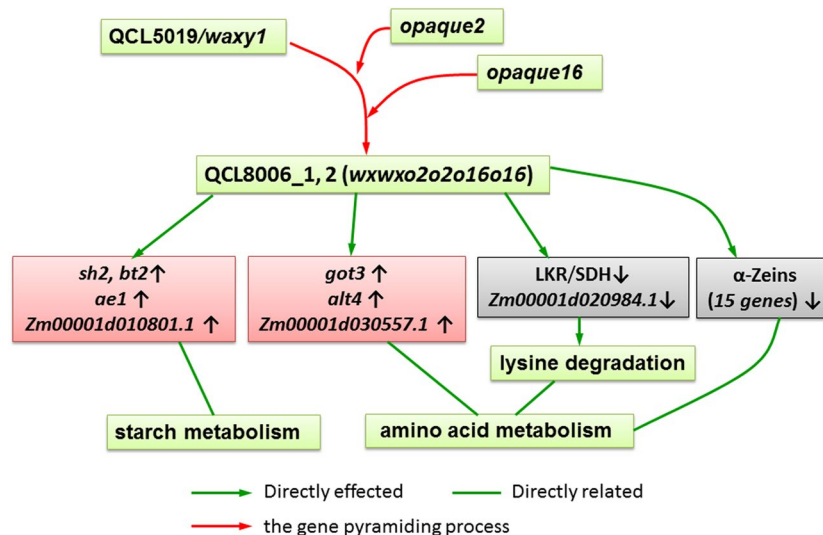
**Figure 7.** (A) The qRT-PCR log<sub>2</sub> ratio of 17 DEGs for QCL5019 vs. QCL8006\_1 and QCL5019 vs. QCL8006\_2. (B) The RNA-Seq log<sub>2</sub> ratio of 17 DEGs for QCL5019 vs. QCL8006\_1 and QCL5019 vs. QCL8006\_2. (C) The qRT-PCR validation of 17 DEGs for QCL5019 vs. QCL8006\_1 identified by RNA-seq. Pearson's  $r = 0.8409$ . (D) The qRT-PCR validation of 17 DEGs for QCL5019 vs. QCL8006\_2 identified by RNA-seq. Pearson's  $r = 0.8377$ . P1-P17, 17 DEGs.

filtered out 272 differentially expressed genes shared between QCL8006\_1 and QCL8006\_2, which minimized the influence of such differences due to different genetic backgrounds and recombination and improved the efficiency of differentially expressed gene identification. Zhan *et al.*<sup>44</sup> discovered that 1863 DEGs were detected in the endosperm of B73O2 versus B73o2 using RNA-seq of 15DAP endosperm, including 1024 DEGs that were downregulated and 839 DEGs upregulated in the *o2* mutant, and the result also showed the 494 DEGs that have been detected by at least one prior study<sup>45–48</sup>. Compared with those 494 DEGs, 66 DEGs were detected in the *o2o2o16o16wxwx* mutants versus *wxwx* parent, including 32 upregulated DEGs and 34 downregulated DEGs.

Lysine ketoglutarate reductase/saccharopine dehydrogenase (LKR/SDH) is a bifunctional lysine-degrading enzyme and a major regulator of free lysine content in plants. Kemper *et al.*<sup>49</sup> found that the transcription level of LKR/SDH mRNA in maize *o2* mutants was decreased by more than 90% and that enzyme activity was significantly decreased, which reduced the degradation of lysine. Kawakatsu *et al.*<sup>50</sup> reported that bifunctional rice OsLKR/SDH exists mainly in seeds and is directly regulated by the transcription regulatory factors RISBZ1 (rice seed *b*-Zipper 1) and RPBf (rice prolamin box binding factor) of seed storage protein genes; these authors also observed that decreases in RISBZ1 or RPBf lead to a decrease in OsLKR/SDH expression levels and an increase in the free lysine content of rice grains. In the present study, consistent with the above observations, *Zm00001d052079.1* encoding LKR/SDH and *Zm00001d020984.1* encoding sarcosine oxidase in the lysine degradation pathway were down-regulated, thereby inhibiting lysine degradation and increasing the grain lysine content of the *o2o2o16o16wxwx* lines.

Hartings *et al.*<sup>42</sup> reported that transcription levels of 10 kDa  $\gamma$ -zein, 19 kDa and 22 kDa  $\alpha$ -zein, and 27 kDa and 50 kDa  $\gamma$ -zein are significantly decreased in *o2o2o7o7* lines. Li *et al.*<sup>51</sup> discovered that O2 not only directly binds to the promoters of known targets (22 kDa  $\alpha$ -zein, 19 kDa  $\alpha$ -zein and 14 kDa  $\beta$ -zein genes), but also to other zein genes, except for the 16 and 18 kDa zeins. Zhan *et al.*<sup>44</sup> discovered that O2 directly regulates 23 zein genes, which were 13 genes of 19 kDa  $\alpha$ -zein, 8 genes of 22 kDa  $\alpha$ -zein, each one gene of 15 kDa  $\beta$ -zein, 18 kDa  $\delta$ -zein, 27 kDa  $\gamma$ -zein and 50 kDa  $\gamma$ -zein. In our study, we obtained similar results, where 15 genes encoding  $\alpha$ -zein were down-regulated in *o2o2o16o16wxwx* lines compared with *wxwx* parent (Supplementary Table S10), and of them 5  $\alpha$ -zein genes were detected by Li and Zhan *et al.*<sup>44,51</sup>. Therefore, the increased level of lysine in *o2o2o16o16wxwx* lines largely depends on the reduction of  $\alpha$ -zein synthesis which excludes lysine, while several non-zein proteins which accounts for most of the higher percentage of lysine are associated with varying degrees of increased accumulation<sup>49</sup>. In the *o2o2o16o16wxwx* lines, *sh2* and *bt2* encoding adenosine diphosphate glucose pyrophosphorylase, and *ae1* encoding starch-branching enzyme were all upregulated, agreeing with data from Zhou *et al.*<sup>41</sup>. These changes are likely responsible for the formation of their wrinkling kernel and farinaceous endosperm, because of *sh2sh2* mutant with shrunken kernels, *bt2bt2* mutant with brittle kernels and *ae1ae1* mutant with wrinkled kernels.





**Figure 8.** A proposed model of the regulatory network of waxy corn following the introgression of the *o2* and *o16* alleles. MAS, Marker-assisted selection; FS, Foreground selection; BS, Background selection; RNA-seq, RNA sequencing; ↑, -up-regulated; ↓, down-regulated.

Hidehiko *et al.*<sup>52</sup> cloned the cDNA of alanine aminotransferase (AlaAT) from mature rice seeds and confirmed that AlaAT is involved in rice-seed nitrogen metabolism and storage protein synthesis. In addition, Shrawat *et al.*<sup>53</sup> has shown that increased levels of alanine transferase can increase crop yields. In the present study, *alt4* and *Zm00001d030557.1* encoding alanine aminotransferase and *got3* encoding glutamic oxaloacetic transaminase were up-regulated in *o2o16o16wxwx* lines, which indicates that the introduction of *o2* and *o16* alleles into *wx* maize can increase the expression of alanine transferase (ALT) and glutamate oxalate transaminase (GOT) genes, promote amino acid transformation, facilitate dry matter accumulation and alleviate the contradiction between quality and yield.

In this study, the key genes responsible for the increase in lysine content and formation of opaque endosperm after introgression of *o2* and *o16* alleles into waxy maize were identified at the transcriptional level. Compared with the *wxwx* parent, in the *o2o16o16wxwx* lines, 15 genes encoding  $\alpha$ -zein were down-regulated, which resulted in the reduction of  $\alpha$ -zein synthesis and correspondingly in the increment of non-zein, thereby increased lysine content. In addition, *lkr/sdh1* and *Zm00001d020984.1* genes involved in the lysine degradation pathway were down-regulated, thereby inhibited lysine degradation. Furthermore, *sh2*, *bt2* and *ae1* genes involved in starch metabolism were upregulated, leading to wrinkling kernel and farinaceous endosperm (Fig. 8).

## Conclusions

In this study, the *o2* and *o16* alleles were introgressed into the waxy corn using MABB, and QCL8006\_1 and QCL8006\_2 were acquired. Lysine contents of QCL8006\_1 and QCL8006\_2 were improved compared with that of recurrent parent. In the lines of QCL8006\_1 and QCL8006\_2, 272 genes were differentially expressed relative to recurrent parent by RNA-seq. GO and KEGG enrichment analyses revealed that these genes were mainly related to biomass metabolism. Among them, 12 genes were enriched for amino acid metabolic pathways, including down-regulated *lkr/sdh1* and *Zm00001d020984.1*, which inhibited the degradation of lysine. In addition, 15 genes encoding  $\alpha$ -zein were down-regulated, which resulted in the reduction of  $\alpha$ -zein synthesis, while several non-zein proteins which accounts for most of the higher percentage of lysine are associated with varying degrees of increased accumulation. Furthermore, 21 genes were related to carbohydrate metabolism; these included *sh2*, *bt2* and *ae1* genes upregulated, led to wrinkling kernel and farinaceous endosperm. These will help to uncover the transcriptional regulation mechanism in lysine content increased by *o2* and *o16* allele introgression into waxy maize.

## Materials and Methods

**Materials.** A three-way  $F_1$  cross was performed using QCL3024 (*o16o16*) and Taixi19 (*o2o2*) as donor parents and QCL5019 (*wxwx*) as the recurrent parent. Using SSRs linked to *wx*, *o2* and *o16* as markers for *phi027*, *umc1066*, and *umc1121*, respectively, we selected individual plants with the *O2o2O16o16Wxwx* genotype from the  $F_1$  generation, and  $BC_1F_1$  plants were obtained by backcrossing with QCL5019. In the  $BC_1F_1$  generation, individuals with the genotype *O2o2O16o16wxwx* were selected, and  $BC_2F_1$  plants were obtained by backcrossing with QCL5019. Next,  $BC_2F_1$  generation individuals with the *O2o2O16o16wxwx* genotype were selected and then self-pollinated to obtain  $BC_2F_2$  plants. In the  $BC_2F_2$  generation, individual plants with the genotype *o2o2o16o16wxwx* were selected, and  $BC_2F_3$  individuals were obtained by self-pollination. After analysing genetic background, lysine and waxy quality, and plant field performance, we selected two *o2o2o16o16wxwx* gene-pyramiding lines, which were named QCL8006\_1 and QCL8006\_2, following more than 20 generations of self-pollination to genetic stability (Supplementary Fig. S4). Lysine and amylopectin contents of grains were respectively 0.49% and 0.54% in QCL8006\_1 and 99.87% and 99.25% in QCL8006\_2 (Inspection and Testing Centre for

Quality of Cereals and their Products, Ministry of Agriculture, Harbin, China). According to a 55 K SNP microarray analysis, recovery rates of the genomic genetic background of QCL8006\_1 and QCL8006\_2 were 95.22% and 95.12% respectively, higher than the theoretical background recovery rate (87.50%), because of selecting individual plants with high genetic background recovery rate using the whole-genome SSR markers in backcross generations, this result indicated that a comparative transcriptome analysis of *o2o2o16o16wxwx* mutant lines and the *wxwx* parent would be beneficial.

**Kernel characteristics and submicroscopic structure.** Grains of the *o2o2o16o16wxwx* mutant lines and their recurrent parent were selected for analysis. The grains were checked for the presence of a smooth vs. wrinkled surface under natural light, and grain transparency was observed using a light box. After grains were cut open with a blade, the grain cross-sections were observed to determine the degree of starchiness or waxiness and then photographed.

Maize kernels were decapped and sliced in the centre of the kernel with a razor blade. A small piece of endosperm was isolated from each kernel, coated with platinum using an E-1010 ion sputterer, and observed by scanning electron microscopy. The submicroscopic structure was observed by the Guizhou Key Laboratory of Agricultural Biotechnology Guiyang, China (N26°30'14" and E106°39'21").

**Analyses of amino acid, protein and starch contents.** Amino acid contents of mature maize grains were analysed using an automatic amino acid analyzer (InfracTM1241 Grain Analyzer, Made in Sweden). Approximately 120–150 mg dried powder of each sample was weighed and transferred to a small tube, and 5 mL of 6 M HCl was added. Each tube was incubated at 110 °C in a water bath for 24 h; after cooling to room temperature, the pH was adjusted to 2.0 with approximately 4.8 mL of 6 M NaOH, and the solution was diluted to a volume to 100 mL using ddH<sub>2</sub>O. Each sample was then filtered through a 0.45 µm membrane into a 2 mL liquid chromatographic sample bottle and used for determination of the contents of 17 FAAs.

Protein and starch contents of maize mature grains were assayed by using near infrared analyzers. Three replicates of each sample were performed.

**RNA-seq library construction and sequencing.** Eighteen days after pollination, total RNA of whole grains was isolated using a plant RNA kit (Omega) according to the “difficult sample” protocol of the manufacturer. The concentration and quality of each RNA sample was checked using a NanoDrop 1000 instrument. After rRNA removal and enrichment of mRNA using Oligo(dT) magnetic beads, mRNA was reverse transcribed into cDNA using random primer N6; this was followed by cDNA second-strand synthesis to form double-stranded DNA. Following adapter fusion, the fragments were amplified using specific primers by PCR. The PCR products were denatured into single strands, and a single-stranded DNA library was obtained by cyclization of the single-stranded DNA with a bridge primer. Transcriptome sequencing (single end 50 bp, SE50) was carried out on a BGISEQ-500 sequencing platform by Shenzhen Huada Gene Technology Co. Three replicates of each sample were sequenced.

**Identification of differentially expressed genes.** Low-quality reads and reads containing adapters or having an unknown base percentage > 10% were removed from the original sequencing data. The resulting clean reads were aligned to the reference genome (B73 version4)<sup>54</sup> using HISAT2<sup>55</sup> with the following parameters: -p 8-phred64-sensitive -I 1 -X 1000, and aligned to reference gene using Bowtie2<sup>56</sup> with the following parameters: -q-phred64-sensitive-dpad 0-gbar 99999999-mp 1,1-np 1-score-min L,0,-0.1 -p 16 -k 200, and to calculate the gene alignment rate. FPKM was calculated as the expression level of genes and transcripts by using RSEM software<sup>57</sup>. Differentially expressed genes were identified using NOISeq<sup>58</sup> according to the following criteria: fold change  $\geq 2$  and corrected  $P \leq 0.05$ .

**GO and pathway enrichment analysis of DEGs.** Significant differentially expressed genes were mapped to the terms of the Gene Ontology database (<http://www.geneontology.org/>). The gene number of each GO term was calculated, and terms that were significantly enriched in differentially expressed genes were identified by the hypergenometric test using WEGO software<sup>43</sup>. The calculated  $P$  values were subjected to Bonferroni<sup>59</sup> correction, and GO terms enriched in differentially expressed genes were identified based on a  $P \leq 0.05$ . KEGG<sup>60</sup> pathway enrichment was calculated in the same way as in the GO functional enrichment analysis. In this study, pathways, with a  $P$  value  $\leq 0.05$ , were considered to be significantly enriched in differentially expressed genes.

**qRT-PCR validation.** Seventeen DEGs were selected for qRT-PCR verification. The primers, shown in Table S7, were designed online (<http://www.primer3plus.com/cgi-bin/dev/primer3plus.cgi>), cDNA was synthesized using a reverse transcription kit (Thermo Scientific RevertAid First Strand cDNA Synthesis Kit, K1622) in a 20 µL reaction volume, and quantitative analysis was conducted on CFX Connect Real-Time PCR System (BIO-RAD, Hercules, CA) according to the method used by Liu *et al.*<sup>61</sup>. In a 10 µL reaction volume, using 2 µL of a twentyfold diluted cDNA solution, 5 µL of SYBR<sup>®</sup> Select Master Mix (ABI), 0.5 µL of each primer (10 mM) and 2 µL ddH<sub>2</sub>O. The thermal cycling conditions were 2 min at 50 °C, 10 min at 95 °C, followed by 39 cycles of 20 s at 95 °C and 1 min at 60 °C. Three replicates were used for each sample and actin was used as an internal standard; the relative expression level was calculated using the  $2^{-\Delta\Delta CT}$  method.

**Accession codes.** The raw data of RNA-seq reads were deposited in the National Center for Biotechnology Information (NCBI) database under accession number (PRJNA 512329), biosample accessions were as follows: SAMN10666504, SAMN10666505, SAMN10666506, SAMN10666507, SAMN10666508, SAMN10666509, SAMN10666510, SAMN10666511 and SAMN10666512.

## References

- Gavazzi, F., Lazzari, B., Ciceri, P., Gianazza, E. & Viotti, A. Wild-type Opaque2 and defective opaque2 polypeptides form complexes in maize endosperm cells and bind the *opaque2*-zein target site. *Plant Physiol* **145**, 933–945 (2007).
- Motto, M., Thompson, R. & Salamini, F. Genetic regulation of carbohydrate and protein accumulation in seeds. *Cellular and molecular biology of plant seed development*. Springer, Dordrecht, 479–522 (1997).
- Mains, E. B. Heritable characters in maize: Linkage of a factor for shrunken endosperm with the *a1* factor for aleurone color. *J Hered* **40**, 21–24 (1949).
- Teas, H. J. & Teas, A. N. Heritable characters in maize: Description and linkage of brittle endosperm-2. *J Hered* **44**, 156–158 (1953).
- Holding, D. R. *et al.* The maize *floury1* gene encodes a novel endoplasmic reticulum protein involved in zein protein body formation. *The Plant Cell* **19**, 2569–2582 (2007).
- Robertson, D. S. A new opaque gene located on chromosome 7. *Maize Genet Coop News Lett* **41**, 94–95 (1967).
- Myers, A. M. *et al.* Maize *opaque5* encodes monogalactosyldiacylglycerol synthase and specifically affects galactolipids necessary for amyloplast and chloroplast function. *The Plant Cell* **23**, 2331–2347 (2011).
- Ashman, R. B. Gene linkages in translocation T9-10a heterozygotes. *Maize Genet Coop News Lett* **42**, 155–156 (1968).
- Ma, Y. & Nelson, O. E. Amino acid composition and storage proteins in two new high-lysine mutants in maize. *Cereal Chem* **52**, 412–419 (1975).
- Wang, G. *et al.* Proline responding1 plays a critical role in regulating general protein synthesis and the cell cycle in maize. *The Plant Cell* **26**, 2582–2600 (2014).
- McWhirter, K. S. A floury endosperm, high lysine locus on chromosome 10. *Maize Genet Coop News Lett* **45**, 184 (1971).
- Misra, P. S. *et al.* Endosperm protein synthesis in maize mutants with increased lysine content. *Science* **176**, 1425–1426 (1972).
- Wang, G. *et al.* *Opaque7* encodes an acyl-activating enzyme-like protein that affects storage protein synthesis in maize endosperm. *Genetics* **189**, 1281–1295 (2011).
- Miclaus, M., Wu, Y., Xu, J., Dooner, H. & Messing, J. The maize high-lysine mutant *opaque7* is defective in an acyl-CoA synthetase-like protein. *Genetics* **189**, 1271–1280 (2011).
- Richardson, D. L. S.-floury a gene affecting protein synthesis. *Maize Genet Coop News Lett* **29**, 45 (1955).
- Pirone, R., Hartings, H., Lauria, M., Rossi, V. & Motto, M. Genetic control of endosperm development and of storage products accumulation in maize seeds. *Maydica* **50**, 515–530 (2005).
- Dannenhoffer, J. M., Bostwick, D. E., Or, E. & Larkins, B. A. *Opaque-15*, a maize mutation with properties of a defective *opaque-2* modifier. *Proc Natl Acad Sci USA* **92**, 1931–1935 (1995).
- Salamini, F. *et al.* *Mucronate*, *Mc*, a dominant gene of maize which interacts with *opaque-2* to suppress zein synthesis. *Theor Appl Genet* **65**, 123–128 (1983).
- Nelson, O. E., Mertz, E. T. & Bates, L. S. Second mutant gene affecting the amino acid pattern of maize endosperm proteins. *Science* **150**, 1469–1470 (1965).
- Coleman, C. E. *et al.* Expression of a mutant alpha-zein creates the *floury2* phenotype in transgenic maize. *Proc Natl Acad Sci USA* **94**, 7094–7097 (1997).
- Nelson, O. E. Jr. The location of *fl3* on chromosome 8. *Maize Genet Coop News Lett* **50**, 114 (1976).
- Li, Q. *et al.* The Maize Imprinted Gene *Floury3* Encodes a PLATZ Protein Required for tRNA and 5S rRNA Transcription through Interaction with RNA Polymerase III. *The Plant Cell* **29**, 2661–2675 (2017).
- Huang, S. *et al.* Improving nutritional quality of maize proteins by expressing sense and antisense zein genes. *J Agric Food Chem* **52**, 1958–1964 (2004).
- Gibbon, B. C. & Larkins, B. A. Molecular genetic approaches to developing quality protein maize. *Trends Genet* **21**, 227–233 (2005).
- Yang, W. P., Zheng, Y. L., Zheng, W. T. & Feng, R. Molecular genetic mapping of a high-lysine mutant gene (*opaque-16*) and the double recessive effect with *opaque-2* in maize. *Mol Breeding* **15**, 257–269 (2005).
- Sarika, K. *et al.* Exploration of novel *opaque16* mutation as a source for high-lysine and-tryptophan in maize endosperm. *Indian J Genet* **77**, 59–64 (2017).
- Sarika, K. *et al.* *Opaque16*, a high lysine and tryptophan mutant, does not influence the key physico-biochemical characteristics in maize kernel. *PLoS ONE* **13**, e0190945 (2018).
- Sarika, K. *et al.* Marker-assisted pyramiding of *opaque2* and novel *opaque16* genes for further enrichment of lysine and tryptophan in sub-tropical maize. *Plant Sci* **272**, 142–152 (2018).
- Tian, Q. Z., Li, X. H., Li, M. S., Jiang, W. & Zhang, S. H. Molecular markers assisted selection to quality protein maize. *J Maize Sci* **12**, 108–110 (2004).
- Zhang, W. L. *et al.* Molecular marker-assisted selection for *o2* introgression lines with *o16* gene in corn. *Acta Agronomica Sinica* **36**, 1302–1309 (2010).
- Andrés, J. M. & Bascialli, P. C. Caracteres hereditarios aislados en maíces cultivados en la Argentina. *Univ Buenos Aires Inst Genet* **2**, 1 (1941).
- Nelson, O. E. & Rines, H. W. The enzymatic deficiency in the waxy mutation of maize. *Biochem Bioph Reseco* **9**, 297–300 (1962).
- Sprague, G. F., Brimhall, B. & Nixon, R. M. Some effects of the waxy gene in corn on properties of the endosperm starch. *J Am Soc Agron* **35**, 817–822 (1943).
- Yang, Y. F., Guo, Q., Cheng, J., Zheng, X. Y. & Lin, C. M. Analysis of genetic and quality traits of waxy corn inbred lines in china temperate zone. *Acta Botan Boreali-Occident Sin* **29**, 2213–2220 (2009).
- Zhang, S. K., Teng, H. S., Su, Q. & Yang, Y. J. Selection of waxy-QPM maize lines with SSR. markers. *Guangxi Agric Sci* **10**, 1279–1283 (2009).
- Zhang, X. X. *et al.* Construction of waxy maize *opaque2* near-isogenic lines. *Acta Agronomica Sinica* **43**, 1760–1766 (2017).
- Yang, L. Q., Wang, W., Yang, W. P. & Wang, M. Marker-assisted selection for pyramiding the waxy and *opaque-16* genes in maize using cross and backcross schemes. *Mol Breeding* **31**, 767–775 (2013).
- Zhang, W. L. *et al.* Increasing lysine content of waxy maize through introgression of *opaque-2* and *opaque-16* genes using molecular assisted and biochemical development. *PLoS ONE* **8**, e56227 (2012).
- Dang, N. C., Munsch, M., Aulinger, I., Renlai, W. & Stamp, P. Inducer line generated double haploid seeds for combined waxy and *opaque 2* grain quality in subtropical maize (*Zea mays* L.). *Euphytica* **183**, 153–160 (2012).
- Jia, H. *et al.* Comparison of transcript profiles in wild-type and *o2* maize endosperm in different genetic backgrounds. *Crop Sci* **47**, S45–S59 (2007).
- Zhou, Z. Q. *et al.* Introgression of *opaque2* into waxy maize causes extensive biochemical and proteomic changes in endosperm. *PLoS ONE* **11**, e0158971 (2016).
- Hartings, H., Lauria, M., Lazzaroni, N., Pirone, R. & Motto, M. The *Zea mays* mutants *opaque-2* and *opaque-7* disclose extensive changes in endosperm metabolism as revealed by protein, amino acid, and transcriptome-wide analyses. *BMC Genomics* **12**, 41 (2011).
- Ye, J. *et al.* WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res* **34**, W293–W297 (2006).
- Zhan, J. *et al.* *Opaque-2* regulates a complex gene network associated with cell differentiation and storage functions of maize endosperm. *The Plant Cell* **30**, 2425–2446 (2018).
- Hunter, B. G. *et al.* Maize opaque endosperm mutations create extensive changes in patterns of gene expression. *The Plant Cell* **14**, 2591–2612 (2002).

46. Frizzi, A. *et al.* Compositional and transcriptional analyses of reduced zein kernels derived from the *opaque2* mutation and RNAi suppression. *Plant Mol. Biol* **73**, 569–585 (2010).
47. Jia, M. *et al.* Identification and characterization of lysine-rich proteins and starch biosynthesis genes in the *opaque2* mutant by transcriptional and proteomic analysis. *BMC Plant Biol* **13**, 60 (2013).
48. Zhang, Z., Zheng, X., Yang, J., Messing, J. & Wu, Y. Maize endosperm-specific transcription factors O2 and PBF network the regulation of protein and starch synthesis. *Proc Natl Acad Sci USA* **113**, 10842–10847 (2016).
49. Kemper, E. L. *et al.* The role of *opaque2* in the control of lysine-degrading activities in developing maize endosperm. *The Plant Cell* **11**, 1981–1993 (1999).
50. Kawakatsu, T. & Takaiwa, F. Differences in transcriptional regulatory mechanisms functioning for free lysine content and seed storage protein accumulation in rice grain. *Plant Cell Physiol* **51**, 1964–1974 (2010).
51. Li, C. *et al.* Genome-wide characterization of cis-acting DNA targets reveals the transcriptional regulatory framework of Opaque2 in maize. *The Plant Cell* **27**, 532–545 (2015).
52. Kikuchi, H., Hirose, S., Toki, S., Akama, K. & Takaiwa, F. Molecular characterization of a gene for alanine aminotransferase from rice (*Oryza sativa*). *Plant Mol Biol* **39**, 149–159 (1999).
53. Shrawat, A. K., Carroll, R. T., DePauw, M., Taylor, G. J. & Good, A. G. Genetic engineering of improved nitrogen use efficiency in rice by the tissue-specific expression of alanine aminotransferase. *Plant Biotechnol J* **6**, 722–732 (2008).
54. Jiao, Y. *et al.* Improved maize reference genome with single-molecule technologies. *Nature* **546**, 524 (2017).
55. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357 (2015).
56. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).
57. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
58. Tarazona, S., García-Alcalde, F., Dopazo, J., Ferrer, A. & Conesa, A. Differential expression in RNA-seq: A matter of depth. *Genome Res* **21**, 2213–2223 (2011).
59. Abdi, H. The bonferroni and Sidak corrections for multiple comparisons. *Encyclopedia Meas Stat* **3**, 103–107 (2007).
60. Kanehisa, M. *et al.* KEGG for linking genomes to life and the environment. *Nucleic Acids Res* **36**, 480–484 (2008).
61. Liu, Y. B., Qin, L. J., Han, L. Z., Xiang, Y. & Zhao, D. G. Overexpression of maize *SDD1* (*ZmSDD1*) improves drought resistance in *Zea mays* L. by reducing stomatal density. *The Plant Cell* **122**, 147–159 (2015).

## Acknowledgements

This research was supported by grants from the National Natural Science Foundation of China (No. 31160307), the Genetically Breeding Major Project of the Ministry of Agriculture of China (No. 2016ZX08010-003), the China Agriculture Research System (No. CARS-02-082), the Research project in Guizhou Province Science and Technology department (Nos MSP20136022, SE20184005, CSL20184003 and S20192299), Guizhou Top Level Innovation Talents Cultivation Project (Nos 20154017 and 20164003), Talent Base for Germplasm Resources Utilization and Innovation of Characteristic Plant in Guizhou (No: RCJD2018-14), Natural Science Foundation of Guizhou (No. 20191451), the Research project in Guizhou Academy of Agricultural Sciences (No. S2014008) and the Open Project of State Key Laboratory of Crop Genetics and Germplasm Enhancement (No. ZW201907).

## Author Contributions

W.Y. and D.Z. -conceived and designed the experiments; W.W., S.N., Y.D. and M.W. -performed the experiments and analyzed the data; W.Y., W.W. and Y.L. -contributed reagents/analysis tools; W.W. and W.Y. -wrote the paper.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-48478-6>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019