

RESEARCH ARTICLE

Exploring the basis of 2-propenyl and 3-butenyl glucosinolate synthesis by QTL mapping and RNA-sequencing in *Brassica juncea*Aimal Nawaz Khattak[☉], Tianya Wang[☉], Kunjiang Yu, Renqin Yang, Wei Wan, Botao Ye, Entang Tian[✉]*

Oil Crops Research Institute of Guizhou University, Agricultural College of Guizhou University, Guizhou University, Guiyang, China

☉ These authors contributed equally to this work.

* erictian121@163.com

OPEN ACCESS

Citation: Khattak AN, Wang T, Yu K, Yang R, Wan W, Ye B, et al. (2019) Exploring the basis of 2-propenyl and 3-butenyl glucosinolate synthesis by QTL mapping and RNA-sequencing in *Brassica juncea*. PLoS ONE 14(10): e0220597. <https://doi.org/10.1371/journal.pone.0220597>

Editor: Maoteng Li, Huazhong University of Science and Technology, CHINA

Received: July 17, 2019

Accepted: October 4, 2019

Published: October 18, 2019

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0220597>

Copyright: © 2019 Khattak et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Abstract

Brassica juncea is used as a condiment, as vegetables and as an oilseed crop, especially in semiarid areas. In the present study, we constructed a genetic map using one recombinant inbred line (RIL) of *B. juncea*. A total of 304 ILP (intron length polymorphism) markers were mapped to 18 linkage groups designated LG01-LG18 in *B. juncea*. The constructed map covered a total genetic length of 1671.13 cM with an average marker interval of 5.50 cM. The QTLs for 2-propenyl glucosinolates (GSLs) colocalized with the QTLs for 3-butenyl GSLs between At1g26180 and BnapPIP1580 on LG08 in the field experiments of 2016 and 2017. These QTLs accounted for an average of 42.3% and 42.6% phenotypic variation for 2-propenyl and 3-butenyl GSLs, respectively. Furthermore, the Illumina RNA-sequencing technique was used to excavate the genes responsible for the synthesis of GSLs in the siliques of the parental lines of the RIL mapping population, because the bulk of the seed GSLs might originate from the siliques. Comparative analysis and annotation by gene ontology (GO) and kyoto encyclopedia of genes and genomes (KEGG) revealed that 324 genes were involved in GSL metabolism, among which only 24 transcripts were differentially expressed genes (DEGs). Among those DEGs, 15 genes were involved in the biosynthesis and transport of aliphatic GSLs, and their expression patterns were further validated by qRT-PCR analysis. Joint QTL mapping and RNA-sequencing analyses reveal one candidate gene of *ILL1* (LOC106416451) for GSL metabolism in *B. juncea*. These results will be helpful for further fine mapping, gene cloning and genetic mechanisms of 2-propenyl and 3-butenyl GSLs in *B. juncea*.

Introduction

Brassica juncea (AABB, $2n = 36$) is an important allotetraploid species that originated from interspecific hybridization between *B. rapa* (AA, $2n = 20$) and *B. nigra* (BB, $2n = 16$) followed

Funding: This work was funded by the National Natural Science Foundation of China (Grant No. 31560422), Agricultural Science and Technology Support Program of Guizhou Province (Qiankehe zhicheng No. [2019]2396), Science and Technology Foundation of Guizhou Province of China (Grant No. Qiankehe J zi [2015]2052), Scientific Research Foundation for Returned Scholars, Ministry of Education of China (Grant No. Jiaowaisiliu [2015]1098), Foundation of Guizhou University (Grant No. Guidarenjihezi [2014]14), Construction, Construction Program of Biology First-class Discipline in Guizhou (Grant No. GNYL [2017]009). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

by chromosome doubling in the natural environment. The crop exhibited better drought and heat tolerance, disease resistance, insect resistance and shattering resistance than *B. napus* [1–7]. In addition to its use as a condiment in Canada and China and as a vegetable in China, great efforts have been made to develop *B. juncea* as an alternative oilseed crop, especially in the semiarid areas. Canola-quality *B. juncea* with less than 2% erucic acid in the seed oil and less than 30- μ M GSLs/g of the deoiled cake was developed through a cross between a zero erucic *B. juncea* line and a low GSLs *B. juncea* line [8]. Hybrid breeding has been successfully used to enhance the yield potential in canola *B. napus*. In *B. juncea*, the Ogura cytoplasmic male sterility (*cms*) and its restorer gene (*Rfo*) were introduced for heterosis utilization [9,10]. Then, the *B. juncea* Ogura *cms* restorer line (*RfoRfo*) was improved with drastically reduced linkage drag, good seed set and high agronomic performance by hybridization with resynthesized *B. juncea* lines and subsequent molecular marker-assisted selection in *B. juncea* [11].

Glucosinolates (GSLs) were first discovered in mustard seeds during an exploration of the chemical origin of their sharp taste in the 17th century. To date at least 120 different GSLs have been identified in sixteen families of dicotyledonous angiosperms [12]. In the Brassicaceae family, GSLs are the major secondary metabolites and could be synthesized in all species of this family by a three-part biosynthetic pathway from methionine, tryptophan and phenylalanine [13–15]. GSLs are mainly divided into aliphatic, indolic and benzyl GSLs in *Brassica* species. Most of the tissues, such as rosette leaves, roots, seeds, inflorescences, contain GSLs. The GSL contents of different tissues are not entirely synthesized locally. The transport of GSLs was suggested more than 40 years ago through a number of studies that indicated that GSLs are produced in maternal tissue and subsequently transported to the seed [16,17]. This could explain why the GSL profiles of hybrid seeds were similar to those of the maternal plants instead of being intermediate between those of the maternal and paternal plants [16]. As the closest organ to the seeds, the silique is the only organ to produce all the GSLs found in the seed [17–19]. Thus, the siliques might be the most important source of seed GSLs.

In seeds of the Brassicaceae, aliphatic GSLs were the major type of GSLs, while 2-propenyl and 3-butenyl GSLs were the major types of aliphatic GSLs. Previous studies have shown that the European type of *B. juncea* mainly contains 2-propenyl GSLs and Indian types contain both 2-propenyl and 3-butenyl GSLs [20–22]. The inheritance of 2-propenyl and 3-butenyl GSLs was studied in F_1 , F_2 and backcrossing populations of *B. juncea*, indicating that these two characteristics had maternal effects and might be controlled by multiple additive alleles at the same loci [21]. Using bulked segregant analysis, one ISSR marker was found to be tightly linked to high 2-propenyl GSLs in *B. juncea* and converted to a SCAR marker [23]. A total of 17 metabolic QTLs for the genetic control of 2-propenyl and 3-butenyl GSLs were identified on LG2, LG3, LG4, LG5, LG6, LG7, LG8 and LG9 in *B. oleracea*, among which 12 regulated 2-propenyl and 3-butenyl GSLs at the same time, 2 were specific for 2-propenyl GSLs and 3 were specific for 3-butenyl GSLs [24]. In *B. napus* seeds, 3 QTLs on A9 and 1 QTL on C2 for the genetic regulation of 3-butenyl GSLs were detected, and each could explain a phenotypic variation between 5.0% and 14.8% [25].

ILP markers utilize variations in intron sequences and are the most easily recognizable type of marker, as they can be detected by PCR with primers designed for the exons flanking the target intron [26]. Furthermore, ILP markers are unique because they are gene-specific, codominant, hypervariable, neutral, convenient and reliable [26,27]. These markers have been used for genetic analysis in many species, such as rice, yellow mustard, foxtail millet, maize, tomato, *B. juncea*, *B. rapa* and *Arabidopsis* [26–32]. In the present study, we successfully used PCR-based ILP markers for the development of a genetic map based on one RIL mapping population in *B. juncea*. Furthermore, we detected one novel major QTL for 2-propenyl and 3-butenyl GSLs on LG08 of the *B. juncea* genome. In addition, we also attempted to explore

the mechanism resulting in the variation in seed GSL contents by RNA-sequencing of siliques from the parental lines of the RIL mapping population.

Materials and methods

Plant materials and field trial

Parents G266 and G302 are DH lines. The seeds of the G266 line had low 2-propenyl and high 3-butenyl GSL contents, while G302 had high 2-propenyl and low 3-butenyl GSL contents. The RILs produced by G266×G302 displayed great variation in agronomic traits, such as flowering time and number of seeds per silique as presented in our earlier study [33]. Three replicates for each of the parental lines G266 and G302, their F₁ and 167 F₆ RILs were planted on the farm of Guizhou University, Guiyang, China in 2016 and 2017. The design of the trial was a randomized complete block. Each plot consisted of two rows with one size of 3.66 m² (3 m×1.22 m). Five grams of seeds from three plants in each plot were harvested at maturity and analyzed for 2-propenyl and low 3-butenyl GSL contents. The average 2-propenyl and low 3-butenyl GSL contents of the three replicates were used for QTL analysis.

DNA extraction and polymerase chain reaction (PCR)

Genomic DNA was extracted from young leaves of the parental lines, F₁ and 167 F₆ RILs using the modified sodium dodecyl sulfate method [34]. PCR of the ILP markers was carried out according to our previous studies [11,35]. Each PCR (20 μl) contained 1× standard PCR buffer (NEB), 1 U of Taq polymerase (NEB), 0.25 μM forward primer, 0.25 μM reverse primer, 100 μM each dNTP and 50 ng of genomic DNA in a total volume of 20 μL. The PCR amplification consisted of an initial denaturation at 94°C for 5 min; 35 cycles consisting of 94°C (45 sec); 55°C (45 sec) and 72°C (1 min); followed by termination at 72°C for 7 min. All PCR products were analyzed by electrophoresis in 2% agarose gels in 1× tri-acetate-ethylene diamine-tetra acetic acid buffer. Gels were visualized by staining in ethidium bromide and photographed on a digital gel documentation system.

Construction of genetic linkage map and QTL analysis

The genetic linkage map of *B. juncea* was constructed by using JoinMap 4.0 software at LOD_≥4.0 [36]. Recombination frequencies were converted to map distances in cm using the Kosambi mapping function and the genetic map was drawn with MapChart software [37]. QTL analysis of 2-propenyl and 3-butenyl GSL contents was performed using the interval mapping method of MapQTL 6.0 software [38]. A permutation test (1,000 replications) was used to determine the significance level for LOD with a genome-wide probability of p<0.05.

Glucosinolate component analysis

The 2-propenyl and 3-butenyl GSL contents of the mature seeds from each plot were analyzed following published methods [39,40] with minor modifications. Each seed sample was crushed and 200 mg of each sample was extracted twice with 2 ml boiling 70% methanol. The concentration of GSLs in the seeds was determined by high-performance liquid chromatography (Waters 2487/600/717) using the ISO9167-1 (1992) standard method.

RNA extraction, preparation, sequencing and data analysis

Total RNA (2 μg) was extracted from fresh seed coats of 20 DAP (days after pollination) siliques of three independent plants for each of the parental lines G266 and G302 using the TRIzol kit (Invitrogen, Carlsbad, CA), according to the manufacturer's instructions. The RNA

purity was checked using the Kaiiao K5500® Spectrophotometer (Kaiiao, Beijing, China), and the RNA integrity and concentration were assessed using the RNA Nano 6000 Assay Kit for the Bioanalyzer 2100 system (Agilent Technologies, CA, USA). Then, the six RNA samples were sent to the ANOROAD GENOME company (<http://www.genome.cn/>) for the construction of cDNA libraries and Illumina deep sequencing according to the paper of Wang et al. [41]. The raw RNA-sequencing data were filtered by a Perl script, following the steps of Wu et al. [42].

Identification and annotation of differentially expressed genes (DEGs)

DESeq2 v1.6.3 was designed for differential gene expression analysis between two samples with three biological replicates under the theoretical basis obeys the hypothesis of negative binomial distribution for the value of count. The p-value was corrected by the BH method. Genes with $q \leq 0.05$ and $|\log_2 \text{ratio}| \geq 1$ were identified as differentially expressed genes (DEGs) [43]. The DEGs obtained were further annotated with Gene Ontology (GO, <http://geneontology.org/>) and analyzed by KEGG (Kyoto Encyclopedia of Genes and Genomes, <http://www.kegg.jp/>) [44,45]. The GO enrichment of DEGs was implemented by the hypergeometric test, in which the p-value is calculated and adjusted to produce the q-value, and the data background is the genes in the whole genome. GO terms with $q < 0.05$ were considered to be significantly enriched. GO enrichment analysis was used to determine the biological functions of the DEGs. KEGG is a database resource containing a collection of manually drawn pathway maps representing our knowledge of molecular interaction and reaction networks. The KEGG enrichment of the DEGs was determined by the hypergeometric test, in which p-value was adjusted by multiple comparisons to produce the q-value. KEGG terms with $q < 0.05$ were considered to be significantly enriched.

Quantitative real time-PCR (qRT-PCR) analysis

Quantitative real-time PCR (qRT-PCR) was used to verify the transcript levels of the RNA-Seq results. Total RNA was extracted using the TRIzol kit (Invitrogen), according to the manufacturer's instructions. Then, the cDNA was synthesized by reverse transcription using Prime-Script RT reagent kits with gDNA Eraser (Takara, Dalian, China) according to the manufacturer's instructions. Sixteen gene-specific primers for qRT-PCR were designed based on reference unigene sequences randomly chosen from the DEGs using Primer Premier 5.0. Real-time PCR was conducted using SsoAdvancedTM Universal SYBRGreen Supermix (Hercules, CA) in a typical 20 μ l PCR mixture. The 20 μ l mixture contained 10 μ l SYBR Green Supermix (2 \times), 0.4 μ l reverse and forward primers (10 μ M), 2 μ l (100 ng) template cDNA, and 7.2 μ l ddH₂O. The qRT-PCR conditions were 95°C for 2 min, followed by 40 cycles of 95°C for 10 s (denaturation), followed by 60°C for 20 s (annealing and extension). The $2^{-\Delta\Delta C_t}$ algorithm was used to calculate the relative level of gene expression. The β -actin gene was used as the internal control, and the T399 samples served as the control. All qRT-PCR were performed with three biological replicates, and run on a Bio-Rad CFX96 Real Time System (Bio-Rad, Hercules, CA, USA).

Results

Polymorphism between the parental lines G266 and G302

A total of 1,272 ILP primers, 284 from *Arabidopsis thaliana* [32], 745 from *B. napus* and 243 from *B. rapa* available in the Potential Intron Polymorphism (PIP) database [27], were used to screen the parental lines G266 and G302 for polymorphic primers. Of the 1,272 ILP primers,

306 (24.1%) generated clear and scorable polymorphic bands between the parental lines varying in size from 150 to 1250 bp. Among the 306 polymorphic primers, 266 (86.9%) amplified one locus, 35 (12.4%) produced two loci, 4 (1.4%, At4g11790, At1g07980, PIP1848 and At3g52990) produced three loci, 2 (0.7%, PIP1202 and PIPR68) revealed four loci and one (0.4%, At1g72890) revealed five loci. In summary, 359 polymorphic markers were amplified by 306 polymorphic primers, including 231 dominant ones and 128 codominant ones. The 359 polymorphic markers were used to construct the linkage map with the RIL population of G266×G302 in *Brassica juncea*.

Construction of one genetic linkage map

A total of 304 polymorphic loci of the 359 polymorphic markers (84.7%) were mapped on 18 linkage groups and covered a genetic length of 1671.13 centiMorgans (cM) with an average marker interval of 5.50 cM (Table 1 and Figs 1 and 2). The linkage groups were designated as LG01-LG18.

The map lengths of the 18 linkage groups ranged from 54.41 cM for LG16 to 180.48 cM for LG08 with an average of 92.843 cM. The marker interval ranged from 0.00 cM to 37.39 cM with an average of 5.75 cM. LG03, LG08, LG14 and LG18 had map lengths longer than 100 cM, ranging from 110.44 cM to 180.48 cM. LG18 had the largest average marker interval of 11.04 cM. LG08 had the longest map length of 180.48 cM and the most ILP markers (35 markers). LG07, LG15 and LG17 had similar map lengths ranging from 93.83 cM to 97.63 cM. LG01, LG04, LG09, LG11 and LG12 had similar map lengths ranging from 80.25 cM to 88.37 cM. LG04 had the largest island without markers (37.39 cM). LG10, LG06 and LG13 had similar long map lengths ranging from 75.00 cM to 78.53 cM. LG02, LG05 and LG16 had similar long map lengths ranging from 54.41 cM to 64.38 cM. LG05 had the shortest map length of 54.41 cM and the smallest average marker interval of 1.92 cM.

Table 1. Characterization of the 18 linkage groups in *Brassica juncea*.

| Linkage Group | Map Length (cM) | Marker interval (cM) | | | No. of markers |
|---------------|-----------------|----------------------|-------------------|-------------------|----------------|
| | | Average | Max Distance (cM) | Min Distance (cM) | |
| LG01 | 88.37 | 5.52 | 24.19 | 0.06 | 17 |
| LG02 | 64.38 | 5.36 | 13.79 | 0.35 | 13 |
| LG03 | 134.25 | 6.39 | 19.30 | 0.01 | 22 |
| LG04 | 82.05 | 8.21 | 37.39 | 0.33 | 11 |
| LG05 | 61.31 | 1.92 | 20.96 | 0.22 | 33 |
| LG06 | 77.10 | 7.71 | 33.53 | 0.00 | 10 |
| LG07 | 93.83 | 11.73 | 25.79 | 0.77 | 9 |
| LG08 | 180.48 | 5.31 | 15.07 | 0.00 | 35 |
| LG09 | 80.35 | 3.21 | 16.21 | 0.19 | 26 |
| LG10 | 75.00 | 7.50 | 16.21 | 0.83 | 11 |
| LG11 | 83.13 | 7.56 | 13.69 | 0.74 | 12 |
| LG12 | 80.25 | 10.03 | 16.62 | 5.10 | 9 |
| LG13 | 78.53 | 8.73 | 16.85 | 2.77 | 10 |
| LG14 | 135.16 | 5.88 | 14.43 | 0.01 | 24 |
| LG15 | 94.46 | 4.11 | 24.01 | 0.04 | 24 |
| LG16 | 54.41 | 4.53 | 13.02 | 0.60 | 13 |
| LG17 | 97.63 | 7.51 | 20.67 | 0.10 | 14 |
| LG18 | 110.44 | 11.04 | 26.26 | 0.62 | 11 |
| Total | 1671.13 | - | - | - | 304 |

<https://doi.org/10.1371/journal.pone.0220597.t001>

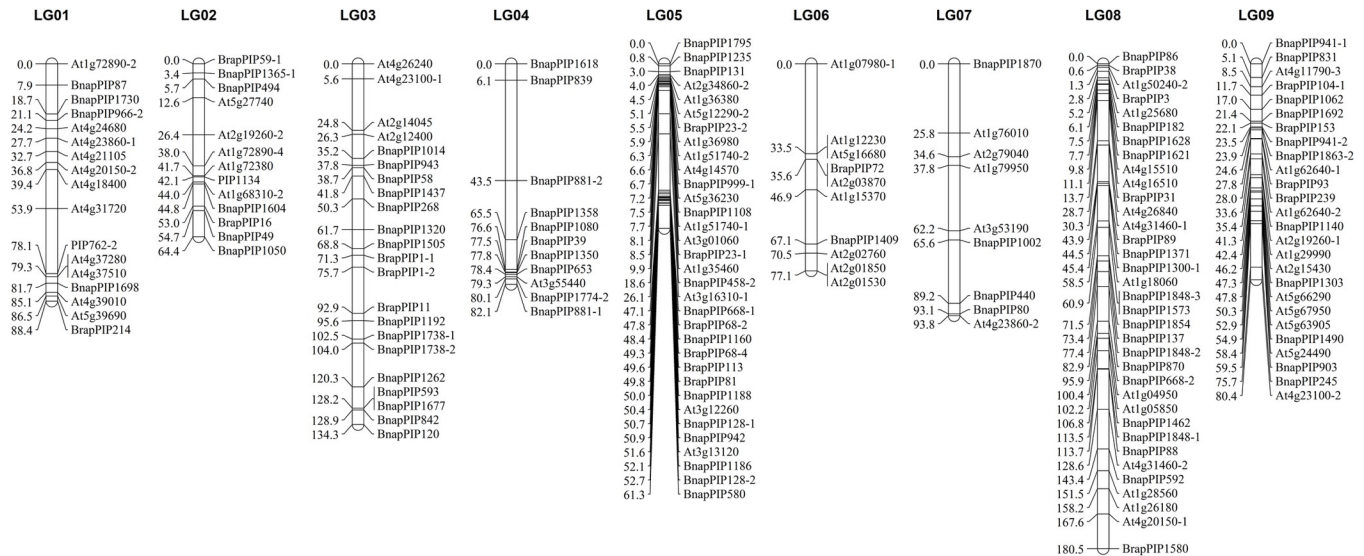


Fig 1. The 9 linkage groups from LG01 to LG09 in *Brassica juncea*. For each linkage group, the ILP markers were shown on the right side and the marker position in centiMorgan on the left side.

<https://doi.org/10.1371/journal.pone.0220597.g001>

QTL mapping of 2-propenyl and 3-butenyl glucosinolate contents

The 167 RILs, their parental lines G266 and G302, and the F₁, were grown in the field with three replications. These lines were grown at Guiyang and distributed normally in 2016 and 2017 (Fig 3). No significant difference across the two years for 2-propenyl and 3-butenyl GSL contents was detected ($p = 0.594$ and $p = 0.888$, respectively). The 2-propenyl GSL contents was significantly negatively correlated with the 3-butenyl GSL contents in 2016 ($r = -0.920$, $p = 0.000$) and 2017 ($r = -0.914$, $p = 0.000$), respectively. The parents of the population differed in 2-propenyl GSL contents, with mean values of 17.29 $\mu\text{mol/g}$ and 180.90 $\mu\text{mol/g}$ for G266 and G302, respectively (Table 2). The mean 2-propenyl GSL contents of F₁ was 67.34 $\mu\text{mol/g}$, which was closer to that of the female parent and slightly higher than the mean value of the RIL mapping population (63.45 $\mu\text{mol/g}$) (Table 2). The range of 2-propenyl GSL contents in the RILs was 10.70~214.36 $\mu\text{mol/g}$ in 2016 and 9.59~215.60 $\mu\text{mol/g}$ in 2017 (Table 2). The

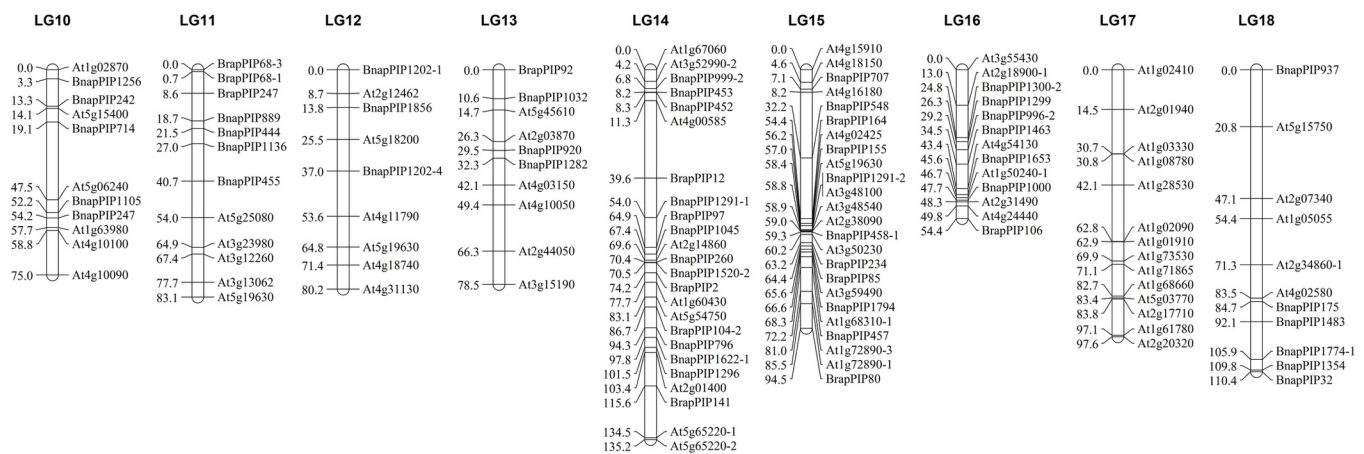


Fig 2. The 9 linkage groups from LG10 to LG18 in *Brassica juncea*. For each linkage group, the ILP markers were shown on the right side and the marker position in centiMorgan on the left side.

<https://doi.org/10.1371/journal.pone.0220597.g002>

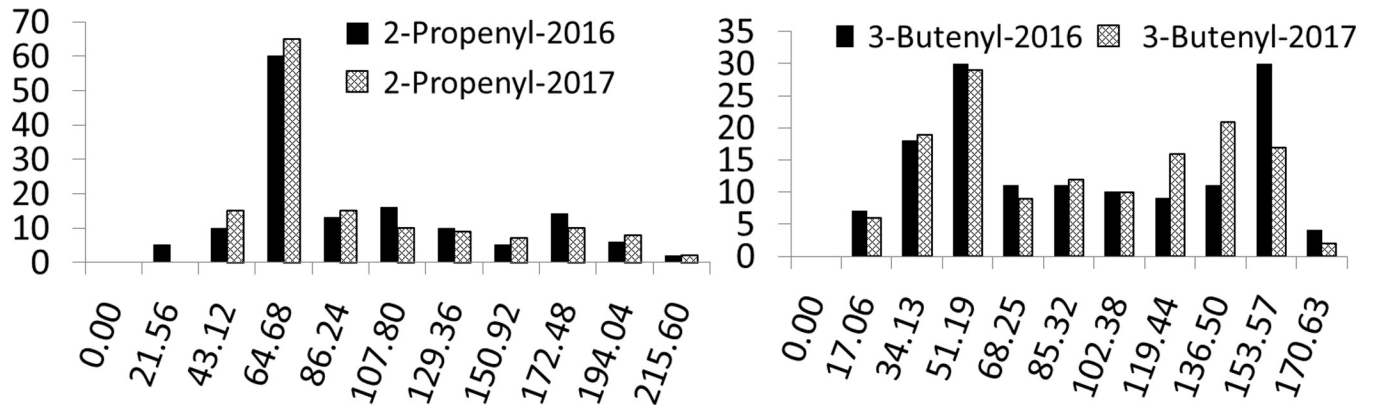


Fig 3. Frequency distribution of 2-propenyl and 3-butenyl glucosinolate content of the recombinant inbred lines (RILs) in 2016 and 2017. Y-axis: the number of the corresponding recombinant inbred lines; X-axis: the glucosinolate content of 2-propenyl (left) and 3-butenyl (right) respectively (u mol/g).

<https://doi.org/10.1371/journal.pone.0220597.g003>

parents of the population differed in 3-butenyl GSL contents, with mean values of 86.98 $\mu\text{mol/g}$ and 15.80 $\mu\text{mol/g}$ for G266 and G302, respectively (Table 2). The mean 3-butenyl GSL contents of F_1 was 63.80 $\mu\text{mol/g}$, more similar to that of the female parent and slightly lower than the mean value in the RIL mapping population (75.98 $\mu\text{mol/g}$) (Table 2). The range of 3-butenyl GSL contents of the RILs was 8.71~170.63 u mol/g in 2016, and 5.82~166.04 $\mu\text{mol/g}$ in 2017 (Table 2).

QTL analysis was performed for 2-propenyl (SIN, sinigrin) and 3-butenyl (GNA, glucanapin) GSL contents in 2016 and 2017 respectively. Two major QTLs of SIN-2016 and SIN-2017 for 2-propenyl GSL contents (LOD = 16.46 and 16.64) were colocalized between At1g26180 and BnapPIP1580 on LG08, accounting for 42.3% and 42.6% of the total variation in 2016 and 2017, respectively (Table 3 and Fig 4). Another two QTLs for 3-butenyl GSL contents were detected and colocalized with GNA-2016 and GNA-2017 between At1g26180 and BnapPIP1580 on LG08 (Table 3 and Fig 4). The two QTLs for 3-butenyl GSL contents explained 31% and 38.4% of the total variation in 2016 and 2017 respectively (Table 3). All these QTLs were mapped to a region adjacent to the ILP marker At4g20150-1 on LG08 (Table 3 and Fig 4).

Syntenic relationships between LG08 and A08 of *B. juncea* and *B. rapa*

LG08 contained a total of 35 ILP markers. Among these, 13 (37.1%) were developed from the single-copy genes of *A. thaliana*, and 17 and 5 were developed from the unique

Table 2. 2-propenyl and 3-butenyl GSL content of the parental lines G266 and G302, F_1 seeds of G266×G302 and the RIL mapping population in 2016 and 2017.

| Traits (u mol/g) | Years | Parents | | F_1 | RILs | |
|------------------|-------|-------------------------|--------------|------------|-------|--------------|
| | | G266 | G302 | | Mean | Range |
| 2-Propenyl | 2016 | 19.14±0.21 ^a | 181.53±7.32 | 73.50±1.12 | 65.45 | 10.70~214.36 |
| | 2017 | 15.43±0.60 | 180.27±11.51 | 61.18±1.91 | 61.44 | 9.59~215.60 |
| | mean | 17.29 | 180.90 | 67.34 | 63.45 | – |
| 3-Butenyl | 2016 | 86.79±1.00 | 20.68±0.16 | 57.66±0.89 | 76.42 | 8.71~170.63 |
| | 2017 | 87.17±1.15 | 10.91±0.63 | 69.93±1.20 | 75.53 | 5.82~166.04 |
| | mean | 86.98 | 15.80 | 63.80 | 75.98 | |

^a: standard deviation.

<https://doi.org/10.1371/journal.pone.0220597.t002>

Table 3. QTLs for 2-propenyl and 3-Butenyl GSL components in the RIL mapping population derived from the cross of G266 and G302 in 2016 and 2017.

| Name of the QTLs ^a | Chromosome | Peak Position | LOD ^b | R ² , % ^c | Additive effect (μmol/g seed) | Nearest ILP and its interval to the peak |
|-------------------------------|------------|---------------|------------------|---------------------------------|-------------------------------|--|
| SIN-2016 | LG08 | 166.21 | 16.46 | 42.3 | 6.09 | At4g20150-1,1.37 |
| SIN-2017 | LG08 | 164.21 | 16.64 | 42.6 | 6.25 | At4g20150-1,3.37 |
| GNA-2016 | LG08 | 165.21 | 11.14 | 31.0 | 6.63 | At4g20150-1,2.37 |
| GNA-2017 | LG08 | 165.21 | 14.53 | 38.4 | 0.10 | At4g20150-1,2.37 |

^a SIN: sinigrin or 2-propenyl GSL; BUT: gluconapin or 3-Butenyl GSL; 2016: the results are from the field experiment in 2016; 2017: the results are from the field experiment in 2017.

^b LOD: logarithm of the odds score for QTLs calculated by composite interval mapping.

^c R²: the phenotypic variation explained by a QTL in percentage.

<https://doi.org/10.1371/journal.pone.0220597.t003>

transcript fragments of *B. napus* and *B. rapa*, respectively [27]. To validate the exact linkage group of LG08 with QTLs for 2-propenyl and 3-butenyl GSL contents, a synteny analysis was performed. Eight of the 13 *A. thaliana* markers prefixed “At” on LG08 showed synteny between LG08 and A08 of *B. juncea* (Fig 4). Furthermore, the sequences of the unique transcript fragments for developing the 22 *B. napus* and *B. rapa* markers were used to blast against the *Brassica rapa* genome (*Brassica rapa* cultivar Chiifu, Brapa_1.0) in NCBI (<https://www.ncbi.nlm.nih.gov/>), and 18 (81.8%) were mapped to the A08 chromosome (S1 Table). The synteny analysis indicated that LG08 was exactly chromosome A08 of the A genome.

Transcriptome analysis of the parental siliques as potential

GSL source for seeds. To construct a de novo transcriptome database, three RNA libraries were generated for each of G266 and G302 lines through Illumina sequencing. A total of 139,197,152 and 138,957,982 raw reads were generated from the G266 and G302 libraries, respectively (S2 Table). After removing low quality reads, adapter polluted reads and higher N contents (>5%) reads, a total of 134,953,284 (96.95%, T399) and 134,832,452 (97.03%, T085) clean reads were obtained (S2 Table). After filtering out the genes that contained only one exon or encoded short peptide chains (<50 amino acid residues), a total of 81,826 transcripts were revealed by blasting the reference genome using DESeq2 (v1.6.3). To functionally annotate those transcripts, the 81,826 transcripts were blasted in search of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG). Finally, 8,1694 transcripts (92.64%) were successfully annotated by GO and KEGG, and 324 ones of these transcripts were involved in GSL metabolism (S3 Table).

To identify differentially expressed genes (DEGs) involved in the GSL metabolism of siliques between G266 and G302. A rigorous algorithm with a threshold of “FDR≤0.05 and |Log₂Ratio|≥1” was developed and used as thresholds to judge the significance of differences in transcript abundance. It was found that only 24 transcripts were DEGs related to GSL metabolism. Among those DEGs, 15 genes are involved in the biosynthesis and transport of aliphatic GSLs (Table 4), and 5 and 4 genes are involved in the biosynthesis of indolic GSLs and the breakdown of GSLs, respectively (S4 Table).

To verify the expression of these transcripts detected by RNA-Seq, the 15 candidate genes involved in the aliphatic GSL metabolism pathway were randomly chosen for validation by qRT-PCR. Detailed information about the chosen DEGs and primers is listed in S5 Table. The data obtained by qRT-PCR were consistent with the RNA-Seq results (Fig 5), suggesting the reliability of the transcriptome database.

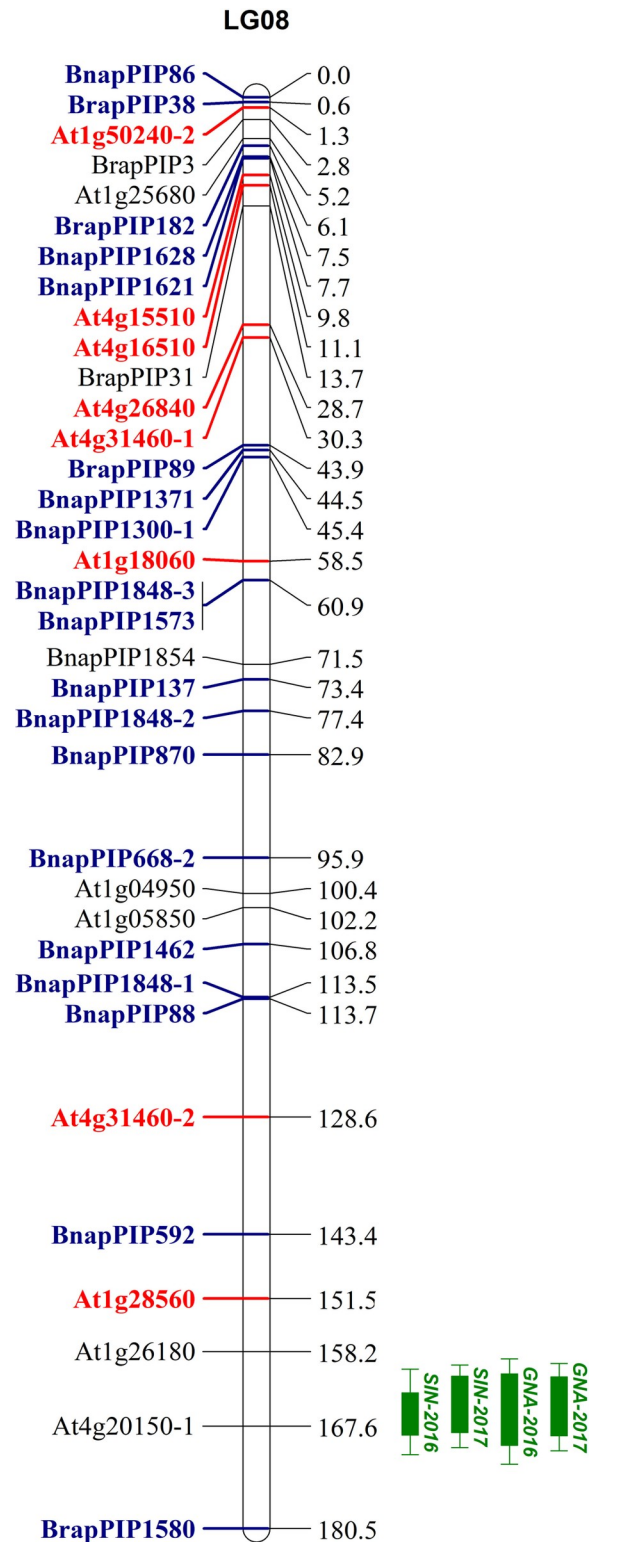


Fig 4. QTL mapping of 2-propenyl and 3-butenyl GSLs and synteny analysis between LG08 in the present study and A08 of *B. juncea* and *B. rapa*. The QTLs for 2-propenyl (SIN, Sinigrin) co-localized with the QTLs for 3-butenyl (GNA, Gluconapin) GSL content. 1-LOD and 2-LOD supporting intervals of each QTL were marked by thick and thin bars, respectively. Eight *A. thaliana* markers prefixed “At” (bold and red) on LG08 show a synteny with A08 of *B. juncea* in publish papers [32]. Eighteen ILP markers prefixed “Bnap and Brap” (bold and blue) on LG08 show a

synteny between LG08 and A08 of *B.rapa* through balsting analysis with Brapa_1.0 of *Brassica rapa* cultivar Chiifu in NCBI (<https://www.ncbi.nlm.nih.gov/>).

<https://doi.org/10.1371/journal.pone.0220597.g004>

Joint QTL mapping and RNA-sequencing analyses reveal the candidate genes for GSL metabolism in *B. juncea*

To integrate the results of QTL mapping and RNA-sequencing, we performs an alignment analysis between the 24 DEGs related to GSL metabolism and the reference genome of *B. juncea* [16] by the BLAST-like alignment tool [54]. Only two DEGs of LOC106429668 encoding MYB28/MYB29/MYB76 (physical position: 23,048,306) and LOC106416451 (physical position: 5,833,626) encoding IIL1 were located on A08 of *B. juncea* genome. In addition, the DNA sequence designed for the PIP markers on A08 also blast the *B. juncea* reference genome [16]. The QTLs for 2-propenyl and 3-Butenyl GSL is located between the physical position of 18,549,777 (BnapPIP592) and the start point of the chromosome, the region of which was overlapped with the position of LOC106416451.

Discussion

Recombinant inbred lines (RILs) are an important resource in the genetic mapping of complex traits in many species. The RIL mapping population was successfully produced in our laboratory, allowing us to construct a genetic linkage map by utilizing ILP markers in *B. juncea*. The ILP primers were designed on the conserved exons flanking the target intron of cDNA/EST sequences to exploit its polymorphism. Each ILP marker locus might represent one gene copy in the studied genome. Taking the polymorphic and monomorphic loci together, approximately 46.8% of the 306 polymorphic ILP primers in the present study revealed more than one locus, indicating a very close ratio revealed by ILP and RFLP primers in the earlier studies [55,56]. The multiple loci revealed by ILP primers confirmed the polyploidy of *B. juncea*. The higher polymorphism ratio of 24.1% in the ILP primers between the parental lines G266 and G302 not only revealed the hypervariability of ILP primers but also suggested a high degree of variation between the parental genomes. This high genetic difference between the parental

Table 4. The candidate genes involved in biosynthesis of and transporting of aliphatic GSLs.

| SN ^a | Gene Name | AGI code | Unigene ID from RNA-Seq | Function |
|-----------------|-----------|-----------|---|--|
| 1 | CYP83A1 | AT4G13770 | LOC106447562(U) ^b ;LOC106391682(U) | aldoxime→s-alkyl-thiohydroximate [46–48] |
| 2 | SOT18 | AT1G74090 | LOC106366617(D);LOC106354324(D);LOC106436726(D) | PAPS-dependent sulfation of desulfo-GSLs→GSLs [49] |
| 3 | SOT17 | AT1G18590 | | |
| 4 | IIL1 | AT4G13430 | LOC106416451(U) | 2-Alkyl-malic acid→3-Alkyl-malic acid [47,50] |
| 5 | IPMI2 | AT2G43100 | LOC106434491(U) | |
| 6 | IPMI SSU1 | AT2G43090 | | |
| 7 | AOP3 | AT4G03050 | LOC106430050(U);LOC106438719(U);LOC106389979(U) | methylsulfinylalkyl GSL→ hydroxyalkyl GSL [47] |
| 8 | SUR1 | AT2G20610 | LOC106440999(D) | s-alkyl-thiohydroximate→ thiohydroximate [47,51,52] |
| 9 | MYB28 | AT5G61420 | LOC106382207(U);LOC106429668(U) | the whole process of biosynthesis of methionine-derived GSL [46] |
| 10 | MYB29 | AT5G07690 | | |
| 11 | MYB76 | AT5G07700 | | |
| 12 | GTR2 | AT5G62680 | LOC106411192(U);LOC106347844(D) | GSL transporting [53] |
| 13 | GTR1 | AT3G47960 | | |

^a: Serial Number.

^b: “U” means up-regulated, “D” means down-regulated.

<https://doi.org/10.1371/journal.pone.0220597.t004>

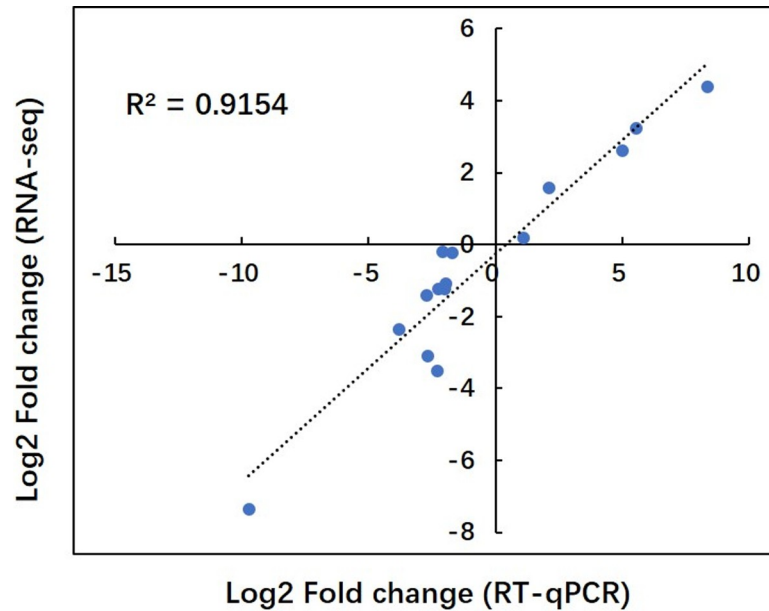


Fig 5. Comparison of gene expression values obtained by RNA-seq and qRT-PCR. Fold changes were calculated for 15 DEGs and a high correlation ($R^2 = 0.92$) was observed between the results obtained using the two techniques.

<https://doi.org/10.1371/journal.pone.0220597.g005>

lines would be convenient to increase the density of genetic markers in different linkage groups in *B. juncea*.

The 2-propenyl and 3-butenyl GSLs are the major glucosinolates found in *B. juncea* [57]. The colocalized QTL regions of 2-propenyl and 3-butenyl GSLs on A08 of *B. juncea* represented one novel QTL first detected on A08 in Brassicaceae [57–59] that explained average phenotypic variations of 42.5% and 42.6%, respectively. The first reason might be that fewer studies focused on 2-propenyl and 3-butenyl GSLs of the germplasm originating in China as the important center of origin in *B. juncea*. Second, the exact order of linkage groups was difficult to obtain before the *B. juncea* genome was sequenced [16]. Another reason might be recombination events between the chromosomes of *B. juncea* in different regions. Although QTL mapping of 2-propenyl and 3-butenyl GSLs has been performed previously in *B. juncea* [57–59], no convenient and reliable primers could be used for gel detection in marker-assisted breeding. In the present study, all 306 primers mapped on *B. juncea* were ILP-type, providing a convenient, specific and rapid detection method for agarose gel electrophoresis in marker-assisted breeding. The ILP marker At4g20150-1 on A08 had the nearest distance of 1.37–3.37 cM to the peak. The ILP primer At4g20150 produced six fragments between 400 bp and 700 bp in the present study and mostly represented six copies of one gene, among which two copies were polymorphic and mapped to LG01 (At4g20150-2) and LG08 (At4g20150-1), respectively. In another genetic map of *B. juncea*, the primer amplified three copies from A3, B7 and B8 [32]. The specific marker of At4g20150 was codominant, clear and simple to score by agarose gel, which resulted in one 700 bp fragment. The novel QTL and the linked marker of At4g20150-1 can be helpful in exploiting the metabolic mechanism of 2-propenyl and 3-butenyl GSLs. The marker tightly linked to QTLs can also be used for marker-assisted selection (MAS). For example, Xu et al. (2018) transferred a thermostable β -amylase from wild barley into a commercial variety, and identified several elite lines with MAS [60], and a major QTL for resistance to *Fusarium* head blight was transferred from *Thinopyrum elongatum* onto durum wheat 7AL chromosome arm by MAS [61].

GSL synthesis in seeds is nearly nonexistent, so these compounds are mainly imported from other tissues [17,62–64]. As the closest GSL source and the only organ with similar types of GSLs, the siliques might be the source of most seed GSLs in Brassicaceae. In the present study, the RNA-Seq technique was used to screen the key genes for aliphatic GSL synthesis in the siliques of the parental lines G266 and G302. In the siliques of G302, which has high aliphatic GSL contents, 9 DEGs associated with *CYP83A1*, *ILL1*, *IPMI2*, *IPMI SSU1*, *AOP3*, *MYB28*, *MYB29* and *MYB76* were upregulated, resulting in its high GSL contents. Surprisingly, *SURI* and *SOT17-18* were downregulated for unknown reasons. Furthermore, *GTR2* is upregulated in the siliques of G302, and thus, it might play a more important role in GSL transport from siliques to seeds than the downregulated gene *GTR1*. The process of GSL metabolism is complicated, and few advances have been achieved. The main reasons might be the transport of GSLs among different organs. The combined method of QTL mapping and RNA-Seq should be helpful for the future fine mapping and gene cloning of 2-propenyl and 3-butenyl GSLs.

The joint QTL mapping and RNA-sequencing analyses reveal one candidate gene of LOC106416451 encoding *ILL1*. *ILL1* is mainly responsible for the isomerization of 2-alkyl-malic acid to form 3-alkyl-malic acid [47,50,65,66]. In the present study, *ILL1* is significantly highly (Log_2 Fold Change = 9.71, $p = 3.58\text{E-}15$) expressed in the siliques of G302 with high aliphatic GSLs than that in G266 with low aliphatic GSLs. The primary work validates that the *ILL1* might be the key gene for GSL regulation in the present RIL mapping population. However, more work is needed to narrow the QTL region and validate the candidate gene of LOC106416451 in our future study.

In addition, the mapping population used in the present study displayed great variation in agronomic and quality traits in this study and our earlier study [33]. The constructed genetic map would be useful in QTL mapping, gene cloning and marker-based precision breeding of more important traits in *B. juncea*. Because the number of traditional genetic markers is limited, we would sequence the RILs to develop more SNPs (single nucleotide polymorphism) and create a unified, saturated genetic map of *B. juncea* in the future.

Supporting information

S1 Table. The synteny analysis between unique transcript of ILP primers and the Brapa_1.0 genome.

(DOCX)

S2 Table. The candidate genes involved in biosynthesis of and transporting of aliphatic GSLs.

(DOCX)

S3 Table. The 324 transcripts involved in GSL metabolism.

(XLSX)

S4 Table. The candidate genes involved in biosynthesis of indolic GLS.

(DOCX)

S5 Table. The primers for qRT-PCR validation.

(DOCX)

Acknowledgments

The authors thank Shuchun Lin, the agronomist of the teaching and experiment farmer of Guizhou University, for his hard work in field experiment.

Author Contributions

Conceptualization: Kunjiang Yu.

Investigation: Aimal Nawaz Khattak, Tianya Wang, Wei Wan, Botao Ye.

Resources: Entang Tian.

Supervision: Entang Tian.

Writing – original draft: Aimal Nawaz Khattak, Tianya Wang, Entang Tian.

Writing – review & editing: Aimal Nawaz Khattak, Tianya Wang, Kunjiang Yu, Renqin Yang, Wei Wan, Botao Ye, Entang Tian.

References

1. Wilson RA, Sangha MK, Banga SS, Atwal AK, Gupta S (2014) Heat stress tolerance in relation to oxidative stress and antioxidants in Brassica juncea. *Journal of environmental biology* 35: 383–387. PMID: [24665766](https://pubmed.ncbi.nlm.nih.gov/24665766/)
2. Kumar A, Singh DP (1998) Use of Physiological Indices as a Screening Technique for Drought Tolerance in Oilseed Brassica Species. *Annals of Botany* 81: 413–420.
3. Phutela A, Jain V, Dhawan K, Nainawatee HS (2000) Proline Metabolism Under Water Stress in the Leaves and Roots of Brassica juncea Cultivars Differing in Drought Tolerance. *Journal of Plant Biochemistry and Biotechnology* 9: 35–39.
4. Srivastava SK (1987) Peroxidase and Poly-Phenol Oxidase in Brassica juncea Plants Infected with Macrophomina phaseolina (Tassai) Goid. and their Implication in Disease Resistance. *Journal of Phytopathology* 120: 249–254.
5. Roy NN (1984) Interspecific transfer of Brassica juncea-type high blackleg resistance to Brassica napus. *Euphytica* 33: 295–303.
6. Østergaard L, Kempin SA, Bies D, Klee HJ, Yanofsky MF (2006) Pod shatter-resistant Brassica fruit produced by ectopic expression of the FRUITFULL gene. *Plant Biotechnology Journal* 4: 45–51. <https://doi.org/10.1111/j.1467-7652.2005.00156.x> PMID: [17177784](https://pubmed.ncbi.nlm.nih.gov/17177784/)
7. Liu Y, Wei W, Ma K, Darmency H (2013) Spread of introgressed insect-resistance genes in wild populations of Brassica juncea: a simulated in-vivo approach. *Transgenic Research* 22: 747–756. <https://doi.org/10.1007/s11248-012-9679-y> PMID: [23250587](https://pubmed.ncbi.nlm.nih.gov/23250587/)
8. Love HK, Rakow G, Ranney JP, Downey RK. Breeding improvements towards canola quality Brassica juncea; 1991; Saskatoon, Canada. pp. 164–169.
9. Delourme R, Foisset N, Horvais R, Barret P, Champagne G, et al. (1998) Characterisation of the radish introgression carrying the Rfo restorer gene for the Ogu-INRA cytoplasmic male sterility in rapeseed (Brassica napus L.). *Theoretical and Applied Genetics* 97: 129–134.
10. Delourme R, Eber F (1992) Linkage between an isozyme marker and a restorer gene in radish cytoplasmic male sterility of rapeseed (Brassica napus L.). *Theoretical and Applied Genetics* 85: 222–228. <https://doi.org/10.1007/BF00222863> PMID: [24197308](https://pubmed.ncbi.nlm.nih.gov/24197308/)
11. Tian E, Roslinsky V, Cheng B (2014) Molecular marker-assisted breeding for improved Ogura cms restorer line (RfoRfo) and mapping of the restorer gene (Rfo) in Brassica juncea. *Molecular Breeding* 34: 1361–1371.
12. Fahey JW, Zalcmann AT, Talalay P (2001) The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* 56: 5–51. [https://doi.org/10.1016/s0031-9422\(00\)00316-2](https://doi.org/10.1016/s0031-9422(00)00316-2) PMID: [11198818](https://pubmed.ncbi.nlm.nih.gov/11198818/)
13. Kjær A, Olesen Larsen P (1973) Non-protein amino acids, cyano-genic glycosides and glucosinolates. In: Geissman T.A. (Ed.), *Specialist Periodical Reports*. The Chemical Society, London. 71–105.
14. Hogge LR, Reed DW, Underhill EW, Haughn GW (1988) HPLC Separation of Glucosinolates from Leaves and Seeds of Arabidopsis thaliana and Their Identification Using Thermospray Liquid Chromatography/Mass Spectrometry*. *Journal of Chromatographic Science* 26: 551–556.
15. Halkier BA, Du L (1997) The biosynthesis of glucosinolates. *Trends in Plant Science* 2: 425–431.
16. Yang J, Liu D, Wang X, Ji C, Cheng F, et al. (2016) The genome sequence of allopolyploid Brassica juncea and analysis of differential homoeolog gene expression influencing selection. *Nature Genetics* 48: 1225. <https://doi.org/10.1038/ng.3657> PMID: [27595476](https://pubmed.ncbi.nlm.nih.gov/27595476/)

17. Nour-Eldin HH, Halkier BA (2009) Piecing together the transport pathway of aliphatic glucosinolates. *Phytochemistry Reviews* 8: 53–67.
18. Brown PD, Tokuhisa JG, Reichelt M, Gershenzon J (2003) Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. *Phytochemistry* 62: 471–481. [https://doi.org/10.1016/s0031-9422\(02\)00549-6](https://doi.org/10.1016/s0031-9422(02)00549-6) PMID: 12620360
19. Wan W, Yu K, Ye B, Khattak AN, Wang T, et al. (2019) Screening, cloning and molecular characterization of the candidate genes for determinate inflorescence in *Brassica napus*. *Plant Genomics in China XX Nan Chang*.
20. Sodhi YS, Mukhopadhyay A, Arumugam N, Verma JK, Gupta V, et al. (2002) Genetic analysis of total glucosinolate in crosses involving a high glucosinolate Indian variety and a low glucosinolate line of *Brassica juncea*. *Plant Breeding* 121: 508–511.
21. Love HK, Rakow G, Raney JP, Downey RK (1990) Genetic control of 2-propenyl and 3-butenyl glucosinolate synthesis in mustard. *Canadian Journal of Plant Science* 70: 425–429.
22. Gland A (1981) Variation of alkenyl glucosinolates in seeds of *Brassica* species. *Z Pflanzenzuchtg* 87: 96–110.
23. Ripley VL, Roslinsky V (2005) Identification of an ISSR Marker for 2-propenyl glucosinolate Content in *Brassica juncea* L. and Conversion to a SCAR Marker. *Molecular Breeding* 16: 57–66.
24. Sotelo T, Soengas P, Velasco P, Rodríguez VM, Cartea ME (2014) Identification of Metabolic QTLs and Candidate Genes for Glucosinolate Synthesis in *Brassica oleracea* Leaves, Seeds and Flower Buds. *PLOS ONE* 9: e91428. <https://doi.org/10.1371/journal.pone.0091428> PMID: 24614913
25. Feng J, Long Y, Shi L, Shi J, Barker G, et al. (2012) Characterization of metabolite quantitative trait loci and metabolic networks that control glucosinolate concentration in the seeds and leaves of *Brassica napus*. *New Phytol* 193: 96–108. <https://doi.org/10.1111/j.1469-8137.2011.03890.x> PMID: 21973035
26. Wang X, Zhao X, Zhu J, Wu W (2005) Genome-wide Investigation of Intron Length Polymorphisms and Their Potential as Molecular Markers in Rice (*Oryza sativa* L.). *DNA Research* 12: 417–427. <https://doi.org/10.1093/dnares/dsi019> PMID: 16769698
27. Yang L, Jin G, Zhao X, Zheng Y, Xu Z, et al. (2007) PIP: a database of potential intron polymorphism markers. *Bioinformatics* 23: 2174–2177. <https://doi.org/10.1093/bioinformatics/btm296> PMID: 17545179
28. Javidfar F, Cheng B (2013) Single locus, multiallelic inheritance of erucic acid content and linkage mapping of *FAE1* gene in yellow mustard. *Crop Sci* 53: 825–832.
29. Muthamilarasan M, Venkata Suresh B, Pandey G, Kumari K, Parida SK, et al. (2014) Development of 5123 Intron-Length Polymorphic Markers for Large-Scale Genotyping Applications in Foxtail Millet. *DNA Research* 21: 41–52. <https://doi.org/10.1093/dnares/dst039> PMID: 24086082
30. Liu H, Lin Y, Chen G, Shen Y, Liu J, et al. (2012) Genome-scale identification of resistance gene analogs and the development of their intron length polymorphism markers in maize. *Molecular Breeding* 29: 437–447.
31. Wang Y, Chen J, Francis DM, Shen H, Wu T, et al. (2010) Discovery of intron polymorphisms in cultivated tomato using both tomato and *Arabidopsis* genomic information. *Theoretical and Applied Genetics* 121: 1199–1207. <https://doi.org/10.1007/s00122-010-1381-y> PMID: 20552324
32. Panjabi P, Jagannath A, Bisht N, Padmaja K, Sharma S, et al. (2008) Comparative mapping of *Brassica juncea* and *Arabidopsis thaliana* using intron polymorphism (IP) markers: homoeologous relationships, diversification and evolution of the A, B and C brassica genomes. *BMC Genomics* 9: 1–19. <https://doi.org/10.1186/1471-2164-9-1> PMID: 18171476
33. Tian E, Li L, Jia S, Lin S (2018) Variation and Statistic Analysis of Agronomic Traits of Recombinant Inbred Lines of *Brassica juncea* L. *Asian Agricultural Research* 10.
34. Somers DJ, Friesen KRD, Rakow G (1998) Identification of molecular markers associated with linoleic acid desaturation in *Brassica napus*. *Theor Appl Genet* 96: 897–903.
35. Tian E, Zeng F, MacKay K, Roslinsky V, Cheng B (2014) Detection and Molecular Characterization of Two FAD3 Genes Controlling Linolenic Acid Content and Development of Allele-Specific Markers in Yellow Mustard (*Sinapis alba*). *PLOS ONE* 9: e97430. <https://doi.org/10.1371/journal.pone.0097430> PMID: 24823372
36. Van Ooijen JW (2006) JoinMap® Version 4.0: Software for the calculation of genetic linkage maps. Plant Research International, Wageningen, The Netherlands.
37. Voorrips RE (2002) MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *Journal of Heredity* 93: 77–78. <https://doi.org/10.1093/jhered/93.1.77> PMID: 12011185
38. Van Ooijen JW (2009) MapQTL 6: Software for the mapping of quantitative trait loci in experimental populations of diploid species. *Kyazma BV: Wageningen, Netherlands*.

39. Agerbirk N, Olsen CE, Nielsen JK (2001) Seasonal variation in leaf glucosinolates and insect resistance in two types of *Barbarea vulgaris* ssp. *arcuata*. *Phytochemistry* 58: 91–100. [https://doi.org/10.1016/S0031-9422\(01\)00151-0](https://doi.org/10.1016/S0031-9422(01)00151-0) PMID: 11524118
40. Wathelet J-P, Iori R, Leoni O, Rollin P, Quinsac A, et al. (2004) Guidelines for glucosinolate analysis in green tissues used for biofumigation. *Agroindustria* 3.
41. Wang B, Tseng E, Regulski M, Clark TA, Hon T, et al. (2016) Unveiling the complexity of the maize transcriptome by single-molecule long-read sequencing. *Nature Communications* 7: 11708. <https://doi.org/10.1038/ncomms11708> PMID: 27339440
42. Wu W, Huang Z, Li Z, Zhang S, Liu X, et al. (2016) De novo transcriptome sequencing of *Cryptotermes domesticus* and comparative analysis of gene expression in response to different wood species. *Gene* 575: 655–666. <https://doi.org/10.1016/j.gene.2015.09.055> PMID: 26410413
43. Wang L, Feng Z, Wang X, Wang X, Zhang X (2010) DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26: 136–138. <https://doi.org/10.1093/bioinformatics/btp612> PMID: 19855105
44. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28: 511–515. <https://doi.org/10.1038/nbt.1621> PMID: 20436464
45. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7: 562–578. <https://doi.org/10.1038/nprot.2012.016> PMID: 22383036
46. Gigolashvili T, Berger B, Flügge U-I (2009) Specific and coordinated control of indolic and aliphatic glucosinolate biosynthesis by R2R3-MYB transcription factors in *Arabidopsis thaliana*. *Phytochemistry reviews* 2009 v. 8 no.1: pp. 3–13.
47. Sønderby IE, Geu-Flores F, Halkier BA (2010) Biosynthesis of glucosinolates—gene discovery and beyond. *Trends in Plant Science* 15: 283–290. <https://doi.org/10.1016/j.tplants.2010.02.005> PMID: 20303821
48. Seo MS, Jin M, Sohn SH, Kim JS (2017) Expression profiles of BrMYB transcription factors related to glucosinolate biosynthesis and stress response in eight subspecies of *Brassica rapa*. *FEBS Open Bio* 7: 1646–1659. <https://doi.org/10.1002/2211-5463.12231> PMID: 29123974
49. Hirai MY (2005) Elucidation of gene-to-gene and metabolite-to-gene networks in *Arabidopsis* by integration of metabolomics and transcriptomics. *J Biol Chem* 280: 25590–25595. <https://doi.org/10.1074/jbc.M502332200> PMID: 15866872
50. Lachler K, Imhof J, Reichelt M, Gershenzon J, Binder S (2015) The cytosolic branched-chain aminotransferases of *Arabidopsis thaliana* influence methionine supply, salvage and glucosinolate metabolism. *Plant Mol Biol* 88: 119–131. <https://doi.org/10.1007/s11103-015-0312-3> PMID: 25851613
51. Zhang X, Liu T, Duan M, Song J, Li X (2016) De novo Transcriptome Analysis of *Sinapis alba* in Revealing the Glucosinolate and Phytochelatin Pathways. *Front Plant Sci* 7: 259. <https://doi.org/10.3389/fpls.2016.00259> PMID: 26973695
52. Seo MS, Kim JS (2017) Understanding of MYB Transcription Factors Involved in Glucosinolate Biosynthesis in Brassicaceae. *Molecules* 22.
53. Andersen TG, Halkier BA (2014) Upon bolting the GTR1 and GTR2 transporters mediate transport of glucosinolates to the inflorescence rather than roots. *Plant signaling & behavior* 9: e27740.
54. Kent WJ (2002) BLAT—the BLAST-like alignment tool. *Genome research* 12: 656–664. <https://doi.org/10.1101/gr.229202> PMID: 11932250
55. Javidfar F, Cheng B (2013) Construction of a genetic linkage map and QTL analysis of erucic acid content and glucosinolate components in yellow mustard (*Sinapis alba* L.). *BMC Plant Biol* 13: 142. <https://doi.org/10.1186/1471-2229-13-142> PMID: 24066707
56. Nelson MN, Lydiate DJ (2006) New evidence from *Sinapis alba* L. for ancestral triplication in a crucifer genome. *Genome* 49: 230–238. <https://doi.org/10.1139/g05-099> PMID: 16604105
57. Cheung WY, Landry BS, Raney P, Rakow GFW. Molecular Mapping of Seed Quality Traits in *Brassica juncea* L. Czern. and Coss; 1998. International Society for Horticultural Science (ISHS), Leuven, Belgium. pp. 139–148.
58. Ramchiary N, Bisht NC, Gupta V, Mukhopadhyay A, Arumugam N, et al. (2007) QTL analysis reveals context-dependent loci for seed glucosinolate trait in the oilseed *Brassica juncea*: importance of recurrent selection backcross scheme for the identification of 'true' QTL. *Theor Appl Genet* 116: 77–85. <https://doi.org/10.1007/s00122-007-0648-4> PMID: 17898985
59. Mahmood T, Ekuere U, Yeh F, Good AG, Stringam GR (2003) RFLP linkage analysis and mapping genes controlling the fatty acid profile of *Brassica juncea*; using reciprocal DH

- populations. *TAG Theoretical and Applied Genetics* 107: 283–290. <https://doi.org/10.1007/s00122-003-1244-x> PMID: 12669199
60. Xu Y, Zhang X-Q, Harasymow S, Westcott S, Zhang W, et al. (2018) Molecular marker-assisted back-crossing breeding: an example to transfer a thermostable β -amylase gene from wild barley. *Molecular Breeding* 38: 63.
 61. L. K, S. T, R. C, P. F, A. B, et al. TRANSFER OF A MAJOR QTL FOR RESISTANCE TO FUSARIUM HEAD BLIGHT FROM THINOPYRUM ELONGATUM ONTO DURUM WHEAT 7AL CHROMOSOME ARM AND ITS PYRAMIDING WITH OTHER USEFUL GENES FROM TH. PONTICUM; 2017; Pisa, Italy.
 62. Nour-Eldin HH, Andersen TG, Burow M, Madsen SR, Jorgensen ME, et al. (2012) NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds. *Nature* 488: 531–534. <https://doi.org/10.1038/nature11285> PMID: 22864417
 63. Bradley D, Ratcliffe O, Vincent C, Carpenter R, Coen E (1997) Inflorescence Commitment and Architecture in Arabidopsis. *Science* 275: 80–83. <https://doi.org/10.1126/science.275.5296.80> PMID: 8974397
 64. Ellerbrock B, Kim JH, Jander G (2007) Contribution of glucosinolate transport to Arabidopsis defense responses. *Plant Signal Behav* 2: 282–283. <https://doi.org/10.4161/psb.2.4.4014> PMID: 19704682
 65. Mirza N, Crocoll C, Erik Olsen C, Ann Halkier B (2016) Engineering of methionine chain elongation part of glucoraphanin pathway in *E. coli*. *Metabolic Engineering* 35: 31–37. <https://doi.org/10.1016/j.ymben.2015.09.012> PMID: 26410451
 66. Imhof J, Huber F, Reichelt M, Gershenzon J, Wiegrefe C, et al. (2014) The small subunit 1 of the Arabidopsis isopropylmalate isomerase is required for normal growth and development and the early stages of glucosinolate formation. *PLoS One* 9: e91071. <https://doi.org/10.1371/journal.pone.0091071> PMID: 24608865