National Institute for Applied Statistics Research Australia

University of Wollongong, Australia

Working Paper

11-21

Genetic Basis of Water Use Efficiency in Canola: Multi-Environment QTL Analysis Delineates a Major Locus Associated with Homoeologous Exchanges for Water-Use Efficiency and Seed Yield in Canola

Harsh Raman, Rosy Raman, Ramethaa Pirathiban, Brett McVittie, Niharika Sharma, Shengyi Liu, Yu Qiu, Anyu Zhu, Andrzej Killian, Brian Cullis, Graham D. Farquhar, Hilary S. Williams, Rosemary White, David Tabah, Andrew Easton, and Yuanyuan Zhang

National Institute for Applied Statistics Research Australia, University of Wollongong, Wollongong NSW 2522, Australia Phone +61 2 4221 5435, Fax +61 2 4221 4998.
Email: karink@uow.edu.au

Short title: Genetic basis of water use efficiency in canola

Multi-environment QTL analysis delineates a major locus associated with homoeologous exchanges for water-use efficiency and seed yield in canola

Harsh Raman[1*], Rosy Raman[1], Ramethaa Pirathiban[2], Brett McVittie[1], Niharika Sharma[3], Shengyi Liu[4], Yu Qiu[1], Anyu Zhu[5], Andrzej Killian[5], Brian Cullis[2], Graham D. Farquhar[6], Hilary S. Williams[6], Rosemary White[7], David Tabah[8] Andrew Easton[8] and Yuanyuan Zhang[4*]

[1]NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Wagga Wagga, NSW 2650, Australia
[2]Centre for Biometrics and Data Science for Sustainable Primary Industries, National Institute for Applied Statistics Research Australia, University of Wollongong NSW 2522, Australia
[3]NSW Department of Primary Industries, Orange Agricultural Institute, ORANGE, NSW 2800, Australia

[4]Oil Crops Research Institute, Chinese Academy of Agricultural Sciences, Wuhan, Hubei, 430062, China

[5] Diversity Arrays Technology P/L, University of Canberra, ACT 2601, Australia

[6]Research School of Biology, Australian National University, Canberra ACT 2601, Australia
[7]CSIRO, Canberra, ACT 2601, Australia
[8]Advanta Seeds Pty Ltd, 268 Anzac Avenue, Toowoomba, QLD 4350, Australia

Date of submission: 15th July 2021

Number of tables: 1

Figures: 6

Word count (start of the introduction to the end of the acknowledgements): 6475

Supplementary figures: 3

Supplementary tables: 15

Corresponding authors:

Harsh Raman

NSW Department of Primary Industries

Wagga Wagga Agricultural Institute, PMB, WAGGA WAGGA, NSW 2650, Australia

Tel: +61269381925

Email: harsh.raman@dpi.nsw.gov.au

Yuanyuan Zhang

Oil Crops Research Institute,

Chinese Academy of Agricultural Sciences, , Wuhan, Hubei 430062, China

Email: zhangyy@caas.cn

Summary

Canola varieties exhibit discernible variation in drought avoidance and drought escape traits, suggesting its adaptation to water-deficit environments. However, the underlying mechanisms are poorly understood. A doubled haploid (DH) population was analysed to identify QTL associated with water use efficiency (WUE) related traits. Based on the resequenced parental genome data, we developed sequence-capture based markers for fine mapping. mRNA-Seq was performed to determine the expression of candidate genes underlying QTL for carbon isotope discrimination ($\Delta^{13}C$). QTL contributing to main and QTL × Environment interaction effects for $\Delta^{13}C$ and for agronomic WUE were identified. One multi-trait QTL for $\Delta^{13}C$, days to flower, plant height and seed yield was identified on chromosome A09, in the vicinity of *ERECTA*. Interestingly, this QTL region was overlapped with a homoeologous exchange event (HE), suggesting its association with the major QTL. Transcriptome analysis revealed several differentially expressed genes between parental lines, including in HE regions. This study provides insights into the complexity of WUE related genes in the context of canola adaptation to water-deficit conditions. Our results suggest that alleles for high $\Delta^{13}C$ contribute positively to canola yield. Genetic and genomic resources developed herein could be utilised to make genetic gains for improving canola WUE.

Key words: water use efficiency, carbon isotope discrimination, drought avoidance, genetic analysis, physiology, yield

Introduction

Drought is the major abiotic stress that reduces the yield potential of various crops, especially in arid and semi-arid regions, of which 89% of regions are prevalent in Oceania (Koohafkan & Stewart, 2008). No doubt the impact of drought stress on crop productivity can be alleviated through irrigation at the 'critical' stages of plant development. However, in recent years fresh water, suitable for irrigation, is becoming scarce for crop production, required to meet the demand of a burgeoning human population (Gleick, 2000). Predicted climatic patterns such as debilitating drought and heat-wave episodes and their possible increased frequency further pose a significant threat to crop production (Smith & De Smet, 2012; Mills et al., 2018). The proportion of arable land per capita is also decreasing at a significant rate due to population growth and land degradation (http://www.fao.org/sustainability/). Therefore, improving crop varieties that have high yield potential and utilise water more effectively or require less water could provide a part of the solution to reduce the negative impacts of drought stress and increase productivity and food security (Passioura, 1977; Kijne et al., 2003; Blum, 2009; Bertolino et al., 2019; Leakey et al., 2019).

In nature, to cope with water-deficit conditions, plants have evolved different strategies such as drought escape, drought avoidance and drought tolerance (Levitt, 1980; Ludlow, 1989; Zhu et al., 2016; Rodrigues et al., 2019). Through tiny microscopic pores in the surface of leaves called stomata, plants assimilate CO_2 for photosynthesis by trading-off water, required for transpiration and other biological processes. This close intimacy between productivity and water use contributes to the adaptation of plants to their growing environments. Therefore, genetic variation in WUE and transpiration efficiency (TE, biomass production/transpirational water loss) that occurs as a result of intentional (via breeding/selection) and unintentional selection in nature provides an opportunity to identify and assemble useful alleles for improving the productivity of various crops.

WUE can be measured at the single leaf level as intrinsic WUE (*i*WUE), defined as the ratio of the photosynthetic CO_2 assimilation rate (A) over transpirational water loss (stomatal conductance, g_{sw}) or as whole-plant vegetative WUE, as the ratio of total dry matter production to total water transpired or as an integrated whole-plant WUE, as the

ratio of biomass or seed yield to evapotranspiration (Farquhar & Richards, 1984; Zhengbin et al., 2011; Leakey et al., 2019; Raman et al., 2019). *i*WUE assessments using the gas-exchange method are very challenging to be accurately performed, particularly in the large breeding populations, as WUE is regulated by a myriad of plant development, physiological, biochemical and molecular networks (Moore et al., 2009; Takahashi et al., 2018). Farquhar and Richards (1984) proposed $\Delta^{13}C$ as a time-integrated surrogate trait for measuring TE both at the single leaf level and at the whole plant level, as C_3 plants discriminate less against ^{13}C during photosynthesis with increased water deficit stress. The negative relationship between WUE and $\Delta^{13}C$ has been verified in *A. thaliana* (Masle et al., 2005) and in some agricultural crop plants (Farquhar et al., 1982; Ehleringer, 1993; Hall et al., 1994; Rebetzke et al., 2008; Des Marais et al., 2014; Raman et al., 2019), with some exceptions where nil or weak associations were observed (Hammer et al., 1997; Monneveux et al., 2007; Devi et al., 2011; Raman et al., 2020b).

Canola, being the second most important oilseed crop grown worldwide with a global production of 75 million tons (FAO STAT, http://www.fao.org/), is often cultivated in arid and semi-arid regions and faces periodic drought. Despite its economic significance to the oilseed industry as well as being an essential rotational crop in agricultural production systems, little research has been conducted on traits contributing to its drought tolerance (McVetty et al., 1989; Knight et al., 1994; Matus et al., 1995; Fletcher et al., 2015; Fletcher et al., 2016; Pater et al., 2017; Hossain et al., 2020; Raman et al., 2020a; Raman et al., 2020b). More recently, it was shown that two canola inbred lines, BC1329 and BC9102 vary by ~ 2‰ in their $\Delta^{13}C$ signatures (Hossain et al., 2020). However, the genetic basis of variation in $\Delta^{13}C$ and other integrated WUE traits such as plant biomass, flowering time and seed yield was not deciphered. Thus, a comprehensive understanding of the genetic and physiological bases underlying WUE is central to developing strategies for resilience to water deficit conditions.

Herein, through comprehensive analyses based on extensive phenotypic and physiological measurements, genetic and genomic studies, we demonstrate that multiple genetic and environmental determinants underlie plasticity in multi-dimensional drought avoidance traits such as $\Delta^{13}C$, early vigour, plant height and seed yield, and drought escape traits such as flowering time in canola. We also show that one of the QTL for multi-traits; $\Delta^{13}C$,

days to flower, plant height and seed yield on chromosome A09 is subjected to homoeologous recombination.

Materials and methods

Plant materials

In total, 223 doubled haploid (DH) lines derived from the F_1 cross between advanced breeding lines 'BC1329' (maternal parent) and 'BC9102' (paternal parent) were utilised for different genetic analysis experiments. An F_2 population comprising 744 lines derived from a single F_1 plant from BC1329/BC9102 was employed for fine mapping/verification of QTL associated with $\Delta^{13}C$.

Phenotypic evaluation for WUE traits

Four experiments were conducted in order to (i) determine the genomic regions that influence the expression of the traits associated with WUE (Experiments 1-3) and (ii) determine the relationship between $\Delta^{13}C$, *i*WUE and integrated WUE related traits under wet and dry conditions (Experiment 4). Experiments 1 and2 were performed under natural field conditions to measure WUE at the plot level; Experiment 3 is a pot experiment for single plant level WUE measurements and Experiment 4 is a rain-out shelter experiment with wet and dry irrigation regimes for measuring WUE at the single leaf level. Details of the experimental designs are presented in Table S4. Monthly weather statistics for average atmospheric temperatures and rainfall are also presented (Fig. S1).

Phenotypic trait measurements

Several agronomic, gas exchange and other physiological traits were measured for genetic analysis. A summary of the experiments in terms of their aim, trial layout, genetic material evaluated, and the traits measured are presented (Table S1, Fig. S2). Details of trait measurements are given in our recent study (Raman et al., 2020b) and summarised in Table S2. A brief description of the traits measured is given below.

Plant development and agronomic traits

$\Delta^{13}C$, flowering time, plant height and seed yield were measured for Experiments 1-4 and normalised difference in the vegetative index (NDVI) was measured only for Experiment 2. $\Delta^{13}C$ was determined from multi-phase experiments with appropriate experimental designs (Smith et al., 2006)) to account for the variations attributed to field/pot and laboratory conditions. The $\delta^{13}C$ composition was determined using Vienna Pee Dee Belemnite (VPDB) as the ultimate reference. $\Delta^{13}C$ was calculated from the $\delta^{13}C$ values assuming the isotopic composition of CO_2 in the air to be -7.8‰ on the VPDB scale, as described previously (Farquhar & Richards, 1984). Fresh and dry weights of the leaf and leaf thickness were also measured from F_2 plants and row plots under wet conditions in Experiment 4.

Physiological traits

The gas exchange measurements were taken at the single leaf level for the plots under wet conditions in Experiment 4, as this relationship varies under different water-deficit levels. We determined *i*WUE by measuring light-saturated assimilation rate (*A*) and stomatal conductance to the diffusion of water vapour (g_{sw}). The 5th fully expanded leaf of each of the 72 lines of BC1329/BC9102 DH population (06-5101DH), including parental lines, was tagged and utilised for gas exchange measurements.

Light microscopy

A leaf disc (9.08 cm^2 size) was taken from each of two replicate canola lines from Experiment 4 (wet block), fixed and stored in 70% ethanol as detailed (Table S4). Leaf sections were stained using a method modified from Rae *et al.* (2020) and were imaged using 488 nm excitation and 500-560 nm emission on a Leica SP8 confocal microscope.

Genotyping and linkage map construction

Genotyping of DH lines was carried-out using the genotyping-by-sequencing (GBS) based DArTseq approach (Raman et al., 2014). Sequence polymorphisms were used for linkage map construction following the method detailed in Raman *et al* (2016). The markers that showed complete segregation between each other were 'binned' into a unique locus and the resulting 'bin' map was used to identify trait-marker associations. To obtain the

physical position of markers, DArTseq sequences were aligned with the Darmor-*bzh* reference assembly version 4.1 using the default parameter settings with the Bowtie program.

Statistical methods

Commensurate with the aims of the experiments and the structure of the data sets, for Experiments 1-3 whole genome, single-step quantitative trait loci (QTL) analyses were performed on each trait using an extension of the approach developed by Verbyla and Cullis (2012) within a multi-environment trial (MET) analysis framework using factor analytic linear mixed models (FA-LMM) (Smith et al., 2015). Whereas, each trait measured on Experiment 4 is analysed individually using appropriate linear mixed models (LMM). A detailed description of the methods is presented (Table S4).

All analyses were performed in *ASReml-R* (Butler et al., 2018), which provides residual maximum likelihood (REML) estimates of variance parameters, empirical best linear unbiased predictions (EBLUPs) of random effects and empirical best linear unbiased estimates (EBLUEs) of fixed effects. The extent of genetic control of traits was investigated by calculating line mean H^2 (broad-sense heritability) as the mean of the squared accuracy of the predicted DH line effects as described previously (Cullis et al., 2006) and found to be dependent on the environment. The across environment summary measure of Overall performance (OP) proposed by Smith and Cullis (2018) was used to identify lines of interest. We examined the relationships of $\Delta^{13}C$ with agronomic traits (seed yield, days to flowering, plant height and NDVI) using pair-wise correlations of Overall performance estimates from the MET analysis of each trait or the EBLUPs from the LMM analysis of each trait.

Identification of candidate genes for WUE

Arabidopsis thaliana genes which had been annotated with various WUE-related terms were retrieved from the TAIR 10 database (https://www.arabidopsis.org/). These genes were then used to identify putative homologues in canola.

Resequencing and structural variation analysis of parental lines

Libraries from high-quality genomic DNA from both parental lines, BC1329 and BC9102, were constructed using the Illumina TruSeq DNA preparation kit, following the manufacturer's instructions (Illumina). Whole-genome resequencing (2 x 150 bp) was performed at the Novogene facility (Novogene Co., Ltd, Hong Kong) using the Illumina HiSeq 2000 sequencing platform. The coverage of the parental lines ranged from 77.6× (BC1329, 102.6 Gb) to 83.8× (BC 9102, 112.4 Gb). Read mapping to the 'Darmor-*bzh*' reference assembly (version 4.1, http://www.genoscope.cns.fr/brassicanapus/data/), SNP and InDel (< 50-bp) calling, structural variation (SV, ≥ 50-bp) detection and identification of HE event (≥ 10-kb windows) was performed as described in Raman *et al.* (2021).

Development of sequence-capture based DArTAg markers

We processed sequence data for target QTL regions on A09 and C09 chromosomes (Table S12) and selected 154 SNPs for DArTag oligo-synthesis. Oligos were synthesised by IDT (Ultramer DNA Oligos, http://idtdna.com) at 200 pmol scale, pooled in the equimolar amount into a single assay and used for processing 8 plates of DNA with the F_2 population and a control canola sample using a proprietary DArTag assay (Targeted Genotyping - Diversity Arrays Technology) using 384 plate format. For each plate, a sample of the pooled product was also run on agarose gel and compared against positive control before proceeding with the sequencing process. The libraries were sequenced on Illumina Hiseq2500 with an average volume of sequencing per sample at 43,225 sequencing reads (median at 46,389) and average read depth per assay at 280. Marker data were extracted using DArT PL's proprietary algorithm deployed a plugin in KDCompute application framework (https://www.kddart.org/kdcompute.html).

RNA sequencing and differential gene expression analysis

Parental lines, BC1329 and BC9102 of DH population were grown in three replicates under both wet (100% field capacity) and dry (50% field capacity) treatments in a glasshouse (Table S4). The clean sequence reads (100 bp single-end reads) for 12 samples that had per base sequence quality with >96% bases above Q30 were aligned against the *B. napus* reference Darmor-*bzh* (Version 4.1), using STAR aligner (v2.5.3a) (https://github.com/alexdobin/STAR/blob/master/doc/STARmanual.pdf). The raw counts

of reads mapping to each known gene was used to perform differential expression analysis using edgeR (version 3.30.3) (https://bioconductor.org/packages/release/bioc/html/edgeR.html) using R version 4.0.3. A generalised linear model approach was then used to quantify the differential expression between the groups. The differentially expressed genes (DEGs) were obtained using a false discovery rate (FDR < 0.05). Heatmaps showing the expression pattern of genes in A09 and C09 QTL regions were produced using the ComplexHeatmap R package (Gu et al., 2016).

Results

Substantial genetic variation in $\Delta^{13}C$ and other WUE traits

We observed high levels of genetic variation in $\Delta^{13}C$ and other WUE related traits in the DH mapping population. The significant source of genetic variation was from the additive component (genetic markers), which ranged from 21.5% for NDVI to 79.1% for days to flower (Table S5, Additive M1, %). Broad sense heritability estimates for $\Delta^{13}C$ and other integrated WUE related traits (plant height, NDVI, flowering time and seed yield) were variable, ranging from low (56%) to high (98%), depending on the nature of trait and growing environment (Table S6). Estimated additive and total (additive plus non-additive) genetic correlations between environments revealed that there are strong correlations between environments for both additive and total genetic variance with values greater than 0.89 and 0.83, respectively, for all traits (Table S7). Overall performance estimates for $\Delta^{13}C$ ranged from 18.73 to 21.25‰ and displayed transgressive segregation among DH lines across phenotypic environments (Fig. 1a, Table S8). Up to 2.52‰ variation in $\Delta^{13}C$ was observed among DH lines that equates to a 5-fold increase compared with the parental lines.

Relationships between WUE traits at plot level

To determine the relationships between $\Delta^{13}C$ and other WUE related traits, pair-wise correlations were obtained using the genotype Overall Performance estimates across environments (Fig. 1b). The $\Delta^{13}C$ showed a negative correlation with days to flower (r = - 0.58), while positive correlations were observed with NDVI, a proxy for plant vigour (r = 0.37), plant height (r = 0.45) and seed yield (r = 0.59). Flowering time showed a negative

correlation with seed yield (r = - 0.63). The promising DH lines that had high WUE of yield (high $\Delta^{13}C$) for use in canola breeding programs based on the Overall performance estimates are presented in Fig. 1c. DH line 06-5101-137 had the maximum $\Delta^{13}C$ (21.25‰) among the DH progenies.

Relationships between physiological WUE (single leaf level) and integrated WUE (whole plant level)

Significant variation for both *A* and g_{sw} was observed, although H^2 estimate of *i*WUE was low (Table S9). This may have occurred due to variable VPD across gas exchange measurements during the experiment, highlighting the plasticity of *i*WUE and $\Delta^{13}C$ as traits. Genotype EBLUPs for *A* and g_{sw} ranged from 4.97 to 17.15, and 0.11 to 0.38, respectively (Table S9). Pairwise correlations revealed that both *A* and g_{sw} are dependent on each other with a correlation of 0.56 (Fig. 2a). We observed a negative correlation between $\Delta^{13}C$ and *iWUE* (r = - 0.16), indicating that DH lines with low $\Delta^{13}C$ have higher *iWUE*, consistent with the findings made earlier (Farquhar & Richards, 1984). There was a more negative correlation between *iWUE* and g_{sw} (r = - 0.46) in comparison to *A* (r = - 0.27), suggesting that g_{sw} is the predominant driver for variation in *i*WUE parameters.

This study showed that $\Delta^{13}C$ correlates negatively with *i*WUE but it ($\Delta^{13}C$) correlates positively with seed yield (Fig. 2a). Under well-watered conditions, there were negative correlations between $\Delta^{13}C$ and days to flower, *A* and *iWUE*. We further investigated relationships between leaf water content (LWC) at a single leaf level and WUE traits at the whole plant level and found that LWC show a negative relationship with $\Delta^{13}C$, but it did not show any relationship with seed yield (Fig. 2a). Further, the estimated genetic correlations between wet and dry blocks for seed yield (Fig. 2b) and plant height, the only two traits measured after imposing water stress at the first flowering stage, were very high (0.93 for both traits). This suggests that genotype by irrigation block interaction is small. High $\Delta^{13}C$ lines revealed higher yield across irrigation blocks compared to low $\Delta^{13}C$ lines.

Genetic basis underlying $\Delta^{13}C$ and WUE related traits

We constructed a linkage map that includes 8,985 DArTseq markers onto 24 linkage groups (LGs), representing all the 19 chromosomes of *B. napus* (Table S10). To reduce computation time for genetic analysis, we produce a 'bin' map of 1793 markers that spanned a total of 1965.29 cM, with an average interval of 1.10 cM between adjacent loci.

Multi-environment QTL analysis identified a total of 29 QTL (15 QTL for main-effects and 14 for QTL (Q) × Environment (E) interactions) for variation in $\Delta^{13}C$ and other WUE related traits (Table 1, Table S11). For leaf $\Delta^{13}C$, three QTL main effects that showed statistically significant (LOD ≥3) associations were identified on chromosomes A08, A09 and C09, while one 'suggestive' QTL (LOD >2.5 but less than 3) was located on chromosome A07 (Table 1, Fig. 3a). We identified QTL for phenotypic plasticity in different traits between three growing environments (Q × E effects) on A02, A05, A08, A09, A10, C02, C03, C06, C07 and C09 chromosomes (Table S11). For $\Delta^{13}C$ plasticity, two QTL were identified on chromosomes A02 and C06, although the size of allelic effects were environment-dependent (Table S11). Collectively, QTL explained 38% of genotypic variation in $\Delta^{13}C$ (Table S5, VAF_m).

Comparative localisation of QTL

Three QTL for multi-traits on chromosomes A01, A08 and A09 were colocalised to the same genomic regions (Table 1). One QTL delimited with marker 3153720 for variation in $\Delta^{13}C$ was colocated with days to flower, plant height and seed yield on chromosome A09 (Table 1, Fig. 3a). We further sought a correlation between allelic effects of markers and variation in $\Delta^{13}C$, days to flower, plant height and seed yield (Fig. 3b-e). Up to 68% of allelic effects were explained by the same marker allele (Fig. 3e), suggesting pleiotropic relationships between these traits and/or tight genetic linkage between them.

Verification of QTL for $\Delta^{13}C$

We validated the genetic control, the linkage between DArTseq markers and $\Delta^{13}C$ (in DH population) and focused on the identification of candidate gene(s) underlying the majority of genetic variation in $\Delta^{13}C$ at QTL regions on chromosomes A09 and C09 (Table 1). The

$\Delta^{13}C$ values showed a wide range distribution among F_2 lines (Fig. 4a). Unlike DH lines, $\Delta^{13}C$ exhibited a positive correlation with flowering time and LWC, and a negative correlation with SLW (Fig. 4b-d). Our anatomical analysis of leaf discs that revealed both parental lines BC1329 and BC9102 differ in thickness and arrangement of palisade and spongy mesophyll cells: BC1329 (192 μm) had high porosity with large airspaces compared to BC9102 (184 μm, Fig. 4e-f), which may facilitate gas exchange, thus leading to efficient water use.

Genetic analysis revealed that several DArTag markers show significant segregation distortion (deviating from the normal segregation consistent with 1:2:1 ratio for codominance, or 3:1 ratio for dominance) on chromosomes A09 and C09 (Table S12), suggesting that the $\Delta^{13}C$ region could be subjected to structural variation. Genome scan using linear marker regression revealed that DArTag markers positioned at 28,598,612 bp on chromosome A09, and 46318271 bp on C09 of the Darmor-*bzh* genome exhibit statistically significant association with $\Delta^{13}C$ variation (Fig. S3).

Physical mapping and candidate genes associated with WUE near $\Delta^{13}C$ QTL

To identify potential candidate genes involved in the $\Delta^{13}C$ variation, we interrogated genomic regions underlying the significantly associated markers in both the mapping (DH) and validation populations (F_2). In the DH population, DArTseq 3153720 'bin' marker revealed the complete linkage with another 12 markers, which were localised within 1.49 Mb region, spanning 28.35 Mb to 29.35 Mb (Table S10, Fig. S3). Annotation of genomic interval revealed that several genes including *ERECTA* (BnaA09g40540D), *PYL2* (BnaA09g40690D), *H^+ATPase-5* (BnaA09g41340), *LEA18* (BnaA09g42180D) and *Protein Kinase* (BnaA09g42220D) on chromosome A09 and on its homoeologous chromosome C08, and *HAC11* (BnaC09g46960D), floral repressor *FLC (FLC.C09a*; *BnaC09g46500* and *FLC.C09b*; *BnaC09g46540D*), Myc-type BHLH (BnaC09g46950D, BnaC09g47080D on homoeologous group C09/A10 chromosomes are likely candidates to be involved in $\Delta^{13}C$ variation (Table S13, Fig. S3). DArTag marker (physical position on the Darmor-*bzh* genome: 28,598,612 bp) on chromosome A09 was located within 93 kb of the *ERECTA* gene that controls transpiration efficiency in *A. thaliana* (Masle et al., 2005).

$\Delta^{13}C$ QTL region on chromosome A09 is subjected to homoeologous exchange (HE)

We observed significant segregation distortion among marker alleles on chromosomes A09 and C09 in both mapping (DH) and validation (F_2) populations and inconsistency in collinearity across both genetic and physical maps (Table S12). To investigate whether QTL region on A09 is subjected to structural variation**,** we performed HE analysis utilising resequencing data of the parental lines. Sequence mapping revealed 26 genomic regions undergone HE events, varying from 90 kb to 870 kb, including the A09 multi-trait QTL region (29.3 to 29.5 Mb), BC9102 from C08 chromosome, as a result of homoeologous recombination (Fig. 5a, Table S14). However, in the maternal line BC1329**,** no such event was identified (Fig. 5b).

Gene expression changes for $\Delta^{13}C$ variation in the A09 and C09 QTL intervals between the parents

To investigate the expression of candidate genes that underlie the $\Delta^{13}C$ variation on chromosomes A09 and C09, we examined the leaf tissue-specific transcriptome of the two parental lines: BC1329 and BC9102 under wet and dry conditions. We found that a total of 60 genes on A09 and 51 genes on C09 underlying $\Delta^{13}C$ QTL regions were significanty differentially expressed between the two parental lines (Table S15). Of the DEGs, several of them such as Casein Kinase 2 α4 (BnaA09g42220D), Cation-transporting P-type ATPase (BnaA09g41340D, BnaA09g42040D), BEL1-like homeodomain protein 4 (BnaA09g41850D), Spermidine disinapoyl acyltransferase (BnaA09g41960D), Protein Kinase (BnaA09g42220D, BnaA09g41970D), HEC3 (BnaC09g46950D), and serine carboxypeptidase (BnaC09g47000D), are related with water use, water use efficiency and response to water stress (https://www.arabidopsis.org/). We also found that the expression levels of genes in BC9102 (with HE event) such as BnaA09g41850D, BnaA09g41970D (wall-associated receptor kinase-like 14), BnaA09g41990D (cyclin-dependent kinase inhibitor), BnaA09g42000D (nicotinate phosphoribosyltransferase 2), BnaA09g42030D (RNA recognition motif domain), and BnaA09g42040D were significantly higher (at least 2-fold) than those of BC1329 (without HE event) (Fig. 6, Table S15), suggesting that HE may be responsible for expression variation at the $\Delta^{13}C$-QTL region on A09.

Discussion

Canola reveals considerable variation for $\Delta^{13}C$

We found substantial genotypic variation in $\Delta^{13}C$, from 18.78 to 21.23‰ among DH, and 20.9 to 27.2‰ among F_2 lines. An earlier study has shown that an increase of 0.5‰ in $\delta^{13}C$ can lead to 25% more transpiration efficiency (TE = biomass gained/water transpired) in *Arabidopsis* (Juenger et al., 2005). Extrapolating this relationship, which is positive between $\delta^{13}C$ and TE, and negative between $\Delta^{13}C$ and TE, canola F_2 lines with 6.3‰ higher $\Delta^{13}C$ values than parental lines (22.8 to 23.5‰) should reduce WUE theoretically by 315%, which is impossible. It reflects the dependence of the sensitivity on the general level of $\Delta^{13}C$. For example, Masle et al. (2005) found that at the level they saw in *Arabidopsis*, an increase in $\Delta^{13}C$ of 1‰ was associated with a 15% decrease in TE. Previous studies revealed that canola lines display a range of variation in $\Delta^{13}C$ (18.7 to 23.7‰). Triazine tolerant (TT) accessions show higher $\Delta^{13}C$ values compared to conventional open-pollinated varieties and hybrids (Matus et al., 1995; Pater et al., 2017; Hossain et al., 2020; Raman et al., 2020b). In this study, we utilised non-TT accessions for genetic analysis. Our research thus provides an additional genetic resource for understanding the genetic and physiological basis, as well as improving WUE in canola.

Integrated WUE is partly driven by fitness traits

This study showed that DH lines that discriminate less between ^{12}C and ^{13}C as carbon source for photosynthesis (low $\Delta^{13}C$) show higher *i*WUE at the single leaf level (Fig. 2a). However, low Δ^{13} lines did not produce high yield (agronomic WUE; seed yield/unit of water used at the whole plot level) suggesting that selection for low *i*WUE at a single leaf level is useful for improving seed yield (r = 0.34, Fig. 2a), rather than using low $\Delta^{13}C$ as a surrogate trait for predicting high seed yield in canola, consistent with our earlier findings (Raman et al., 2020b). This inconsistent relationship between $\Delta^{13}C$ and seed yield could be due to genotypic variation in WUE being driven by variation in water use rather than by variation in assimilation per unit of water applied (Kobata et al., 1996; Blum, 2005; Sinclair, 2018). WUE, being a multi- dimensional trait can also be driven with other 'fitness' traits that reduce evapo-transpiration rate and crop water use. For example, high $\Delta^{13}C$ lines with faster growth (NDVI, a proxy for plant vigour and plant height) could

provide quicker canopy cover, which enables plants to reduce water loss from soil evaporation, thus increasing seed yield (r = 0.45 to 0.72, Fig. 1b). This is partly supported by in this study showing high correlation between plant fitness and seed yield and tight linkage of corresponding QTL (Fig. 1-3). In addition, $\Delta^{13}C$ exhibited negative correlations with flowering time (r= - 0.58; DH population), and a positive correlation with NDVI, plant height and seed yield (Fig. 1b), suggesting that high $\Delta^{13}C$ lines tend to 'escape' via accelerating growth and flowering - an evolutionary trait for adaptation to terminal drought stress. Our results showed that genotypes with low $\Delta^{13}C$ had less canopy cover, late flowering and lower seed yield; these characteristics are typical for plants with drought avoidance strategy (TE). However, under terminal water-deficit situations, low $\Delta^{13}C$ lines could yield poorly due to the shorter seed filling period, accompanied with high temperatures. It remains to establish how low $\Delta^{13}C$ lines which require a longer season for seed filling, perform in climates that are not prone to environmental constraints (non-water deficit/heat stress).

Genetic and environmental determinants affect phenotypic trait expression

We observed plasticity between $\Delta^{13}C$, and flowering time evaluated under field/pot and rain-out shelter (negative correlation, Fig. 1b, 2A) but a positive correlation under glasshouse conditions (Fig. 4b). This could be due to growing conditions (non- water stress condition, 100% field capacity) and nature of leaf tissue (discs without much vascular tissue) analysed for $\Delta^{13}C$.

Our comprehensive multi-environment QTL analysis showed that by using well-designed multiphase experiments (Table S3), and efficient statistical models (Table S4), both genetic and environmental determinants underpinning phenotypic variation can be deciphered for traits of interest (Table S11). For example, we identified QTL for the main effects (on A01, A07, A08 and A09) and Q x E interaction effects (on A02 and C06) that describe $\Delta^{13}C$ plasticity across different environments (Table S11). Multi-environment based QTL analysis is a more powerful approach to dissect complex traits than the traditional QTL approaches (Zhang et al., 2010) but it was not used to uncover the genetic basis of WUE traits in canola previously. Consistent detection of $\Delta^{13}C$-QTL across three environments suggests that these loci contribute to the adaptive capacity of DH lines to

water-deficit stress conditions and thus translating to economic seed yield (~1 t/ha). Across field environments, DH lines were subjected to water deficit conditions, right from stem elongation to seed maturity (rainfall ranged from 225 to 235 mm over seven months of growing season, Fig. S2). Colocation of QTL for seed yield, $\Delta^{13}C$ and plant height at the same genomic regions and stable allele (BC9102), contributing to trait variation that suggest multi-trait QTL on chromosome A09 are associated with effective water use. Early flowering showed a negative relationship with seed yield (Table 1), reiterating crosstalk between drought stress signalling and flowering time pathways (Des Marais et al., 2012).

It was interesting that none of the $\Delta^{13}C$ QTL that we identified for main effect and Q x E interactions (Table S11) were detected in the Skipton/Ag-Spectrum population (Raman et al., 2020b). In an independent study, Mekonnen *et al.*, (2020) identified three QTL for $\delta^{13}C$ on chromosomes A02, A09, and C08 in the North American *B. napus* mapping population. However, none of the QTL were consistently detected across environments. It is yet to establish whether the genomic region on chromosome A09 or its homoeologous counterpart C08 (QTL for root pulling force, plant height and $\delta^{13}C$) is the same as found in our study, as the authors did not report the physical positions of QTL marker-intervals. In addition, there was a poor marker coverage on chromosome C08 in our genetic mapping population (13 markers, Table S10), which may have led to QTL (if any) being undetected in the unmapped regions, especially in HE region. These studies suggest that several genomic regions on A02, A03, A07, A09, C03, C06, C08, and C09 control variation in $\Delta^{13}C$, thus, genetic architecture of $\Delta^{13}C$ is rather complex.

A priori genes regulating WUE and efficient water use underlie QTL for $\Delta^{13}C$

Coarse and high-resolution mapping approaches utilised herein facilitated the validation of genomic regions for Δ^{13} variation and delimited candidate genes in canola, which are implicated in leaf-level WUE (Hersen et al., 2008; Cutler et al., 2010; Youn et al., 2016; Tao et al., 2018; Menéndez et al., 2019). For example, this study identified and validated a QTL that influences multiple traits; $\Delta^{13}C$, days to flower, plant height and seed yield on chromosome A09 that map within 92 kb of the *ERECTA* gene (Table S13). In different plant species, *ERECTA* and *ERECTA Like 1,2* genes encoding leucine-rich repeat protein

kinases, regulate stomatal density and patterning, inflorescence architecture, ovule development, transpiration, and thermo-tolerance (Torii et al., 1996; Godiard et al., 2003; Shpak et al., 2003; Masle et al., 2005; Meng et al., 2012; Pillitteri & Torii, 2012; Bemis et al., 2013; Shen et al., 2015; Guo et al., 2020). *ERECTA* is also shown to control spikelet number- a component trait of grain yield via crosstalk between a Mitogen-activated protein kinase (MAPK) signalling pathway and cytokinin metabolism in rice. However, we did not find any difference in the level of expression of *ERECTA* between parental lines differing in Δ^{13} (unpublished data). We also localised several stress-responsive genes, including DEGs that may contribute to drought avoidance strategies via signal transduction pathways, encoding functional proteins (LEA18, RD20, glycine metabolism, CAT) and regulatory proteins, including transcription factors (bHLH, MYB, TINY2, ATHB6), protein kinases (Tyrosine protein kinase, Wall-associated receptor kinase-like 14, MAPK, SNF1-related protein kinase) and receptors (ABA receptor PYL12), phosphatases (PP2C), and calmodulins (CPK17) (Jonak et al., 2002; Des Marais et al., 2014; Jagodzik et al., 2018; Yong et al., 2019) within QTL intervals associated with Δ^{13}C variation (Table 1, Table S13, S15). Plant expressing PYL12, and SRK2C genes are shown to improve the water use and drought tolerance (Yang et al., 2016) whereas ABC transporter (ABCG22) and ABA responsive kinase gene, MPK12 reduced the WUE (Des Marais et al., 2014). Our data hint that genes affecting stomatal characteristics (RD20, ERECTA), leaf thickness and water-deficit responsive genes described above likely underlie WUE and drought avoidance traits, while Q x E interactions are likely driven by environmental cues (PHYTOCHROME C was mapped with 6.2 kb from Δ^{13}C-QTL on C06, Table S13).

Our results suggest that a QTL region underlying Δ^{13}C, flowering time, plant height and seed yield on chromosome A09 may be subjected to HE. Homoeologous recombination is associated with presence-absence variation (Nicolas et al., 2007; Hurgobin et al., 2018). Recently, a major QTL for homoeologous recombination, *BnaPh1* was mapped on A09 (Higgins et al., 2021) and this was located within 5 Mbp of the QTL region that is associated with multiple traits. It is possible that the same genomic region may be involved in regulating WUE in diverse canola accessions and require further research.

In summary, this current study demonstrates that measures of *i*WUE, Δ^{13}C and integrated WUE are complex and modulated by environmental and genetic determinants, including

those subject to homoeologous exchange. Our findings on identification of useful variation in $\Delta^{13}C$ (up to 6.3‰) and its underlying basis of variation in WUE traits, including their plasticity across environments, and identification of favourable alleles for increasing WUE would provide potential resources for developing new drought tolerant varieties for drier-environments to continue making genetic gains in the breeding programs.

Sequence data availability

The raw sequence data reported in this paper has been deposited in the National Center for Biotechnology Information Sequence Read Archive (Accession no. PRJNA743730 for RNA-Seq data, PRJNA743989 for whole-genome resequencing data).

Acknowledgements

This study was supported by the Australian Grains Research and Development Corporation and NSW Department of Primary Industries and partners (projects: DAN00117, and DAN00208. We thank Dr Simon Diffey (Apex Biometry) for multi-phase experimental designs for $\Delta^{13}C$ and Dr. Alison Smith (UOW) for constructive discussions on the statistical methods.The authors are grateful to Mr. Warren Bartlett and Mr. Dean McCullum, for their assistance in sowing and management of field experiments; Hannah Roe and Wayne Pitt for grinding leaf samples for $\Delta^{13}C$ analysis and Advanta for providing F_1 cross.

Author contributions

HR designed research; HR, RR, BM and YQ, performed field experiments; HSW analysed samples for carbon isotope discrimination; BC developed the statistical methods; RP, YZ, NS, HR, and SL analysed data; HR, NS, AZ and AK developed DArTags; BM, HR and GF performed/interpreted gas exchange measurements; RW investigated microscopic analysis; AE and DT provided seeds of F_1; HR prepared manuscript with inputs from others. All authors read and approved this manuscript for publication.

Reference

Bemis SM, Lee JS, Shpak ED, Tori KU. 2013. Regulation of floral patterning and organ identity by Arabidopsis ERECTA-family receptor kinase genes. *Journal of Experimental Botany, Vol. 64, No. 17, pp. 5323–5333, doi:10.1093/jxb/ert270*

Bertolino LT, Caine RS, Gray JE. 2019. Impact of stomatal density and morphology on water-use efficiency in a changing world. *Front Plant Sci* **10**: 225.

Blum A. 2005. Drought resistance, water-use efficiency, and yield potential—Are they compatible, dissonant or mutually exclusive? Aust. J. Agr. Res. 56:1159–1168.

Blum A. 2009. Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. Field Crops Research 112, 119–123.

Butler DG, Cullis BR, Gilmour AR, Gogel BJ, Thompson R. 2018. ASReml-R Reference Manual Version 4. Technical report, VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Cullis BR, Smith AB, Coombes NE. 2006. On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics 11, 381–393. doi: 10.1198/108571106X154443*.

Cutler SR, Rodriguez PL, Finkelstein RR, Abrams SR. 2010. Abscisic Acid: Emergence of a Core Signaling Network. *Annual Review of Plant Biology* **61**(1): 651-679.

Des Marais DL, Auchincloss LC, Sukamtoh E, McKay JK, Logan T, Richards JH, Juenger TE. 2014. Variation in *MPK12* affects water use efficiency in Arabidopsis and reveals a pleiotropic link between guard cell size and ABA response. *Proc Natl Acad Sci U S A* **111**(7): 2836-2841.

Des Marais DL, McKay JK, Richards JH, Sen S, Wayne T, Juenger TE. 2012. Physiological genomics of response to soil drying in diverse Arabidopsis accessions. *Plant Cell* **24**(3): 893-914.

Devi MJ, Bhatnagar-Mathur P, Sharma KK, Serraj R, Anwar SY, Vadez V. 2011. Relationships between transpiration efficiency and its surrogate traits in the rd29A:DREB1A transgenic lines of groundnut. *Journal of Agronomy and Crop Science* **197**(4): 272-283.

Ehleringer JR. 1993. Gas-exchange implications of isotopic variation in arid-land plants. In H Griffiths, J Smith, eds, Plant Responses to Water Deficit. BIOS Scientific Publishers, London, pp 265–284.

Farquhar GD, O'Leary MH, Berry JA. 1982. On the relationship between carbon isotope discrimination and intercellular carbon diooxide concentration in leaves. *Aust. J. Plant Physiol.* **9**: 121-137.

Farquhar GD, Richards RA. 1984. Isotopic composition of plant carbon correlates with water-use efficiency of wheat genotypes. Australian Journal of Plant Physiology 11, 539–552.

Fletcher RS, Herrmann D, Mullen JL, Li Q, Schrider DR, Price N, Lin J, Grogan K, Kern A, McKay JK. 2016. Identification of polymorphisms associated with drought adaptation QTL in *Brassica napus* by resequencing. *G3: Genes|Genomes|Genetics* **6**(4): 793-803.

Fletcher RS, Mullen JL, Heiliger A, McKay JK. 2015. QTL analysis of root morphology, flowering time, and yield reveals trade-offs in response to drought in *Brassica napus*. *Journal of Experimental Botany* **66**(1): 245–256, https://doi.org/210.1093/jxb/eru1423.

Gleick PH. 2000. Ed., In: The world water 2000-2001: The Biennial Report on Freshwater Resources (Island Press, 2000). p 53.

Godiard L, Sauviac L, Torii KU, Grenon O, Mangin B, Grimsley NH, Marco Y. 2003. ERECTA, an LRR receptor-like kinase protein controlling development pleiotropically affects resistance to bacterial wilt. *Plant J* **36**(3): 353-365.

Gu Z, Eils R, Schlesner M. 2016. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**(18): 2847-2849.

Guo T, Lu Z-Q, Shan J-X, Ye W-W, Dong N-Q, Lin H-X. 2020. *ERECTA1* acts upstream of the OsMKKK10-OsMKK4-OsMPK6 cascade to control spikelet number by regulating cytokinin metabolism in rice. *The Plant Cell*: tpc.00351.02020.

Hall AE, Richards RA, Condon AG, Wright GC, Farquhar GD. 1994. Carbon isotope discrimination and plant breeding. *Plant Breed. Rev. 12, 81–113*.

Hammer GL, Farquhar GD, Broad IJ. 1997. On the extent of genetic variation for transpiration efficiency in sorghum. *Australian Journal of Agricultural Research* **48**(5): 649-656.

Hersen P, McClean MN, Mahadevan L, Ramanathan S. 2008. Signal processing by the HOG MAP kinase pathway. *Proceedings of the National Academy of Sciences* **105**(20): 7165-7170.

Higgins EE, Howell EC, Armstrong SJ, Parkin IAP. 2021. A major quantitative trait locus on chromosome A9, *BnaPh1*, controls homoeologous recombination in *Brassica napus*. *New Phytologist* **229**(6): 3281-3293.

Hossain SM, Masle J, Easton A, Hunter MN, Godwin ID, Farquhar GD, Lambrides CJ. 2020. Genetic variation for leaf carbon isotope discrimination and its association with transpiration efficiency in canola (*Brassica napus*). *Funct Plant Biol* **47**(4): 355-367.

Hurgobin B, Golicz AA, Bayer PE, Chan C-KK, Tirnaz S, Dolatabadian A, Schiessl SV, Samans B, Montenegro JD, Parkin IAP, et al. 2018. Homoeologous exchange is a major cause of gene presence/absence variation in the amphidiploid *Brassica napus*. *Plant biotechnology journal* **16**(7): 1265-1274.

Jagodzik P, Tajdel-Zielinska M, Ciesla A, Marczak M, Ludwikow A. 2018. Mitogen-Activated Protein Kinase cascades in plant hormone signaling. *Front Plant Sci* **9**: 1387.

Jonak C, Okrész L, Bögre L, Hirt H. 2002. Complexity, cross talk and integration of plant MAP kinase signalling. *Curr Opin Plant Biol* **5**(5): 415-424.

Juenger TE, McKay JK, Hausmann N, Keurentjes JJB, Sen S, Stowe KA, Dawson TE, Simms EL, Richards JH. 2005. Identification and characterization of QTL underlying whole-plant physiology in *Arabidopsis thaliana*: δ13C, stomatal conductance and transpiration efficiency. *Plant, Cell & Environment* **28**(6): 697-708.

Kijne JW, Barker R, (Eds.) MD. 2003. Water productivity in agriculture: limits and opportunities for improvement. *Wallingford, UK: CABI; Colombo, Sri Lanka: International Water Management Institute (IWMI). xix, 332p. (Comprehensive Assessment of Water Management in Agriculture Series 1)*.

Knight JD, Livingston NJ, Van Kessel C. 1994. Carbon isotope discrimination and water-use efficiency of six crops grown under wet and dryland conditions. *Plant, Cell & Environment* **17**(2): 173-179.

Kobata T, Okuno T, Yamamoto T. 1996. Contributions of capacity for soil water extraction and water use efficiency to maintenance of dry matter production in rice subjected to drought. *Japanese Journal of Crop Science, 65(4), 652-662. https://doi.org/10.1626/jcs.65.652*.

Koohafkan P, Stewart BA. 2008. Water and Cereals in Dryland. *The Food and Agriculture Organisation of the United Nations and Earthscan, London, Sterling, VA*.

Leakey ADB, Ferguson JN, Pignon CP, Wu A, Jin Z, Hammer GL, Lobell DB. 2019. Water use efficiency as a constraint and target for improving the resilience and productivity of c3 and c4 crops. *Annual Review of Plant Biology* **70**(1): 781-808.

Levitt J. 1980. Responses of plants to environmental stresses. Volume II. Water, Radiation, Salt, and Other Stresses. Academic Press, New York.

Ludlow MM. 1989. Strategies in response to water stress, p. 269–281. In: H.K. Kreeb, H. Richter, and T.M. Hinkley (eds.). Structural and functional response to environmental stresses: Water shortage. SPB Academic Press, The Netherlands.

Masle J, Gilmore SR, Farquhar GD. 2005. The *ERECTA* gene regulates plant transpiration efficiency in Arabidopsis. *Nature* **436**(7052): 866-870.

Matus A, Slinkard A, van Kessel C. 1995. Carbon isotope discrimination: Potential for indirect selection for seed yield in canola. *Crop Sci.* **35**(5): 1267-1271.

McVetty PBE, Austin RB, Morgan CL. 1989. A comparison of the growth, photosynthesis, stomatal conductance and water use efficiency of *Moricandia* and *Brassica* species *Ann. Bot.* **64**: 87-94.

Mekonnen MD, Mullen JL, Arathi HS, Assefa Y, McKay JK, Byrne PF. 2020. Quantitative trait locus mapping for carbon isotope ratio and root pulling force in canola. *Agrosystems, Geosciences & Environment* **3**(1): e20095.

Menéndez AB, Calzadilla PI, Sansberro PA, Espasandin FD, Gazquez A, Bordenave CD, Maiale SJ, Rodríguez AA, Maguire VG, Campestre MP, et al. 2019. Polyamines and legumes: joint stories of stress, nitrogen fixation and environment. *Frontiers in Plant Science* **10**(1415).

Meng X, Wang H, He Y, Liu Y, Walker JC, Torii KU, Zhang S. 2012. A MAPK cascade downstream of ERECTA receptor-like protein kinase regulates arabidopsis inflorescence architecture by promoting localized cell proliferation. *The Plant Cell* **24**(12): 4948-4960.

Mills G, Sharps K, Simpson D, Pleijel H, Frei M, Burkey K, Emberson L, Uddling J, Broberg M, Feng Z, et al. 2018. Closing the global ozone yield gap: Quantification and cobenefits for multistress tolerance. *Global Change Biology* **24**(10): 4869-4893.

Monneveux P, Sheshshayee MS, Akhter J, Ribaut J-M. 2007. Using carbon isotope discrimination to select maize (Zea mays L.) inbred lines and hybrids for drought tolerance. *Plant Science* **173**(4): 390-396.

Moore JP, Le NT, Brandt WF, Driouich A, Farrant JM. 2009. Towards a systems-based understanding of plant desiccation tolerance. *Trends Plant Sci* **14**(2): 110-117.

Nicolas SD, Mignon GL, Eber F, Coriton O, Monod H, Clouet V, Huteau V, Lostanlen A, Delourme R, Chalhoub B, et al. 2007. Homeologous recombination plays a major role in chromosome rearrangements that occur during meiosis of *Brassica napus* haploids. *Genetics* **175**(2): 487-503.

Passioura JB. 1977. Grain yield, harvest index and water use of wheat. J. Aust. Inst.Agric. Sci. 43, 117–121.

Pater D, Mullen JL, McKay JK, Schroeder JI. 2017. Screening for natural variation in water use efficiency traits in a diversity set of brassica napus l. identifies candidate variants in photosynthetic assimilation. *Plant Cell Physiol* **58**(10): 1700-1709.

Pillitteri LJ, Torii KU. 2012. Mechanisms of stomatal development. *Annu Rev Plant Biol* **63**: 591-614.

Rae AE, Wei X, Flores-Rodriguez N, McCurdy DW, Collings DA. 2020. Super-resolution fluorescence imaging of *Arabidopsis thaliana* transfer cell wall ingrowths using pseudo-Schiff labelling adapted for the use of different dyes. *Plant and Cell Physiology 61: 1775–1787*.

Raman H, Raman R, Kilian A, Detering F, Carling J, Coombes N, Diffey S, Kadkol G, Edwards D, McCully M, et al. 2014. Genome-wide delineation of natural variation for pod shatter resistance in *Brassica napus*. *PLoS ONE 9(7): e101673. https://doi.org/10.1371/journal.pone.0101673* **9**(7): e101673.

Raman H, Raman R, Mathews K, Diffey S, Salisbury P. 2020a. QTL mapping reveals genomic regions for yield based on an incremental tolerance index to drought stress and related agronomic traits in canola. *Crop and Pasture Science 71(6) 562-577* **https://doi.org/10.1071/CP20046**.

Raman H, Raman R, McVittie B, Borg L, Diffey S, Yadav AS, Balasubramanian S, Farquhar GD. 2020b. The genetic and physiological basis for phenotypic variation in effective water use in *Brassica napus Food and Energy Security. https://doi.org/10.1002/fes3.237*.

Raman H, Raman R, Qiu Y, Zhang Y, Batley J, Liu S. 2021. The *Rlm13* gene, a new player of *Brassica napus-Leptosphaeria maculans* interaction maps on chromosome C03 in canola *Front. Plant Sci. | doi: 10.3389/fpls.2021.654604*.

Raman H, Uppal RK, Raman R 2019. Genetic solutions to improve resilience of canola to climate change. In: Kole C ed. *Genomic Designing of Climate-Smart Oilseed Crops*. Cham: Springer International Publishing, 75-131.

Raman R, Diffey S, Carling J, Cowley R, Kilian A, Luckett D, Raman H. 2016. Quantitative genetic analysis of yield in an Australian *Brassica napus* doubled haploid population. *Crop & Pasture Science* **67**(4): 298-307.

Rebetzke GJ, Condon AG, Farquhar GD, Appels R, Richards RA. 2008. Quantitative trait loci for carbon isotope discrimination are repeatable across environments and wheat mapping populations. *Theor Appl Genet* **118**(1): 123-137.

Rodrigues J, Inzé D, Nelissen H, Saibo NJM. 2019. Source–Sink Regulation in Crops under Water Deficit. *Trends in Plant Science* **24**(7): 652-663.

Shen H, Zhong X, Zhao F, Wang Y, Yan B, Li Q, Chen G, Mao B, Wang J, Li Y, et al. 2015. Overexpression of receptor-like kinase ERECTA improves thermotolerance in rice and tomato. *Nat Biotechnol* **33**(9): 996-1003.

Shpak ED, Lakeman MB, Torii KU. 2003. Dominant-negative receptor uncovers redundancy in the Arabidopsis ERECTA Leucine-rich repeat receptor-like kinase signaling pathway that regulates organ shape. *Plant Cell* **15**(5): 1095-1110.

Sinclair TR. 2018. Effective water use required for improving crop growth rather than transpiration efficiency. *Frontiers in Plant Science* **9**(1442).

Smith A, Ganesalingam A, Kuchel H, Cullis B. 2015. Factor analytic mixed models for the provision of grower information from national crop variety testing programmes. *Theoretical and Applied Genetics 128:55–72*.

Smith A, Lim P, Cullis B. 2006. The design and analysis of multi-phase plant breeding experiments. *Journal of Agricultural Science 144* **5**: 393-409.

Smith AB, Cullis BR. 2018. Plant breeding selection tools built on factor analytic mixed models for multi-environment trial data. *Euphytica 214(8), 143. https://doi.org/10.1007/s10681-018-2220-5*.

Smith S, De Smet I. 2012. Root system architecture: insights from *Arabidopsis* and cereal crops. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**(1595): 1441-1452.

Takahashi F, Kuromori T, Sato H, Shinozaki K. 2018. Regulatory gene networks in drought stress responses and resistance in plants. *Adv Exp Med Biol.* **1081:189-214**.

Tao Y, Chen M, Shu Y, Zhu Y, Wang S, Huang L, Yu X, Wang Z, Qian P, Gu W, et al. 2018. Identification and functional characterization of a novel BEL1-LIKE homeobox transcription factor GmBLH4 in soybean. *Plant Cell, Tissue and Organ Culture (PCTOC)* **134**(2): 331-344.

Torii KU, Mitsukawa N, Oosumi T, Matsuura Y, Yokoyama R, Whittier RF, Komeda Y. 1996. The Arabidopsis ERECTA gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *The Plant Cell* **8**(4): 735-746.

Verbyla AP, Cullis BR. 2012. Multivariate whole genome average interval mapping: QTL analysis for multiple traits and/or environments. *Theoretical and Applied Genetics 125, 933–953. https://doi.org/10.1007/s00122-012-1884-9*.

Yang Z, Liu J, Tischer SV, Christmann A, Windisch W, Schnyder H, Grill E. 2016. Leveraging abscisic acid receptors for efficient water use in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **113**(24): 6791-6796.

Yong Y, Zhang Y, Lyu Y. 2019. A MYB-related transcription factor from *Lilium lancifolium L. (LlMYB3)* is involved in anthocyanin biosynthesis pathway and enhances multiple abiotic stress tolerance in *Arabidopsis thaliana*. *Int J Mol Sci* **20**(13).

Youn H-S, Kim TG, Kim M-K, Kang GB, Kang JY, Lee J-G, An JY, Park KR, Lee Yj, Im YJ, et al. 2016. Structural insights into the quaternary catalytic mechanism of hexameric human quinolinate phosphoribosyltransferase, a key enzyme in de novo NAD biosynthesis. *Scientific Reports* **6**(1): 19681.

Zhang Y, Li Y-x, Wang Y, Liu Z-z, Liu C, Peng B, Tan W-w, Wang D, Shi Y-s, Sun B-c, et al. 2010. Stability of QTL across environments and QTL-by-environment interactions for plant and ear height in maize. *Agricultural Sciences in China* **9**(10): 1400-1412.

Zhengbin Z, Ping X, Hongbo S, Mengjun L, Zhenyan F, Liye C. 2011. Advances and prospects: biotechnologically improving crop water use efficiency. *Crit Rev Biotechnol* **31**(3): 281-293.

Zhu M, Monroe JG, Suhail Y, Villiers F, Mullen J, Pater D, Hauser F, Jeon BW, Bader JS, Kwak JM, et al. 2016. Molecular and systems approaches towards drought-tolerant canola crops. *New Phytol* **210**(4): 1169-1189.

Legends of Figures

Fig. 1: Genetic variation in WUE traits and their relationships among doubled haploid lines derived from the cross, BC1329/BC9102. **a:** Frequency distribution of the Overall performance estimates for $\Delta^{13}C$. Estimates for the parental lines are shown with arrows; **b:** Pair-wise correlations of the Overall performance estimates between $\Delta^{13}C$ (‰) and other WUE related traits; **c:** Top four DH lines that showed the highest $\Delta^{13}C$ based on Overall performance estimates across environments in relation to control commercial varieties of canola and the parental lines are shown.

Fig. 2: Relationships between $\Delta^{13}C$, gas exchange measurements (CO_2 assimilation (A), stomatal conductance (g_{sw}), and intrinsic water use efficiency (*i*WUE), plant developmental and agronomic traits (LWC: leaf water content; DTF: Days to flower; PH: Plant height and SY: Seed yield) of selected 70 DH lines of the BC1329/BC9102 population, representing extremes (High and low values) in $\Delta^{13}C$ and their parents. **a:** Pair-wise correlations of the genotype EBLUPs are plotted. DH lines were grown under rain-out shelter with wet and dry conditions. **b**: Relationships between $\Delta^{13}C$ and seed yield for wet and dry blocks. Genotype EBLUPs for $\Delta^{13}C$ and seed yield are plotted. Parental lines and the DH lines with high and low $\Delta^{13}C$ are labelled.

Fig. 3: Distribution and relationships between Overall performance estimates of $\Delta^{13}C$, days to flower (DTF), plant height (PH) and seed yield (SY) and DArTseq marker alleles for the QTL (3153720) that colocalized in the same genomic region on chromosome A09. Manhattan plot showing LOD scores for associations between DArTseq markers and $\Delta^{13}C$ (a). QTL main effects are labelled with the respective trait (for days to flower, plant height and seed yield only the 3153720 QTL is shown) and QTL x Environment interactions are labelled with the trait followed by 'Q × E' (only shown for $\Delta^{13}C$). LOD scores presented in the Manhattan plot are from the genome scan for the QTL main effects where the LOD scores of the significant QTL are replaced with the ones from the final model. The black dash line indicates the threshold value for significant SNPs at LOD ≥ 3. Box plots showing the distribution of the Overall performance estimates for $\Delta^{13}C$, days to flower, plant height and seed yield partitioned into allele combinations, 'AA (BC1329)' and 'BB (BC9102)',

for the SNP marker 3153720 (**b**). Pair-wise correlations of Overall performance estimates between $\Delta^{13}C$ vs days to flower (**c**), $\Delta^{13}C$ vs plant height (**d**) and $\Delta^{13}C$ vs seed yield (**e**) are partitioned into different allelic combinations.

Fig. 4: Distribution and relationships of the traits measured for an F_2 validation population derived from the BC1329/BC9102, grown under non-stress conditions. The frequency distribution of $\Delta^{13}C$ (‰) among 744 F_2 lines (**a**). Pair-wise correlations between $\Delta^{13}C$ and DTF (**b**), $\Delta^{13}C$ and LWC (**c**) and $\Delta^{13}C$ and SLW (**d**) are shown. $\Delta^{13}C$: Carbon isotope discrimination; DTF: Days to flower; LWC: Leaf water content; SLW: Specific leaf weight. Leaf sections showing differences in air spaces (AS, marked with arrow) between parental lines BC1329 (**e**) and BC9102 (**f**). EP: epidermis; PM: palisade mesophyll (comparatively regular elongated cells); SM: spongy mesophyll (irregular cells)

Fig. 5: Homoeologous exchange (HE) events detected between parental lines of doubled haploid population derived from the BC1329/BC9102. Genomic sequences that undergone HE are shown in Table S14. Substituted and 'translocated' reads are highlighted in Blue and Red colour, respectively.

Fig. 6: Expression profiles of differentially expressed genes (DEGs) in A09 (a) and C09 9b) QTL regions under water-deficit and water non-deficit conditions of the parental lines of the doubled haploid population derived from the BC1329/BC9102. The normalised read counts were plotted as a heatmap and genes were clustered according to the basis of their expression pattern. The genes in the heatmap were subjected to homoeologous exchange (HE) as well as the genes map within QTL region for $\Delta^{13}C$. DEGs that map within HE regions are highlighted in green boxes.

Table 1: Quantitative trait loci (main effects) for carbon isotope discrimination ($\Delta^{13}C$) and agronomic traits (DTF: Days to flower; NDVI: Normalised difference vegetative difference; PH: Plant height; SY: Seed yield) evaluated in doubled haploid lines from BC1329/BC9102, across three environments. LOD scores, allelic effect, parental allele and percentage of genetic variance explained (R^2) were also provided. QTL x Environment interactions for each environment are presented in supplementary Table S11. Putative candidate genes underlying QTL x Environment interactions are given in Table S13. Suggestive QTL having LOD ≤3 are in italics whereas consistent markers that were associated with multiple traits are in bold font.

Trait	**Marker**	**Chromosome**	**Physical map position of 'Top' marker on Darmor-*bzh* genome version 4.1**	**LOD**	**R^2 (%)**	**Allelic effect**	**Parental allele**	**Physical distance from putative candidate genes (Kb) within LD block**
NDVI	*3117901|F|0-11:T>C-11:T>C*	*A01*	*2276244*	*2.79*	*5.87*	*0.02*	*BC1329*	*Tyrosine-protein kinase-BnaA01g04900D* (3.15) *ASP5-BnaA01g04910D* (0.19)
SY	27390647|F|0-42:G>A-42:G>A	A01	3895510	4.17	8.37	-0.06	BC9102	Pentatricopeptide repeat-BnaA01g08190D (1.67)

								RmlC-like cupins-BnaA01g08200D (0.09)
DTF	**4106850|F|0-28:G>A-28:G>A**	**A01**	**3911345**	**4.18**	**12.73**	**1.27**	**BC1329**	**RmlC-like cupins-BnaA01g08210D (3.87) AMT1;4-BnaA01g08220D (7.33)**
PH	**4106850|F|0-28:G>A-28:G>A**	**A01**	**3911345**	**4.54**	**11**	**-2.64**	**BC9102**	
$\Delta^{13}C$	*3098654|F|0-12:G>A-12:G>A*	*A07*	*10641571*	*2.76*	*7.46*	*-0.09*	***BC9102***	*Auxin responsive SAUR protein- BnaA07g11440D (6.09)* *Dehydrin-BnaA07g11450D (0.68)*
PH	**3190876|F|0-10:T>A-10:T>A**	**A08**	**11767812**	**6.77**	**8.7**	**3.34**	**BC1329**	**SAM dependent carboxyl methyltransferase - BnaA08g13630D (0.61)** **Tetratricopeptide repeat-BnaA08g13640D (2.17)**
SY	**3190876|F|0-10:T>A-10:T>A**	**A08**	**11767812**	**4.84**	**10.76**	**0.06**	**BC1329**	

								Rab-GTPase-TBC domain-BnaA08g13650D (16.35) **Tyrosine-protein kinase-BnaA08g13660D (19.12)**
$\Delta^{13}C$	3087427|F|0-17:C>G-17:C>G	A08	13909331	4.57	10.86	0.14	BC1329	Tyrosine-protein kinase-BnaA08g17440D (7.75) Kelch-type beta propeller-BnaA08g17450D (2.73)
DTF	**3153720**	**A09**	**29356333**	**3.01**	**11.84**	**1.04**	**BC1329**	**LEA18- BnaA09g42180D (10.08) SDR3-BnaA09g42190D (5.64)** **Tyrosine-protein kinase-BnaA09g42220D (4.61)** **Embryo-specific 3-BnaA09g42230D (9.32)**
$\Delta^{13}C$	**3153720**	**A09**	**29356333**	**9.57**	**21.3**	**-0.19**	**BC9102**	
PH	**3153720**	**A09**	**29356333**	**6.91**	**10.09**	**-3.36**	**BC9102**	
SY	**3153720**	**A09**	**29356333**	**3.87**	**7.73**	**-0.05**	**BC9102**	

								ACA7- BnaA09g42240D (17.43)
DTF	*3148204|F|0-11:A>G-11:A>G*	*A10*	*12591509*	*2.87*	*15.32*	*-1.08*	*BC9102*	*NAD(P)-binding Rossmann-fold superfamily protein- BnaA10g16630D (1.10)* *UTP- BnaA10g16640D (0.52)* *Protein Kinase- BnaA10g16650D (5.20)* *GDH1- BnaA10g16660D (10.53)*
DTF	*4333486*	*C02*	*2170435*	*2.57*	*5.96*	*-0.92*	*BC9102*	*CYSTATIN - BnaC02g04270D (2.35)* *CLT3 - BnaC02g04280D (2.31)* *CPK17- BnaC02g04290D (2.83)* *Non-haem dioxygenase N-terminal domain- BnaC02g04300D (10.09)*

								DGS1- BnaC02g04310D (12.88)
NDVI	3140998|F|0-9:A>G-9:A>G	C06	18209584	4.46	7.27	0.03	BC1329	Armadillo-type fold-BnaC06g15500D (8.07) OMR1-BnaC06g15510D (22.46)
SY	27246553	C06	35071983	6.33	8.81	0.07	BC1329	GDSL-like lipase-BnaC06g36590D (11.75) SKP1-BnaC06g36600D (5.06) Galactose oxidase-BnaC06g36610D (3.40) Hap15-BnaC06g36630D (3.03)
PH	*3147080*	*C07*	*30323351*	*2.36*	*3.66*	*2.41*	*BC1329*	*Lipase-BnaC07g23920D (2.28)* *VH1-interacting kinase - BnaC07g23950D (28.05)*

DTF	*5053011|F|0-62:G>A-62:G>A*	*C08*	*22278532*	*2.06*	*7.12*	*-0.88*	*BC9102*	*CYCB2;3- BnaC08g19340D (23.99)* *SUMO- BnaC08g19350D (8.55)* *CAT3- BnaC08g19360D (2.48)* *RWP-RK- BnaC08g19370D (15.01)* *Tyrosine-protein kinase- BnaC08g19380D (27.23)*
$\Delta^{13}C$	3158874	C09	46623311	6.57	18.51	-0.17	BC9102	Hydroxyproline-rich glycoprotein- BnaC09g47070D (6.48) Myc-type, basic helix-loop-helix (bHLH)- BnaC09g47080D (0.72)

								Phosphate-induced protein 1-BnaC09g47090D (3.74) Epoxide hydrolase-like, alpha/beta-hydrolase-BnaC09g47100D (20.30)
DTF	3152507	C09_ran	3981494	4.21	14.02	1.39	BC1329	TIP1-BnaC09g54140D (0.45) Fatty acid synthase-BnaC09g54150D (10.95)

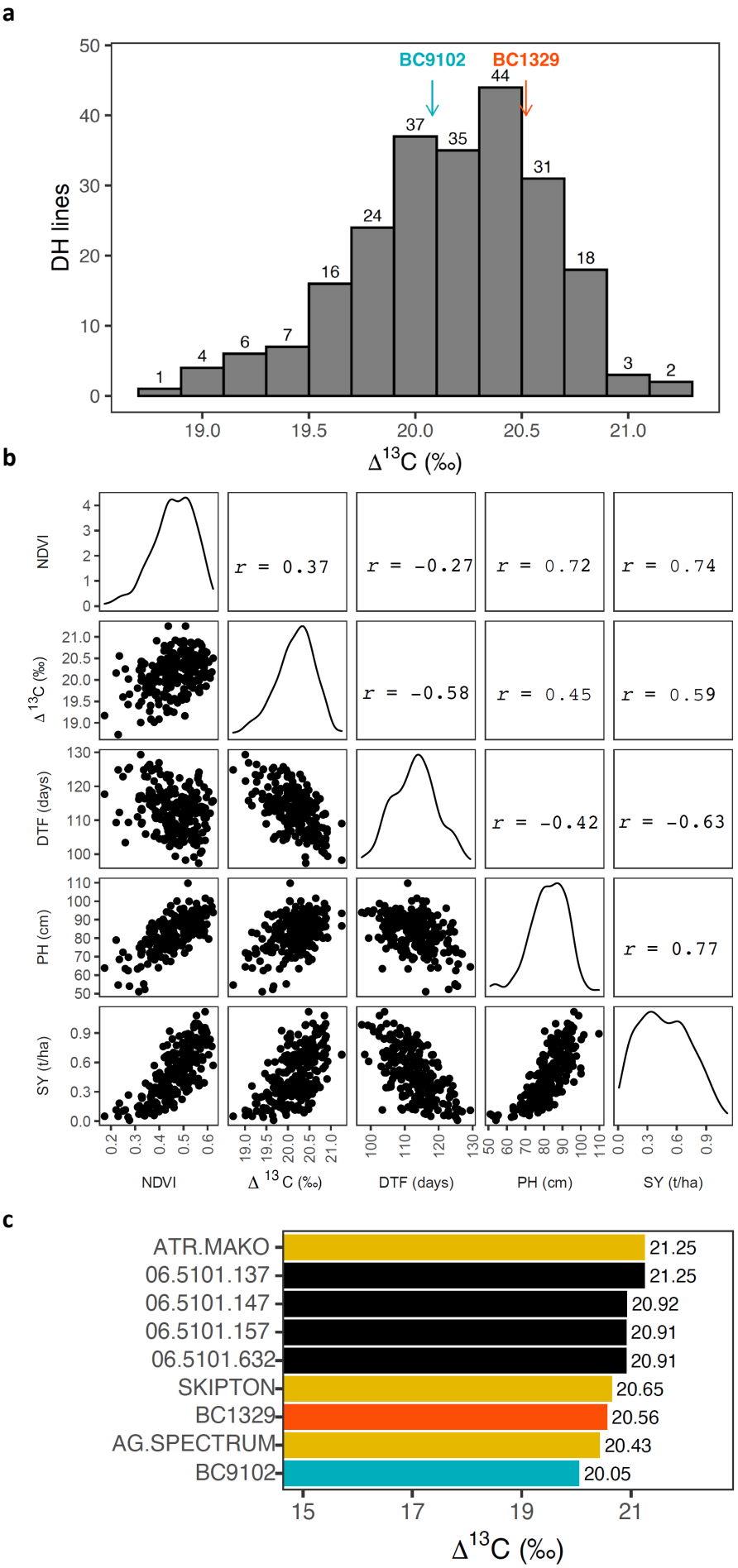
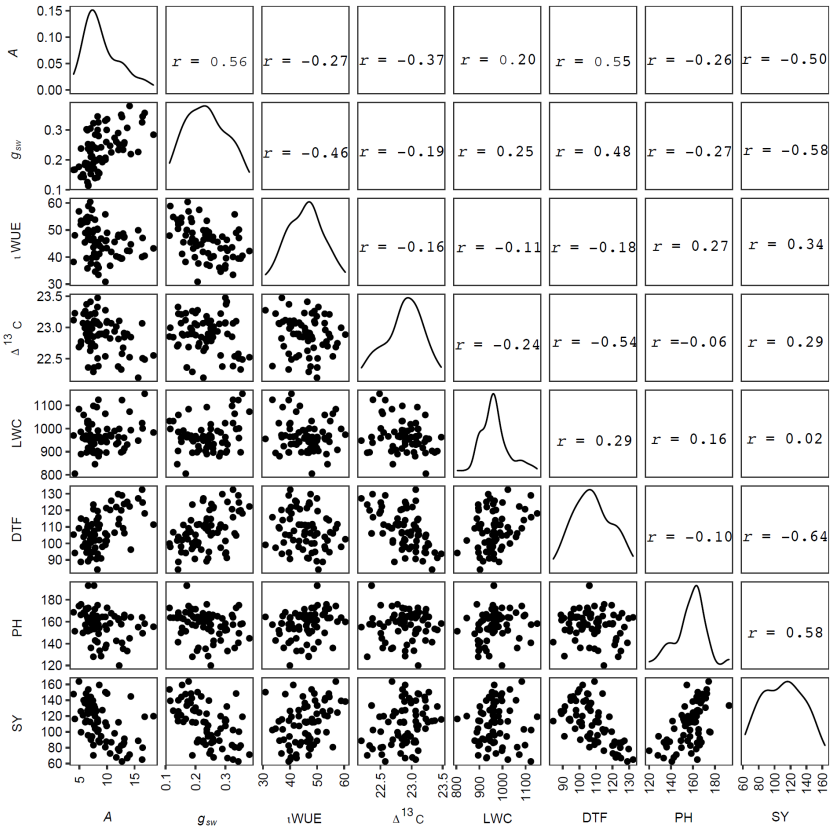

a
BC9102
BC1329
DH lines
Δ13C (‰)
b
NDVI
Δ 13C (‰)
DTF (days)
PH (cm)
SY (t/ha)
r = 0.37
r = -0.27
r = 0.72
r = 0.74
r = -0.58
r = 0.45
r = 0.59
r = -0.42
r = -0.63
r = 0.77
c
ATR.MAKO
06.5101.137
06.5101.147
06.5101.157
06.5101.632
SKIPTON
BC1329
AG.SPECTRUM
BC9102
21.25
21.25
20.92
20.91
20.91
20.65
20.56
20.43
20.05
Δ13C (‰)

Fig. 1

a


A
gsw
tWUE
Δ13C
LWC
DTF
PH
SY
r = 0.56
r = −0.27
r = −0.37
r = 0.20
r = 0.55
r = −0.26
r = −0.50
r = −0.46
r = −0.19
r = 0.25
r = 0.48
r = −0.27
r = −0.58
r = −0.16
r = −0.11
r = −0.18
r = 0.27
r = 0.34
r = −0.24
r = −0.54
r =−0.06
r = 0.29
r = 0.29
r = 0.16
r = 0.02
r = −0.10
r = −0.64
r = 0.58

b

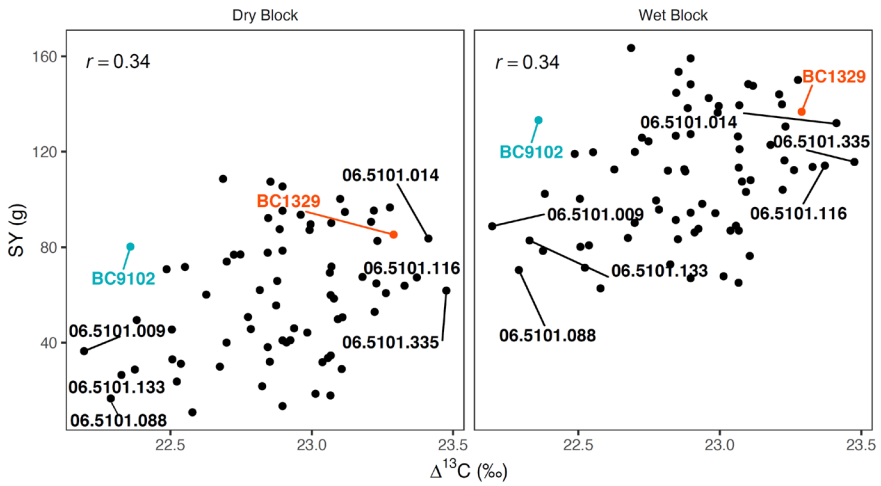

Dry Block
Wet Block
r = 0.34
r = 0.34
SY (g)
Δ13C (‰)
BC1329
BC9102
06.5101.014
06.5101.116
06.5101.335
06.5101.009
06.5101.133
06.5101.088

Fig. 2

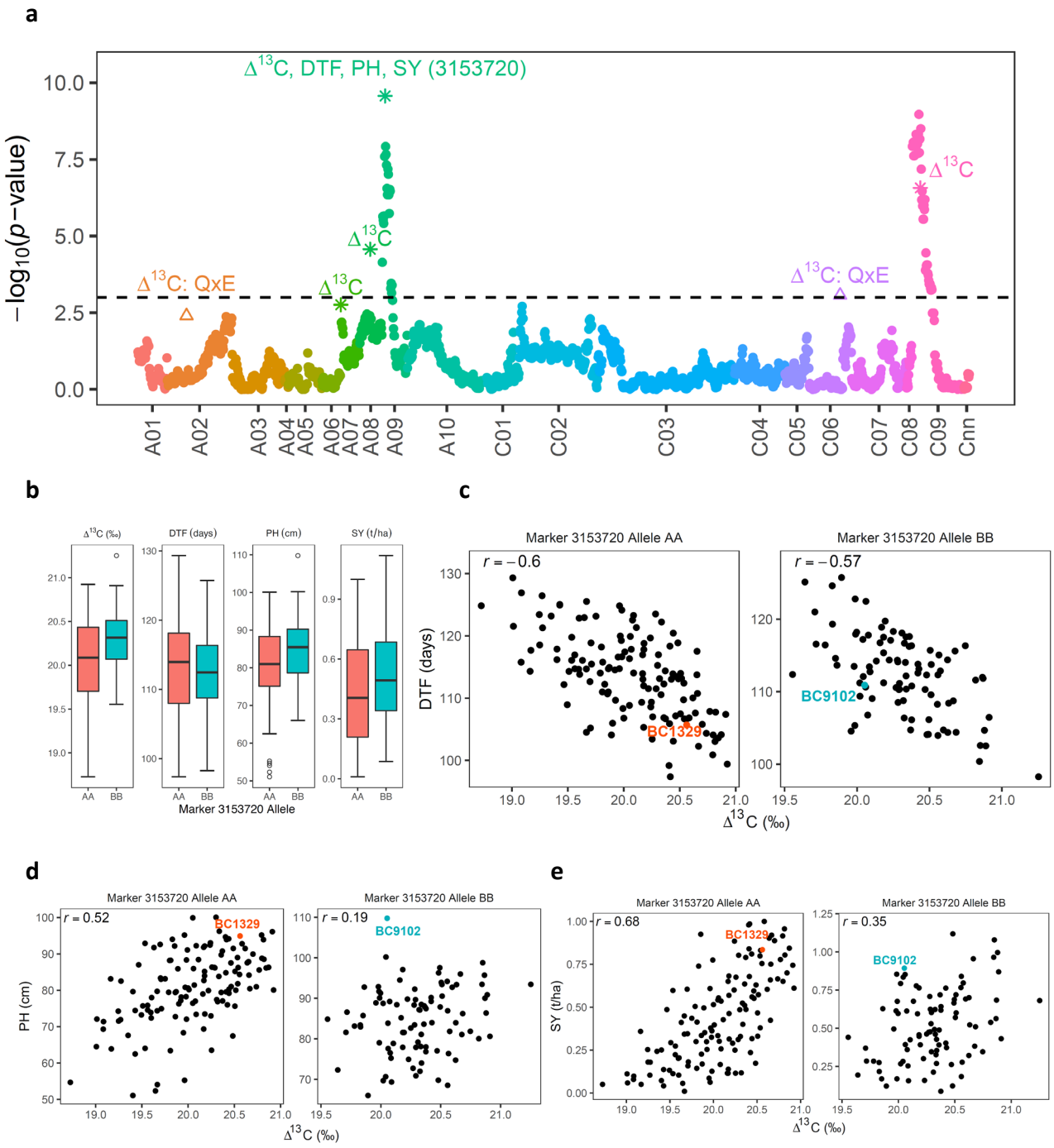

a
Δ13C, DTF, PH, SY (3153720)
Δ13C
Δ13C: QxE
Δ13C
Δ13C: QxE
Δ13C
−log10(p−value)
10.0
7.5
5.0
2.5
0.0
A01
A02
A03
A04
A05
A06
A07
A08
A09
A10
C01
C02
C03
C04
C05
C06
C07
C08
C09
Cnn
b
Δ13C (‰)
DTF (days)
PH (cm)
SY (t/ha)
AA
BB
Marker 3153720 Allele
c
Marker 3153720 Allele AA
Marker 3153720 Allele BB
r = −0.6
r = −0.57
BC1329
BC9102
DTF (days)
Δ13C (‰)
d
Marker 3153720 Allele AA
Marker 3153720 Allele BB
r = 0.52
r = 0.19
BC1329
BC9102
PH (cm)
Δ13C (‰)
e
Marker 3153720 Allele AA
Marker 3153720 Allele BB
r = 0.68
r = 0.35
BC1329
BC9102
SY (t/ha)
Δ13C (‰)

Fig. 3

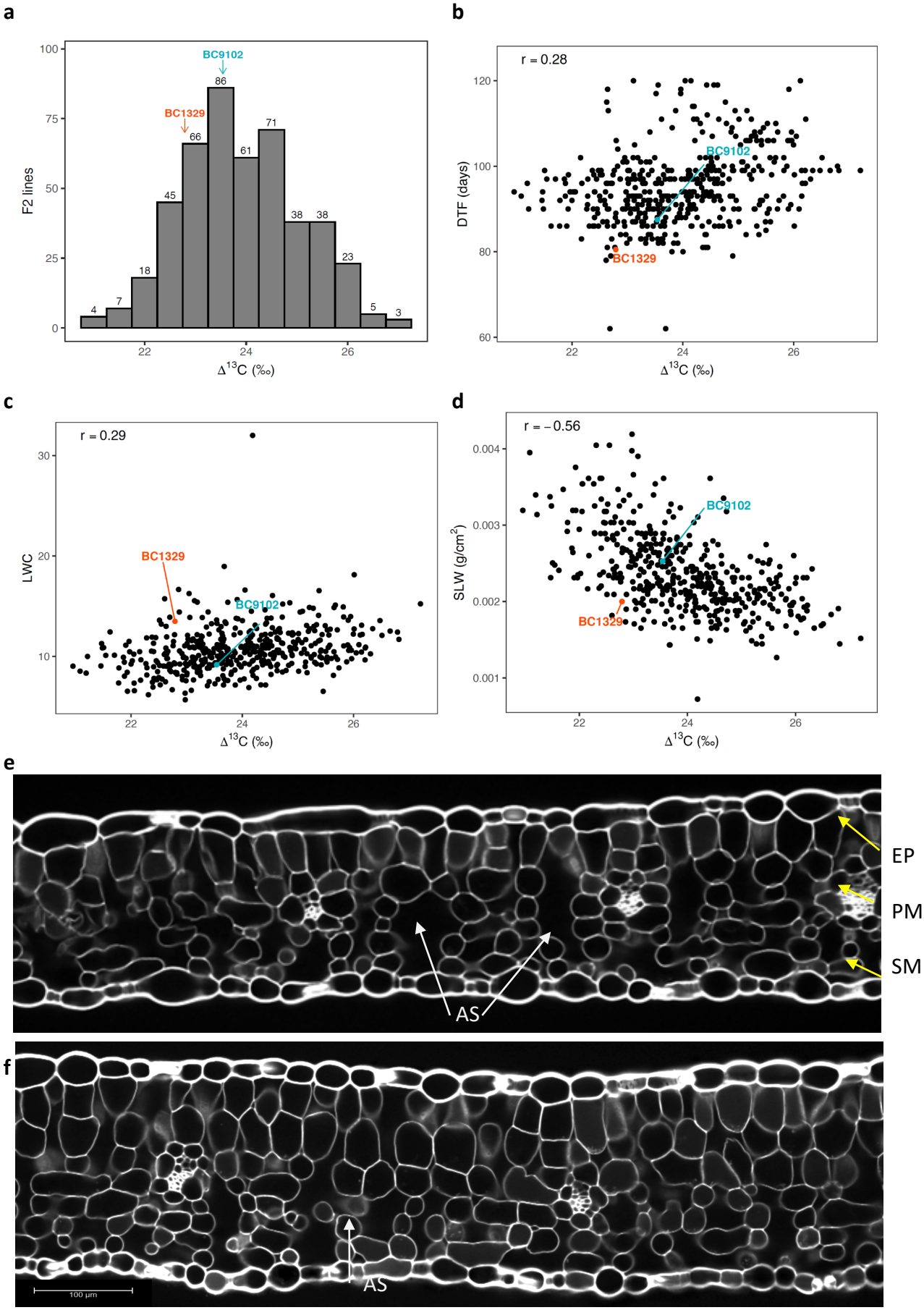

a
b
c
d
e
f
F2 lines
Δ13C (‰)
BC9102
BC1329
r = 0.28
DTF (days)
r = 0.29
LWC
r = −0.56
SLW (g/cm2)
EP
PM
SM
AS
100 µm

Fig. 4

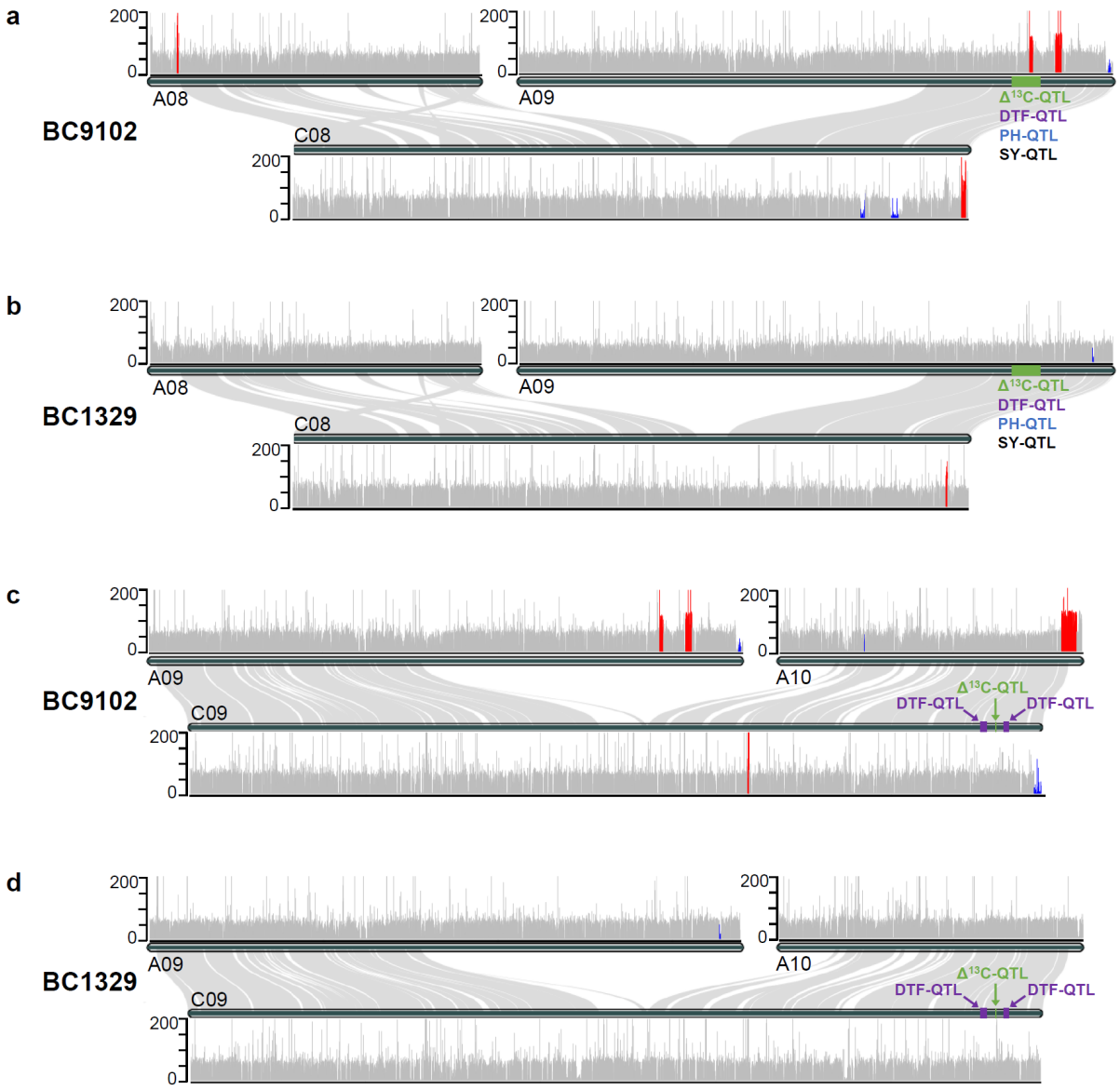

a
BC9102
A08
A09
C08
Δ13C-QTL
DTF-QTL
PH-QTL
SY-QTL
b
BC1329
A08
A09
C08
Δ13C-QTL
DTF-QTL
PH-QTL
SY-QTL
c
BC9102
A09
A10
C09
Δ13C-QTL
DTF-QTL
DTF-QTL
d
BC1329
A09
A10
C09
Δ13C-QTL
DTF-QTL
DTF-QTL
200
0

Fig. 5

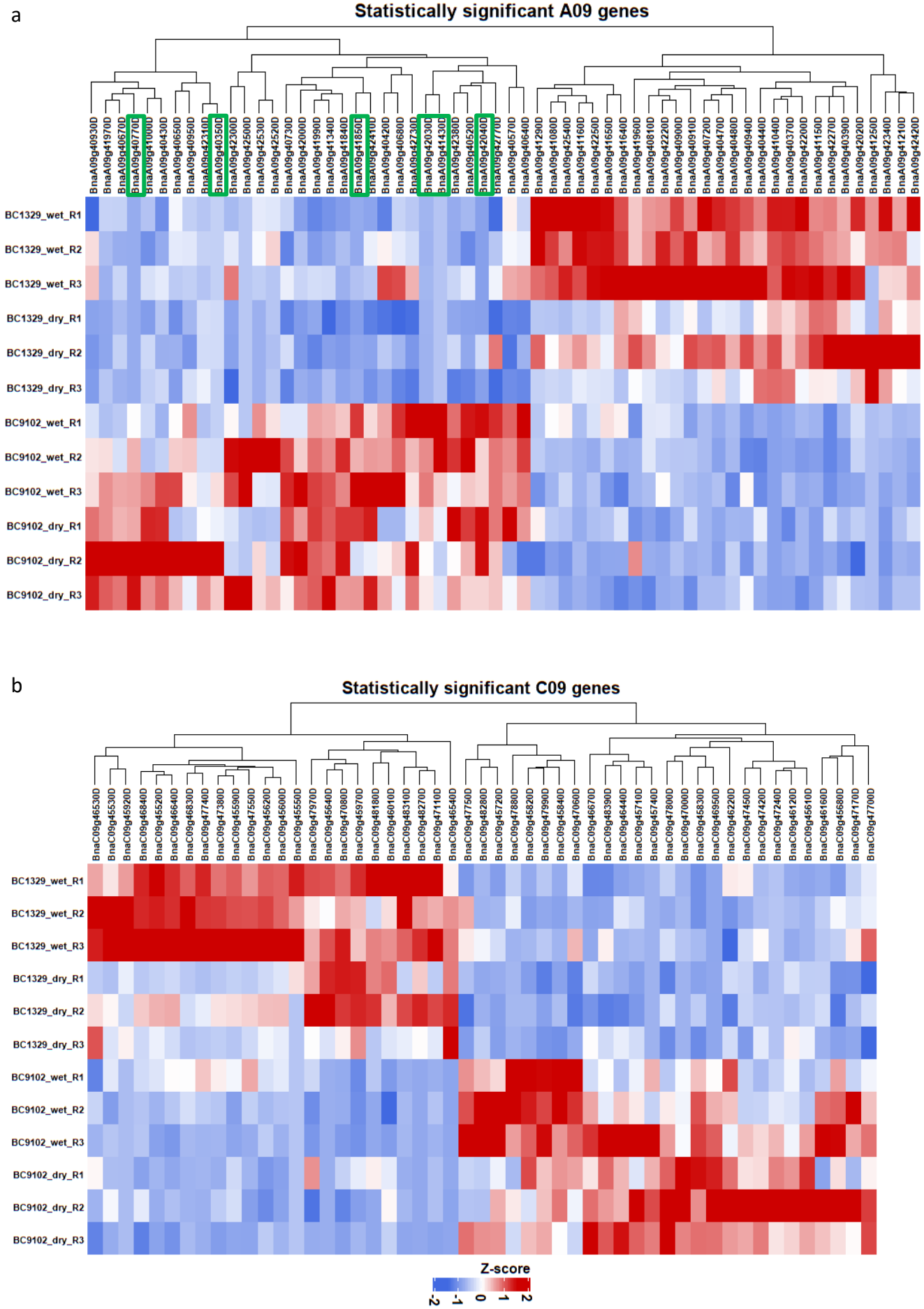

a
Statistically significant A09 genes
BnaA09g40930D
BnaA09g41970D
BnaA09g40670D
BnaA09g40770D
BnaA09g41000D
BnaA09g40430D
BnaA09g40650D
BnaA09g40950D
BnaA09g42310D
BnaA09g40350D
BnaA09g42300D
BnaA09g42500D
BnaA09g42530D
BnaA09g42520D
BnaA09g40730D
BnaA09g42000D
BnaA09g41990D
BnaA09g41340D
BnaA09g41840D
BnaA09g41850D
BnaA09g42410D
BnaA09g40420D
BnaA09g40680D
BnaA09g42730D
BnaA09g42030D
BnaA09g41430D
BnaA09g42380D
BnaA09g40520D
BnaA09g42040D
BnaA09g42770D
BnaA09g40570D
BnaA09g40640D
BnaA09g41290D
BnaA09g41080D
BnaA09g42540D
BnaA09g41160D
BnaA09g42250D
BnaA09g41650D
BnaA09g41640D
BnaA09g41960D
BnaA09g40810D
BnaA09g42220D
BnaA09g40900D
BnaA09g40910D
BnaA09g40720D
BnaA09g40470D
BnaA09g40480D
BnaA09g40940D
BnaA09g40440D
BnaA09g41040D
BnaA09g40370D
BnaA09g42200D
BnaA09g41150D
BnaA09g42270D
BnaA09g40390D
BnaA09g42020D
BnaA09g41250D
BnaA09g42340D
BnaA09g41210D
BnaA09g42420D
BC1329_wet_R1
BC1329_wet_R2
BC1329_wet_R3
BC1329_dry_R1
BC1329_dry_R2
BC1329_dry_R3
BC9102_wet_R1
BC9102_wet_R2
BC9102_wet_R3
BC9102_dry_R1
BC9102_dry_R2
BC9102_dry_R3
b
Statistically significant C09 genes
BnaC09g46530D
BnaC09g45530D
BnaC09g45920D
BnaC09g46840D
BnaC09g45520D
BnaC09g46640D
BnaC09g46830D
BnaC09g47740D
BnaC09g47380D
BnaC09g45590D
BnaC09g47550D
BnaC09g45620D
BnaC09g45600D
BnaC09g45550D
BnaC09g47970D
BnaC09g45640D
BnaC09g47080D
BnaC09g45970D
BnaC09g48180D
BnaC09g46010D
BnaC09g48310D
BnaC09g48270D
BnaC09g47110D
BnaC09g46540D
BnaC09g47750D
BnaC09g48280D
BnaC09g45720D
BnaC09g47880D
BnaC09g45820D
BnaC09g47990D
BnaC09g45840D
BnaC09g47060D
BnaC09g46670D
BnaC09g48390D
BnaC09g46440D
BnaC09g45710D
BnaC09g45740D
BnaC09g47800D
BnaC09g47000D
BnaC09g45830D
BnaC09g46950D
BnaC09g46220D
BnaC09g47450D
BnaC09g47420D
BnaC09g47240D
BnaC09g46120D
BnaC09g45610D
BnaC09g46160D
BnaC09g45680D
BnaC09g47170D
BnaC09g47700D
BC1329_wet_R1
BC1329_wet_R2
BC1329_wet_R3
BC1329_dry_R1
BC1329_dry_R2
BC1329_dry_R3
BC9102_wet_R1
BC9102_wet_R2
BC9102_wet_R3
BC9102_dry_R1
BC9102_dry_R2
BC9102_dry_R3
Z-score
-2
-1
0
1
2

Fig. 6

Table S1: Description of phenotyping experiments conducted to uncover genetic basis of carbon isotope discrimination ($\Delta^{13}C$) and water use efficiency related traits in canola (*Brassica napus* L).

Experiment	Research question	Material	Phenotyping environment	Blocks	Rows	Ranges	Plots	Traits measured
1	Genetic basis of CID variation	223 DH lines from BC1329/BC9102, plus parental lines	Field (2017)	2	45	10	450	$\Delta^{13}C$, days to flower, plant height, seed yield
2		223 DH lines from BC1329/BC9102, plus parental lines and commercial cultivars	Field (2018)	2	76	6	456	$\Delta^{13}C$, plant vigour (NDVI), days to flower, plant height, seed yield
3		217 doubled haploid lines from BC1329/BC9102, plus parental lines	Pot (2017)	4	73	12	876	$\Delta^{13}C$, days to flower, plant height
4	Relationship between physiological (intrinsic water use efficiency, CID) and agronomic water use efficiency (canola productivity) related traits	Selected 70 DH lines representing extreme (Low and High CID values) plus parental lines under wet and dry conditions	Rain-out shelter (2019)	2	48	9	432	$\Delta^{13}C$, days to flower, plant height, seed yield, Photosynthesis (*A*), stomatal conductance (g_{sw}), intrinsic water use efficiency (*i*WUE), Specific leaf weight (SLW), Leaf water content (LWC), Leaf thickness

Table S1: Trait measurement protocols for the plant development, agronomic and physiological traits measured in Experiments 1-4.

Trait	**Measurement protocol**
Plant development and agronomic traits	
$\Delta^{13}C$	For Experiments 1-3, leaf samples were taken from 10 plants of the plots at the vegetative stage (BBCH 40, prior to flowering) and processed as described earlier (Raman *et al.*, 2020b). For Experiment 4, leaf samples were taken only from the wet block. The $\Delta^{13}C$ was measured separately for each experiment and in different "runs", with each run consisting of 3 carousals and 49 samples in each carousal (including five standards at position 1st/2nd, 25th, and 48th/49th of the carousal). At least twenty percent of field plot/pot samples were duplicated for the measurement of $\Delta^{13}C$ in order to account for variation due to the laboratory process and field/pot variation.
Flowering time	Recorded as days to flower when 50% of plants in each plot/pot showed the first flower.
Plant height	At maturity, for Experiments 1-2 and 4, five plants were randomly selected from the middle of each plot from and were measured from the soil surface to the top of main branch, including siliques. For the Experiment 3, plant height was measured for all 5 plants in each pot.
Seed yield	For Experiments 1-2, seed yield was measured on plot basis. Plots were harvested with a small plot header and seed yield was expressed in ton/ha. For Experiment 4, seed yield was expressed in gm and the number of plants per row were counted.
NDVI	Measurements were taken periodically from each plot in Experiment 2 using the GreenSeeker® (NTech Industries Inc., Ukiah, CA, USA) following manufacturer's instructions.
Leaf thickness	This was carried out using 9.02cm^2 (area of leaf disc) cake cutters. Fresh and dry weights were measured.
Leaf water content (LWC)	Determined as the difference between fresh and dry weights, expressed as a percentage of dry weight. Leaves were dried at 80°C for 48 hr, until they reached constant mass. Measurements were made from the same samples that were used for leaf thickness.
Specific leaf weight (SLW)	Expressed as leaf dry weight/leaf area. Leaf area was determined from digital scans of flattened excised leaves using the ImageJ program using thresholding and the magic wand tool (http://www.imagej.nih.gov./ij/).
Physiological traits	
Light-saturated assimilation rate (Photosynthesis, *A*)	The 5th fully expanded leaf of a randomly selected plant in each plot in the wet block of Experiment 4 was tagged and gas exchange measurements were obtained using the gas exchange cuvette (LI-6400XT, LICOR Inc., Lincoln, NE, USA) between 10h and 16h (AEST) following conditions described in Raman et al (2020). The same 5th leaf that was used for gas exchange measurements was used for $\Delta^{13}C$ measurements.
Stomatal conductance to the diffusion of water vapour (g_{sw})	
*i*WUE (A/g_{sw})	Expressed as the ratio between *A* and g_{sw}.

Table S1: Details of multi-phase (field/pot and lab) experiments carried-out to measure $\Delta^{13}C$ in the doubled haploid population from BC1926/BC9102.

Experiment	Phenotypic environment	Field/Pot phase design							Lab phase design													
		Rows	Ranges	Plots/ Pots	Genotypes	Genotypes with *p* replicates			Samples for $\Delta^{13}C$					Runs	Carousals per Run							
									Rows	Ranges	Plots/ Pots	Field plots/Pots with *q* replicates		Total								
						p = 2	*p* = 4	*p* = 6				*q* = 1	*q* = 2			1	2	3	4	5	6	7	8
1	Field (2017)	45	10	450	225	225	0	0	45	10	447*	366	81	588	4	3	3	3	3	-	-	-	-
2	Field (2018)	76	6	456	228	228	0	0	76	4	304	256	48	392	3	3	3	2	-	-	-	-	-
3	Pot (2017)	73	12	876	219	0	219	0	73	12	862*	756	106	1078	8	3	3	3	3	3	3	3	1
4	Rain-out shelter (2019)	48	9	432	72	0	0	72	24	4	96	60	36	146	1	3	-	-	-	-	-	-	-

*Information not available for 3 field plots in Experiment 1 and 12 pots in Experiments 3.

The $\Delta^{13}C$ was measured separately for each experiment and in different "runs", with each run consisting of 3 carousals and 49 samples in each carousal (including five standards at position 1st/2nd, 25th, and 48th/49th of the carousal). At least twenty percent of field plot/pot samples were duplicated for the measurement of $\Delta^{13}C$ in order to account for variation due to the laboratory process and field/pot variation. Field plots/pots were resolvable to carousal, i.e. duplicated samples (which were from the same field plots/pots) were not allocated to the same carousal. The standards used in the lab phase were excluded from the data for analysis

Table S4. Supplementary materials

Experimental designs

(i) Experiments 1-2

The DH lines (n = 223), along with its parental lines (BC1329 and BC9102) and three commercial check cultivars (Skipton, Ag-Spectrum and ATR-Mako) were evaluated in two field experiments; Experiment 1 and Experiment 2 during the 2017 and 2018 canola growing seasons, respectively (autumn, April-November/December). Experiment 1 was sown on 18th May 2017 at the experimental station of Wagga Wagga Agricultural Institute (WWAG) (-35.0598782017231, 147.31173621965058) and Experiment 2 was sown on 14th May 2018 (-35.04540035261563, 147.34480619057837). Both experiments were conducted in rectangular arrays of plots; Experiment 1 with 45 rows by 10 ranges and Experiment 2 with 76 rows by 6 ranges. The trial designs for both experiments were randomised complete block designs (RCBD) with 2 row-wise blocks for Experiment 1 and 2 range-wise blocks for Experiment 2. The lines (treatments) were allocated to plots within blocks where the blocks were resolvable with 1 replicate plot of each line occurring in each block. The plots were of the same width of 2 m in both experiments, consisting 8 rows, spaced 25 cm apart. Lengths of the plots were 10 m in 2017 and 6 m in 2018. The seeds were sown at a depth of approximately 2.5 cm with Impact-in furrow fertiliser to protect against blackleg disease, caused by the *L. maculans*. The seeding rate was 1,400 seeds/plot for Experiment 1 whereas in Experiment 2 it was adjusted to 840 seeds for comparison with Experiment 1. Pre-emergent herbicides were used to control weeds.

(ii) Experiment 3

Experiment 3 is conducted in 2017 using plastic pots (size 15 cm diameter) arranged in a rectangular array of 73 rows by 12 ranges under an area covered by glass on top (-35.04722451361815, 147.33105481807226). Five seeds for each of the 217 DH lines, along with parental lines: BC1329 and BC9102 were sown on 6th June 2017 in the pots and raised to the full maturity stage (BBCH scale 95). The trial design was a randomized complete block design (RCBD) with 4 range-wise blocks. The lines (treatments) were allocated to plots within

blocks where the blocks were resolvable with 1 replicate plot of each line occurring in each block. The plants were watered using poly-pipe drip irrigation system (water flow rate: 3.8L/h emitters) and supplemented with liquid in-line fertiliser "Campbells Diamond Blue" (Campbells Fertilisers Australasia) as required. Plants received water three times a week for 15 minutes.

(iii) Experiment 4

To determine the relationship between $\Delta^{13}C$, *i*WUE (A/g_{sw}) and agronomic WUE related traits such as plant height and seed yield, extreme DH lines were selected from the tails of the population, comprising 35 top and 35 bottom lines on the basis of their seed yield across the field experiments. These 70 DH lines together with their two parental lines were evaluated in a rain-out shelter (15 m wide x 75 m long) under wet and dry conditions in 2019 at the WWAI (-35.04540035261563, 147.34480619057837). The experiment was conducted in two adjacent blocks, wet and dry, where both blocks were managed with the same protocol except for watering. Within each block, plots were arranged in rectangular arrays with 24 rows and 9 ranges. The treatments comprised the factorial combinations of irrigation regimes: wet (control) and dry (stress); and lines. The treatments were allocated to plots in such a way that each irrigation regime was allocated to all plots in a single block (for operational convenience) and lines were allocated to plots within blocks. There were three replicate plots of all lines per block.

A total of 25 seeds were sown in single row-plots (1 m length) with custom-built stand, ensuring proper spacing between plants. Initially, a pre-seeding irrigation of 16 mm of water was applied across both blocks for optimal and uniform germination. Water stress treatment was applied at the stem elongation stage (BBCH scale 50) to the dry block, while stress was not applied to the wet block. Soil moisture was tracked on-line throughout the experiments using 1.5 m moisture probes inserted in ground (ICT) and ensured that plants were subjected to water stress. Three moisture probes were installed in the experiment following the manufacturer's instructions: two in the dry block and one in the wet block; these probes were evenly spaced for reliable estimates of soil moisture across the blocks.

Statistical methods

(i) QTL analysis for the Experiments 1-3

The investigation of various traits on multiple trials required a multi-environment trial (MET) analysis for each trait. The MET data for each trait comprised the two field experiments (Experiments 1-2) and the pot experiment (Experiment 3) where each trial is considered as an "environment". A baseline linear mixed model (LMM) is formulated first and then it is extended to a factor analysis form for the genotype by environment (GE) effects between environments. The GE effects are then partitioned into additive and non-additive GE effects by incorporating the marker data into the model. A working model for QTL analysis is then developed and genome scan is performed on this model to identify genomic regions associated with each trait. The potential set of markers identified from the genome scan is then thinned in order to establish a final multi-QTL model.

Formulation of the baseline model

The baseline models for each trait are presented here using the symbolic model notation of Wilkinson and Rogers (1973) and *ASReml-R* (Butler et al., 2018) syntax.

<u>$\Delta^{13}C$</u>

```
fixed ~  1 + Environment + Environment:Gdrop
random ~ diag(Environment):Gkeep +
               at(Environment):Run +
               at(Environment):Carousal +
               at(Environment):Run:Carousal +
               at(Environment):Carousal:Tube +
               at(Environment):Block +
               at(Environment):Block:Plot
```

<u>Days to flower</u>

```
fixed ~  1 + Environment + Environment:Gdrop
random ~ diag(Environment):Gkeep +
```

```
                at(Environment):Block +
                at(Environment):Row +
                at(Environment):Range
```

Plant height

```
fixed ~  1 + Environment + Environment:Gdrop
random ~ diag(Environment):Gkeep +
                at(Environment):Block +
                at(Environment):Block:Plot +
                at(Environment):Row +
                at(Environment):Range
```

Seed yield

```
fixed ~  1 + Environment + Environment:Gdrop
random ~ diag(Environment):Gkeep +
                at(Environment):Block +
                at(Environment):Row +
                at(Environment):Range
```

NDVI

```
Fixed ~ 1 + Gdrop
random ~ Gkeep + Block + Row + Range
```

There are many characteristics of the above symbolic model notation that require explanation and defining. Firstly, formulation of the baseline model commenced with fitting a model that assumed independence of the GE effects between environments. This model termed the *diagonal (DIAG)* variance model for the GE effects (`diag(Environment):Gkeep`, where `Gkeep` is the lines that were genotyped) is analogous to analysing each environment separately. This baseline model is used to assess whether additional terms are required to account for non-treatment sources of variation as well as investigating the presence of outlier observations. Two additional terms were fitted as fixed effects, representing the main effects of the environments (`Environment`) and the interaction of the environments with those

lines which were not genotyped but had phenotypic data (`Environment:Gdrop`, where `Gdrop` is the lines that were not genotyped) (see Tolhurst et al. (2019) for details). All models also included a term representing the overall mean (`1`) as a fixed effect.

The non-genetic component of the baseline model required the inclusion of random terms, which represent the plot structure of the randomized complete block designs of the experiments (`Block`), as well as accounting for other significant sources of non-genetic variation which occurred in either field/pot experiments (random `Row` or `Range` terms) or in the lab phase (see **Error! Reference source not found.**) for the trait $\Delta^{13}C$ (`Run`, `Carousal`, `Run:Carousal` and `Carousal:Tube`) (see Smith et al. (2006) and Gilmour et al. (1997) for example). The working model also included terms (`Block:Plot`) which partition residual variance for those traits where the experimental unit is not the observational unit (see Bailey (2008)). These traits were plant height and $\Delta^{13}C$. The term `at(Environment)` denotes that variance models for these random model terms allowed for variance heterogeneity between environments. The variance models for the residuals were either independent and identically distributed (iid) effects (for $\Delta^{13}C$ and plant height), or a two dimensional separable first order autoregressive variance model as explained in Cullis and Gleeson (1991) (for days to flower, seed yield and NDVI) with different variance parameters for each environment.

Factor analytic variance models for the GE effects

The baseline LMM for all traits except NDVI were then extended to include factor analytic (FA) variance models for the GE effects. A MET analysis was not necessary for NDVI as this trait was measured only in one environment (Experiment 2). The FA-LMM estimates the genetic variances and covariances between environments using a small number of unknown factors. These models were fitted using the reduced rank form of the FA model introduced by Thompson et al. (2003). Using the symbolic model notation of Wilkinson and Rogers (1973) and *ASReml-R* syntax, we fit

```
fixed ~ 1 + Environment + Environment:Gdrop
random ~ rr(Environment, k): Gkeep +
   diag(Environment): Gkeep
   …
```

where `rr` component models the common genotype by environment (CGE) effect, `diag` component models the specific genotype by environment (SGE) effect, `k` is the order of the model representing the number of unknown factors and "…" denotes other non-genetic terms. The FA modelling process commences with one factor (`k` = 1) and continues until either the limit of the data is reached or the overall percentage variance accounted for reaches 80%. For example, the limit for a MET dataset with $p = 3$ environments is `k` = 1 factor because an increase in the number of factors will result in more parameters being estimated than are possible in the fully unstructured model ($p(p + 1)/2 = 6$). For the FA1 model, there are $pk + p - k(k - 1)/2 = 6$ parameters to estimate.

In our case, an FA model of order 1 (FA1) was fitted for all traits except for seed yield, for which the variance of the SGE effects of one of the environments was constraint to be zero following Cullis & Smith (2016) as there were only two environments for this trait (Experiments 1-2). We refer to these models as baseline FA-LMM. On average, the percentage of genetic variation (%VAF) accounted by the baseline FA-LMM were above 86% for all traits ($\Delta^{13}C$: 88.6%, seed yield: 86.6%, days to flower: 92.3% and plant height: 89.9%). Estimated genetic correlations between environments were all greater than 0.83 for all traits.

Partitioning GE effects into additive and non-additive GE effects

The baseline FA-LMM was then extended to include marker information in order to determine the genomic regions that influence the traits associated with WUE. GE effects were partitioned into additive and non-additive GE effects (Oakey et al., 2007) and each were modelled by separate FA variance models. We utilised a genetic linkage map of 5101 DH population based on 1793 'bin' DArTseq markers representing all 19 chromosomes of *B. napus*. The linkage map was constructed using the package *ASMap* (Taylor and Butler, 2014) in *R* statistical computing environment (R Core Team, 2019) utilising the Minimum spanning tree algorithm (Wu et al. 2008). The missing values in these markers were imputed using the *k*-nearest neighbour method (Troyanskaya et al., 2001). The resulting marker matrix, M, is used to construct a genomic relationship matrix, K, (GRM) (VanRaden, 2008) where $K = MM^T$ using the package *pedicure* (Butler, 2019) in *R* statistical computing environment (R Core Team, 2019). Using the symbolic model notation of Wilkinson and Rogers (1973) and *ASReml-R* syntax, we fit

```
fixed ~ 1 + Environment + Environment:Gdrop
random ~ rr(Environment, 1):vm(Gkeep, K)+
   diag(Environment):vm(Gkeep, K) +
   rr(Environment, 1):ide(Gkeep) +
   diag(Environment):ide(Gkeep) +
   …
```

where `rr(Environment, 1):vm(Gkeep, K)` models the additive CGE effects, `diag(Environment):vm(Gkeep, K)` models the additive SGE effects, `rr(Environment, 1):ide(Gkeep)` and `diag(Environment):ide(Gkeep)` models the same but for non-additive effects, `K` is the GRM and associated with the additive effects and "…" denotes other non-genetic terms. This model is equivalent to the standard genomic best linear unbiased prediction (GBLUP) model used in models for genomic selection (Tolhurst et al., 2019). We refer to this model as baseline FA-LMM with markers.

Both the between environment variance matrices of the GE additive genetic effects and non-additive genetic effects were modeled by FA structures of order 1 (FA1:FA1) for all traits except for seed yield, for which the variance of the additive and non-additive SGE effects of one of the environments was constraint to be zero as before. Table 1 presents a summary for baseline LMM, baseline FA-LMM and the baseline FA-LMM with markers in terms of the variance model for the GE effects, number of estimated parameters (total and genetic), residual log-likelihood and Akaike Information Criterion (AIC). The AIC value of the baseline FA-LMM with markers (or baseline LMM for NDVI) is the lowest for each trait, indicating that this model is the better fit.

Summary of the baseline models fitted for each trait with the variance model for genotype by environment (GE) effects, number of parameters estimated in each model in total (Total) and for the genetic variance (Genetic), log-likelihood (LogLik) and Akaike Information Criterion (AIC). DIAG: Diagonal; FA1: Factor analytic structure of order 1; FA1:FA1: Factor analytic structures of order 1 for both the between environment variance matrices of the GE additive genetic effects and non-additive genetic effects.

Model	**Variance model for GE effects**	**Parameters**		**LogLik**	**AIC**
		Total	**Genetic**		
$\Delta^{13}C$					
Baseline LMM	DIAG	24	3	-28.13016	104.2603
Baseline FA-LMM	FA1	30	6	92.26117	-124.5223
Baseline FA-LMM with markers	FA1:FA1	39	12	112.21808	-146.4362
Days to flower					
Baseline LMM	DIAG	17	3	-3364.616	6763.232
Baseline FA-LMM	FA1	23	6	-2985.602	6017.204
Baseline FA-LMM with markers	FA1:FA1	32	12	-2888.462	5840.924
Plant height					
Baseline LMM	DIAG	18	3	-23051.01	46138.02
Baseline FA-LMM	FA1	24	6	-22853.79	45755.58
Baseline FA-LMM with markers	FA1:FA1	33	12	-22837.59	45741.19
Seed yield					
Baseline LMM	DIAG	13	2	1110.801	-2195.601
Baseline FA-LMM	FA1	16	3	1244.776	-2457.552
Baseline FA-LMM with markers	FA1:FA1	21	6	1265.601	-2489.202
NDVI					
Baseline LMM	-	7	1	938.9046	-1863.809
Baseline LMM with markers	-	8	2	943.0174	-1870.035

Working model for QTL analysis

The approach used herein for the determination of genomic regions that influence the expression of the traits associated with water use efficiency in multiple environments is an extension of the whole genome, single-step, multi-environment QTL analysis approach developed by Verbyla et al. (2012).

Our approach uses an alternative working model where we partition M as $M = [M_1\ M_2 \dots\ M_c]$ and let $K = \sum_{i=1}^{c} M_i M_i^T$. Further, let $M_i = \left(m_{i;jk}\right)$ be the marker matrix for the chromosome i of size $n_g \times n_{m_i}$ where $n_g$is the number of genotypes with marker data, $n_{m_i}$is the number of markers in chromosome i and j and k are the subscripts for markers within chromosomes and genotypes respectively. If $n_m = \sum_{i=1}^{c} n_{m_i}$ then M is $n_g \times n_m$ and K is $n_g \times n_g$.

Let baseline FA-LMM with markers denoted by FA_B. For the QTL analysis we fit the same FA_B model where we replace the GRM, K, by the formulation presented above. Further, let $\hat{\kappa}^T = (\hat{\kappa}^T{}_{vm}, \hat{\kappa}^T{}_{ide}, \hat{\kappa}^T{}_{o})$ be the vector of variance parameters partitioned conformably with the associated model terms and evaluated at the residual maximum likelihood (REML) estimates of $\hat{\kappa}$ from FA_B.

The QTL analysis approach that we use has multiple steps which are described in detail in the following.

Step 1: Genome scan

To avoid proximal contamination (Listgarten et al. 2012) and to reduce computation time (Lippert et al. 2011) we scan the genome by fitting each marker $m_{i;j}$, $i = 1, \dots, c$ and $j = 1, \dots, n_{m_i}$ as a covariate by adding the main effect and the interaction of the marker within the environment. In this fit we reduce computation by holding κ_{vm} at the REML estimates from FA_B and only re-estimate κ_{ide} and κ_o. Furthermore, we replace the GRM, K, by K_{-i} where K_{-i} is given by

$$K_{-i} = \sum_{\substack{l=1, \\ l \neq i}}^{c} M_l M_l^T$$

This avoids the possibility of proximal contamination and is a conservative approach.

For each fit we compute the preceding terms' Wald statistics from the fit of

`fixed ~ Environment + Environment:Gdrop +` $m_{i;j}$ `+` $m_{i;j}$`:Environment` where $m_{i;j}$ represents the covariate $m_{i;j}$ expanded by $m_{i;j} = Z_g m_{i;j}$ where Z_g is the design matrix for `Gkeep`. Two probabilities are obtained for each fit and we denote these by $p_{m;ij}$ and $p_{m;ij.e}$

Step 2: LD block thinning for Marker by Environment ($M \times E_{Q_{ME}}$) interactions

In this step, we determined LD blocks by calculating the squared LD (r^2) for each pair of markers and any markers with an estimated LD of greater than 0.7 were deemed to be within the same LD block. We thin $\{p_{m;ij.e}\}$ by tagging each $p_{m;ij.e}$ by its LD block and choosing the representative marker to be the marker with the lowest *p*-value from the set of markers in each LD block. If we denote the set of thinned *p*-values by ${p_{me}}^T$, we then apply a false discovery rate (FDR) approach (Benjamini and Hochberg, 1995) to this vector to determine the working set of Marker by Environment ($M \times E_{Q_{ME}}$) QTLs.

Step 3: LD block thinning for Marker (M_{Q_M}) main effects

In this step, we apply the same approach from Step 2 to thin $\{p_{m;ij}\}$, except before thinning we discard any marker $m_{i;j}$ that is present in the thinned set from Step 2, i.e. in the working set of $M \times E_{Q_{ME}}$ QTLs. We also exclude all markers from the set of LD blocks associated with the working set of $M \times E_{Q_{ME}}$ QTLs. This provides the working set denoted by M_{Q_M}.

Step 4: Final multi QTL model

We include all markers from Step 2 $M \times E_{Q_{ME}}$ QTL set as "$m_{i;j}$ `:Environment`" and include all markers from Step 3 M_{Q_M} QTL set as "$m_{i;j}$" as fixed terms in the model. In this model we replace the GRM, K, by K_e, where K_e is the GRM using markers from LD blocks which do not contain markers from working QTL sets M_{Q_M} and $M \times E_{Q_{ME}}$. We then use standard backward elimination procedure to determine the final multi-QTL model. During this procedure we remove the $M \times E_{Q_{ME}}$ interactions first and foremost and stop when there are no more non-significant interactions. We then focus on the M_{Q_M} main effects but leaving the main effects for the significant interactions in the model, i.e. only exclude main effects for those non-significant interactions and then exclude main effects only if there is not a partner interaction. Each time when a marker is dropped during the backward

elimination, that marker and its LD block is added back into the GRM. All remaining M_{Q_M} and $M \times E_{Q_{ME}}$ markers in the fixed component of the final multi-QTL model are reported as putative QTLs together with their positions, LOD scores, i.e., $-\log_{10}(p\text{-value})$ and the percentage of genetic variance accounted (R^2).

(ii) LMM analysis for the Experiment 4

The models for each trait for the LMM analysis of Experiment 4 are presented here using the symbolic model notation of Wilkinson and Rogers (1973).

<u>$\Delta^{13}C$</u>

```
fixed ~ 1
random ~ Genotype + Carousel + Rep + Plot
```

<u>Days to flower</u>

```
fixed ~ 1
random ~ Genotype + Block:Rep + Block:Range + Block:Row
```

<u>Plant height</u>

```
fixed ~ 1 + Irrig[Block]
random ~ Genotype + Genotype:Irrig[Block] + Block[Irrig]:Rep  +
Plot + Block[Irrig]:Range + Block[Irrig]:Row
```

<u>Seed yield</u>

```
fixed ~ 1 + Irrig[Block] + nplants
random ~ Genotype + Genotype:Irrig[Block] + Block[Irrig]:Rep  +
Block[Irrig]:Range + Block[Irrig]:Row
```

<u>Gas exchange and physiological traits</u>

```
fixed ~ 1
random ~ Genotype + Rep + Range + Row
```

There are many characteristics of the above symbolic model notation that require explanation and defining. Firstly, the construction of the appropriate genetic component of the model for each trait represented the treatment structure of the trial. As there was a factorial treatment structure (factorial combination of genotypes and irrigation regimes), it is necessary to incorporate this structure in the analysis. Based on the randomisation procedure of the factorial treatment structure described in Method S1, there is aliasing of irrigation regime (`Irrig`) and block (`Block`) effects and hereafter denoted as `Irrig[Block]` or `Block[Irrig]`. Hence, there was no valid inferential framework to test the main effect of irrigation regimes (`Irrig`). This is explained as “pseudo” or “false" replication in Bailey (2008). There was a valid inferential framework to test the genotypes by irrigation regimes interaction (`Genotype:Irrig[Block]`), but caution should be exercised when interpreting the results as these effects may be due to the different blocks (`Block[Irrig]`) (see Gururaj et al., (2020) for details). Thus, in terms of the treatment effects, random genotype main effects (`Genotype`) and random genotype by irrigation block interactions were fitted (`Genotype:Irrig[Block]`) and the `Irrig[Block]` effects were fitted as fixed. However, this treatment structure is only appropriate for the traits plant height and seed yield as days to flower is measured before the irrigation regime treatment being imposed, whereas $\Delta^{13}C$, gas exchange and physiological measurements were taken only for the wet block. Hence, for the remaining traits, only the genotype (`Genotype`) main effects were fitted as random. The seed yield is measured for each plot and the number of plants (`nplants`) available in each plot were measured. Number of plants is fitted as a covariate accounting for the variation in seed yield (p-value = 0.000057507 from Wald test). All models also included a term representing the overall mean (`1`) as a fixed effect.

The non-genetic component of the model for each trait required the inclusion of random terms representing the plot structure of the experiment (`Block` and `Block:Rep`). For plant height and seed yield random block (`Block`) effects were not fitted as the aliasing irrigation regime effects (`Irrig[Block]`) were fitted as fixed. As $\Delta^{13}C$, gas exchange and physiological measurements were taken only for the wet block these traits required inclusion of only the random `Rep` effects. Random row (`Block:Row or Row`) and range (`Block:Range or Range`) effects within blocks were included as necessary. For plant height and $\Delta^{13}C$, the model also included random `Plot` effects which partition residual

variance as the experimental unit of these traits is not the observational unit. For $\Delta^{13}C$, random `Carousel` effects were fitted as it is a source of non-genetic variation which occurred in the lab phase. Variations related to the runs were not necessary as all samples were evaluated in a single run (see **Error! Reference source not found.**). The variance models for the residuals were independent and identically distributed (iid) effects as the experiment was conducted in single row plots.

Light microscopy

A leaf disc (9.08 cm^2 size) was taken from each of two replicate canola lines from Experiment 4 (wet treatment) for anatomical analysis. Discs were collected from leaf blades, at a position between the 2nd and 3rd veins, from plants at the flowering stage and fixed in 70% ethanol. After clearing in 5% household bleach (~0.2% sodium hypochlorite) for five days, they were stored in 70% ethanol before sectioning and staining using a method modified from Rae *et al.* (2020). For imaging, hand sections were made from leaf discs, oxidised in 1% (w/v) periodic acid for ~15 sec, and then rinsed in water ~15 sec. Sections were incubated in 2 µg/ml Rhodamine-123 in water for 30 min and rinsed for 5-10 min in water before mounting in 50% glycerol. Samples were imaged using 488 nm excitation and 500-560 nm emission on a Leica SP8 confocal microscope.

RNA sequencing and differential gene expression analysis

Parental lines, BC1329 and BC9102 of DH population were grown in three replicates under both wet (100% field capacity) and dry (50% field capacity) treatments in a glasshouse, located at Wagga Wagga Agricultural Institute . The clean sequence reads (100 bp single-end reads) for 12 samples that had per base sequence quality with >96% bases above Q30 were aligned against the *B. napus* reference Darmor-*bzh* (Version 4.1), using STAR aligner (v2.5.3a) (https://github.com/alexdobin/STAR/blob/master/doc/STARmanual.pdf). The raw counts of reads mapping to each known gene was used to perform differential expression analysis using edgeR (version 3.30.3) (https://bioconductor.org/packages/release/bioc/html/edgeR.html) using R version 4.0.3. Counts are summarised at gene level using the featureCounts v1.5.3 utility of the subread package (http://subread.sourceforge.net/). The default TMM normalisation method of edeR was used to normalise the counts between samples. A generalised linear model approach was then used to quantify the differential expression

between the groups. The differentially expressed genes (DEGs) were obtained using a false discovery rate (FDR < 0.05). Heatmaps showing the expression pattern of genes in A09 and C09 QTL regions were produced using the ComplexHeatmap R package (Gu et al., 2016).

References for Statistical analysis

Bailey, R. (2008). *Design of Comparative Experiments (Cambridge Series in Statistical and Probabilistic Mathematics).* Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511611483

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J. and Thompson, R. (2018). ASReml-R Reference Manual Version 4. Technical report, VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Cullis, B. R. and Gleeson, A. (1991). Spatial Analysis of Field Experiments-An Extension to Two Dimensions. *Biometrics*, 47(4), 1449-1460. doi:10.2307/2532398

Cullis, B. R. and Smith, A. B. (2016). The analysis of QTL and QTL by treatment experiments using spatial models for marker effects. Technical report.

Gilmour, A. R., Cullis, B. R. and Verbyla, A. P. (1997). Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agricultural, Biological, and Environmental Statistics* 2, 269–293. doi: 10.2307/1400446

Kadkol, G., Smith, A., Cullis, B. and Chenu, K. (2020). Variation in Australian durum wheat germplasm for productivity traits under irrigated and rainfed conditions: Genotype performance for agronomic traits and benchmarking. *The Journal of Agricultural Science* 158(6), 479-495. doi:10.1017/S0021859620000817

Lippert, C., Listgarten, J., Liu, Y., Kadie, C. M., Davidson, R. I. and Heckerman D (2011). FaST linear mixed models for genome-wide association Studies. *Nature methods*, 8, 833–835. https://doi.org/10.1038/nmeth.1681

Listgarten, J., Lippert, C., Kadie, C. M., Davidson, R. I., Eskin, E. and Heckerman, D. (2012). Improved linear mixed models for genome-wide association studies. *Nature methods*, 9(6), 525–526.

Oakey, H., Verbyla, A., Cullis, B. R., Wei, X. and Pitchford, W. (2007). Joint modelling of additive and non-additive (genetic line) efects in multi-environment trials. *Theoretical and Applied Genetics* 114, 1319–1332.

Raman, H., McVittie, B., Pirathiban, R., Raman, R., Zhang, Y., Barbulescu, DM., Qiu, Y., Liu, S. and Cullis B. (2020). Genome-Wide Association Mapping Identifies Novel Loci for Quantitative Resistance to Blackleg Disease in Canola. *Frontiers in Plant Science* 2020; 11:1184. pmid:32849733

Smith, A., Lim, P. and Cullis, B. (2006). The design and analysis of multi-phase plant breeding experiments. *The Journal of Agricultural Science* 144 5, 393–409. doi: 10.1017/ S0021859606006319

Smith, A., Ganesalingam, A., Kuchel, H. and Cullis, B. (2015). Factor analytic mixed models for the provision of grower information from national crop variety testing programmes. *Theoretical and Applied Genetics* 128:55–72

Taylor, J. D. and Butler, D. (2014). ASMap: An (A)ccurate and (S)peedy linkage map construction package for inbred populations that uses the extremely efficient MSTmap algorithm.

Thompson, R., Cullis, B., Smith, A. and Gilmour, A. (2003). A Sparse Implementation of the Average Information Algorithm for Factor Analytic and Reduced Rank Variance Models. *Australian and New Zealand Journal of Statistics* 45, 445-459.

Tolhurst, D. J., Mathews, K. L., Smith, A. B. and Cullis, B. R. (2019). Genomic selection in multi-environment plant breeding trials using a factor analytic linear mixed model. *Journal of Animal Breeding and Genetics* 136, 279–300. doi: 10.1111/ jbg.12404

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., et al. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics* 17, 520–525. doi: 10.1093/bioinformatics/17.6.520

VanRaden, P. M. (2008). Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science,* 91, 4414–4423. https ://doi.org/10.3168/jds.2007-0980

Verbyla, A.P. and Cullis, B.R. Multivariate whole genome average interval mapping: QTL analysis for multiple traits and/or environments. *Theoretical and Applied Genetics* 125, 933–953 (2012). https://doi.org/10.1007/s00122-012-1884-9

Wilkinson, G. N. and Rogers, C. E. (1973). Symbolic description of factorial models for analysis of variance. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 22, 392–399.

Wu, Y., Bhat, P. R., Close, T. J. and Lonardi, S. (2008). Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. *PLoS Genetics* 4.

Table S5: Summary of the partitioning of genetic variance into Additive and Non-additive and REML estimate of the Total (additive plus non-additive) genetic variance before (Baseline FA-LMM with markers: M1) and after identifying putative QTL (Final multi QTL model: M2) for each of the trait and trial.

Trait	Experiment	Phenotypic environment	Additive (M1, %)	Additive (M2, %)	Non-additive (M1, %)	Non-additive (M1, %)	Total (M1)	Total (M2)	VAF_m (%)
$\Delta^{13}C$	1	Field (2017)	47.99	23.46	52.01	78.36	0.49	0.32	-
	2	Field (2018)	50.11	23.98	49.89	77.88	0.23	0.15	-
	3	Pot (2017)	55.65	19.96	44.35	81.59	0.23	0.12	-
		Mean	51.25	22.47	48.75	79.28	0.32	0.20	37.91
DTF	1	Field (2017)	79.41	31.77	20.59	72.66	45.14	17.16	-
	2	Field (2018)	82.44	34.07	17.56	70.68	55.66	19.07	-
	3	Pot (2017)	75.55	13.19	24.45	88.64	46.02	17.08	-
		Mean	79.13	26.35	20.87	77.33	48.94	17.77	63.68
NDVI	2	Field (2018)	21.53	13.92	78.47	86.08	0.01	0.01	11.76
PH	1	Field (2017)	19.04	1.92	80.96	98.22	135.09	108.43	-
	2	Field (2018)	34.42	11.00	65.58	89.79	131.97	88.13	-
	3	Pot (2017)	45.87	22.57	54.13	79.04	69.30	43.63	-
		Mean	33.11	11.83	66.89	89.02	112.12	80.06	28.59
SY	1	Field (2017)	40.96	2.91	59.04	98.69	0.14	0.08	-
	2	Field (2018)	60.18	17.99	39.82	91.92	0.08	0.04	-
		Mean	50.57	10.45	49.43	95.31	0.11	0.06	45.42

VAF_m shows the percentage of genetic variance accounted by the identified putative QTLs. Trait variations were assessed using MET analyses across three environments in 2017 and 2018 under field and pot experiments (Experiments 1-3) for $\Delta^{13}C$, DTF and PH and two environments (Experiments 1-2) for SY; whereas it was based on a single site analysis from only one environment for NDVI (Experiment 2). $\Delta^{13}C$: Carbon isotope discrimination; DTF: Days to flower; NDVI: Normalised difference vegetative difference; PH: Plant height; SY: Seed yield.

Table S1: Summary of heritability, mean, minimum and maximum values for each trait and trial in the doubled haploid population from BC1329/BC9102.

Trait	Experiment	Phenotypic environment	Mean	Minimum	Maximum	Heritability (h^2)
$\Delta^{13}C$	1	Field (2017)	20.47	18.66	21.75	0.86
	2	Field (2018)	18.87	17.59	19.86	0.59
	3	Pot (2017)	21.21	20.10	22.08	0.59
		OP	20.16	18.73	21.25	-
DTF	1	Field (2017)	116.1	101.3	131.5	0.98
	2	Field (2018)	115.00	98.19	132.42	0.98
	3	Pot (2017)	108.07	92.64	124.05	0.91
		OP	113.07	97.37	129.33	-
NDVI	2	Field (2018)	0.4604	0.1734	0.6230	0.92
PH	1	Field (2017)	80.34	46.18	99.61	0.93
	2	Field (2018)	89.26	53.34	109.52	0.91
	3	Pot (2017)	77.84	53.73	91.45	0.56
		OP	82.48	51.08	100.20	-
SY	1	Field (2017)	0.53	0.02	1.25	0.98
	2	Field (2018)	0.41	0.001	0.99	0.98
		OP	0.47	0.01	1.12	-

Trait variations were assessed using MET analyses across three environments in 2017 and 2018 under field and pot experiments (Experiments 1-3) for $\Delta^{13}C$, DTF and PH and two environments (Experiments 1-2) for SY; whereas it was based on a single site analysis from only one environment for NDVI (Experiment 2). CGE-EBLUPs and Overall performance (OP) estimates are summarised for $\Delta^{13}C$, DTF, PH and SY; whereas genotype EBLUPs are summarised for NDVI. $\Delta^{13}C$: Carbon isotope discrimination; DTF Days to flower; NDVI: Normalised difference vegetative difference; PH: Plant height; SY: Seed yield.

Table S7: REML estimate of the Between environment genetic correlation matrices for Additive and Total (additive plus non-additive) effects of each trait from the baseline FA-LMM with markers.

Trait	Variance	Phenotypic environment	Correlation (ρ)		
			Field (2017)	Field (2018)	Pot (2017)
$\Delta^{13}C$	Additive	Field (2017)	1.00	0.96	0.99
		Field (2018)	0.96	1.00	0.98
		Pot (2017)	0.99	0.98	1.00
	Total	Field (2017)	1.00	0.93	0.83
		Field (2018)	0.93	1.00	0.87
		Pot (2017)	0.83	0.87	1.00
DTF	Additive	Field (2017)	1.00	0.98	0.92
		Field (2018)	0.98	1.00	0.94
		Pot (2017)	0.92	0.94	1.00
	Total	Field (2017)	1.00	0.94	0.91
		Field (2018)	0.94	1.00	0.90
		Pot (2017)	0.91	0.90	1.00
PH	Additive	Field (2017)	1.00	0.95	0.94
		Field (2018)	0.95	1.00	0.89
		Pot (2017)	0.94	0.89	1.00
	Total	Field (2017)	1.00	0.90	0.87
		Field (2018)	0.90	1.00	0.94
		Pot (2017)	0.87	0.94	1.00
SY	Additive	Field (2017)	1.00	0.91	-
		Field (2018)	0.91	1.00	-
	Total	Field (2017)	1.00	0.87	-
		Field (2018)	0.87	1.00	-

Trait variations were assessed using MET analyses across three environments in 2017 and 2018 under field and pot experiments (Experiments 1-3) for $\Delta^{13}C$, DTF and PH and two environments (Experiments 1-2) for SY. $\Delta^{13}C$: Carbon isotope discrimination; DTF: Days to flower; PH: Plant height; SY: Seed yield; -: Not applicable.

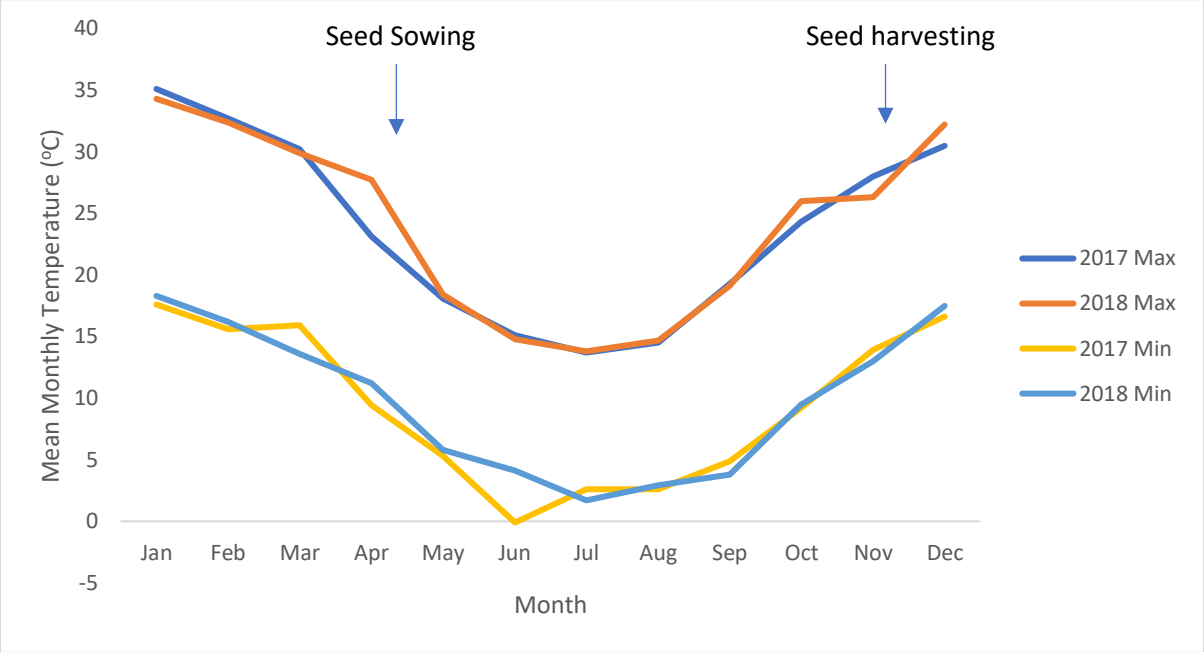
Table S9: Summary of heritability, mean, minimum and maximum values of the genotype EBLUPs for each trait evaluated in the rain-out shelter (Experiment 4) with wet and dry conditions for a doubled haploid population from BC1329/BC9102.

***Trait**	**Units**	**Block**	**Heritability (H^2)**	**Mean**	**Minimum**	**Maximum**
$\Delta^{13}C$	‰	Wet	0.39	22.91	22.17	23.49
A	µmol m^{-2} sec^{-1}	Wet	0.77	10.34	4.97	17.15
g_{sw}	mol H_2O m^{-2} sec^{-1}	Wet	0.58	0.23	0.11	0.38
*i*WUE	µmol/ mol H_2O	Wet	0.31	45.68	30.89	60.41
LWC	-	Wet	0.42	10.14	8.36	12.52
DTF	days	Wet	0.97	107.96	84.20	132.41
PH	cm	Dry	0.93	157.2	122.6	195.6
		Wet		157.3	120.0	193.0
SY	g	Dry	0.89	60.45	10.81	108.63
		Wet		110.47	62.76	163.50

*****$\Delta^{13}C$: Carbon isotope discrimination; *A*: Photosynthesis rate; g_{sw}: stomatal conductance; SLW: *i*WUE: intrinsic water use efficiency; LWC: Leaf water content; DTF: Days to flower; PH: Plant height; SY: Seed yield. Seed yield is predicted at the average value of 9.2 plants per plot.

Fig. S1: Weather conditions prevailed under field conditions (Wagga Wagga, NSW 2650, Australia), where experiments were conducted during canola growing seasons (May to November) in 2017, and 2018. a: Mean monthly maximum and minimum temperatures, b: Monthly cumulative total rainfall.

a



b

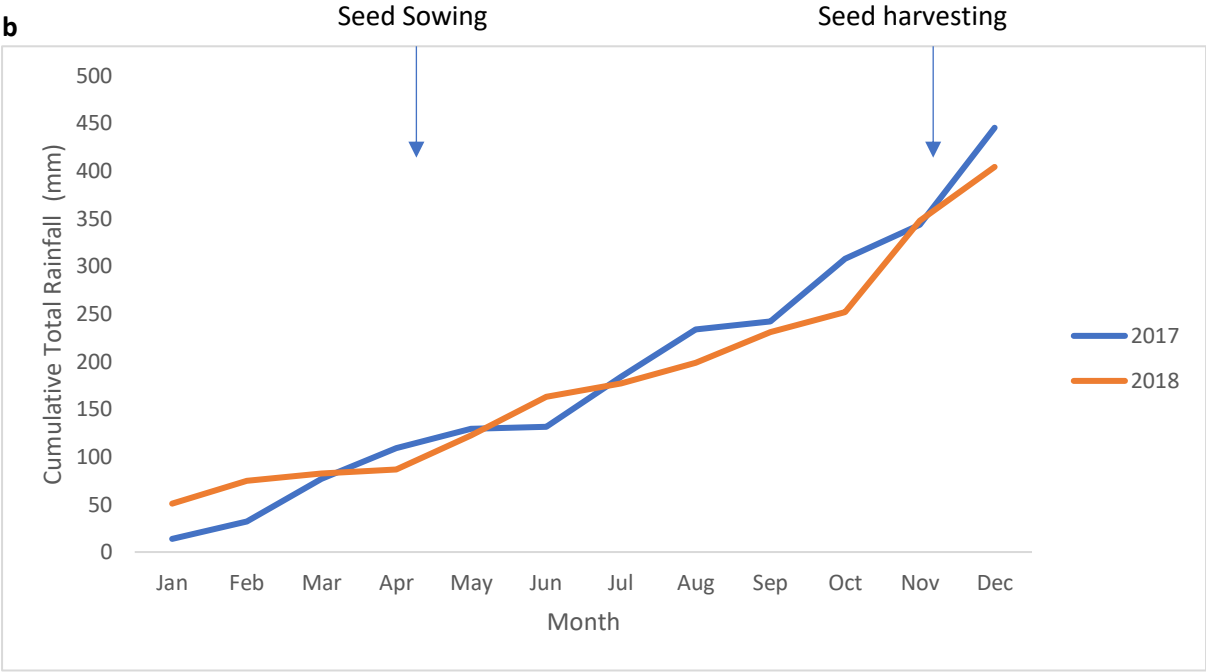
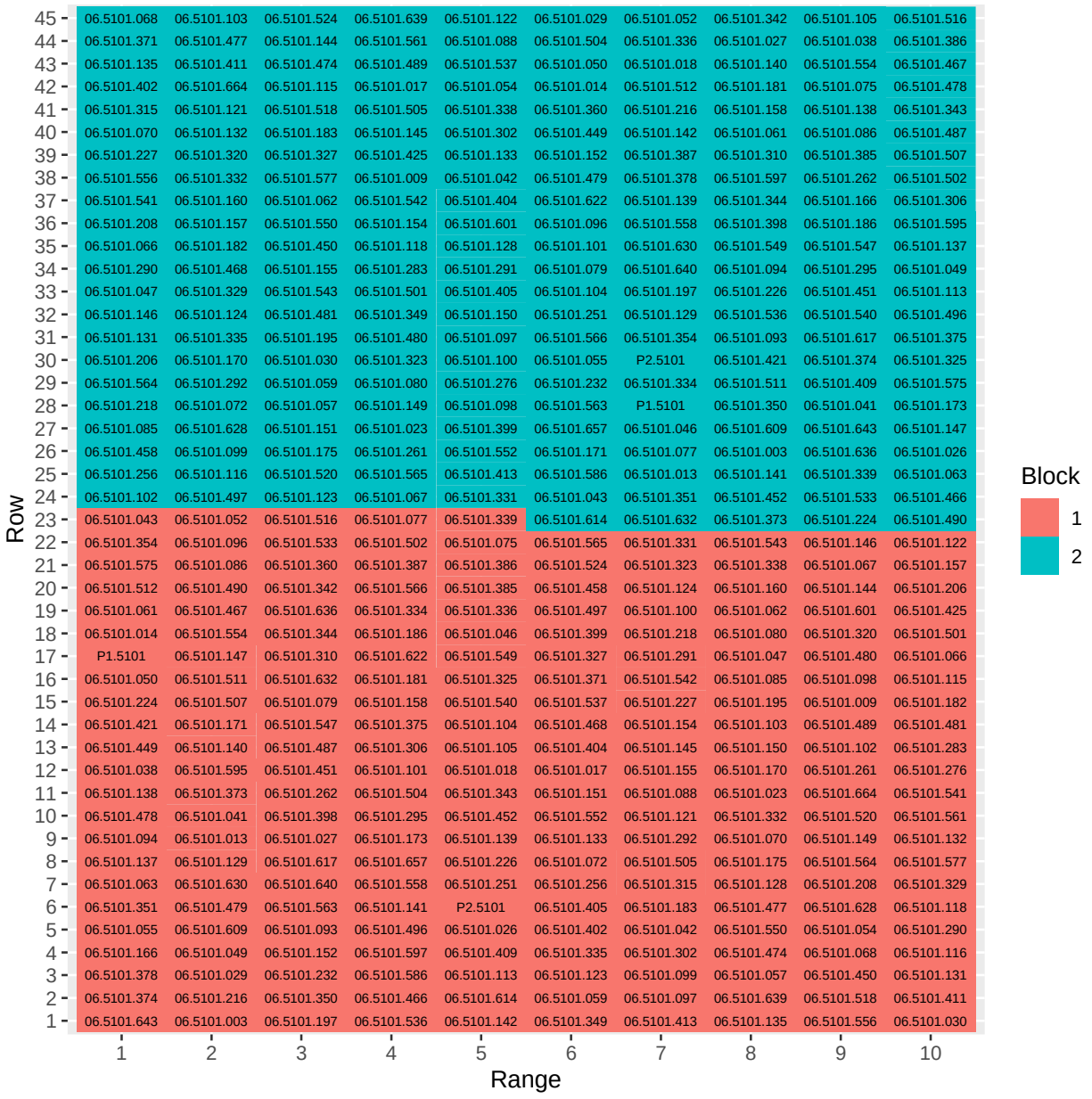
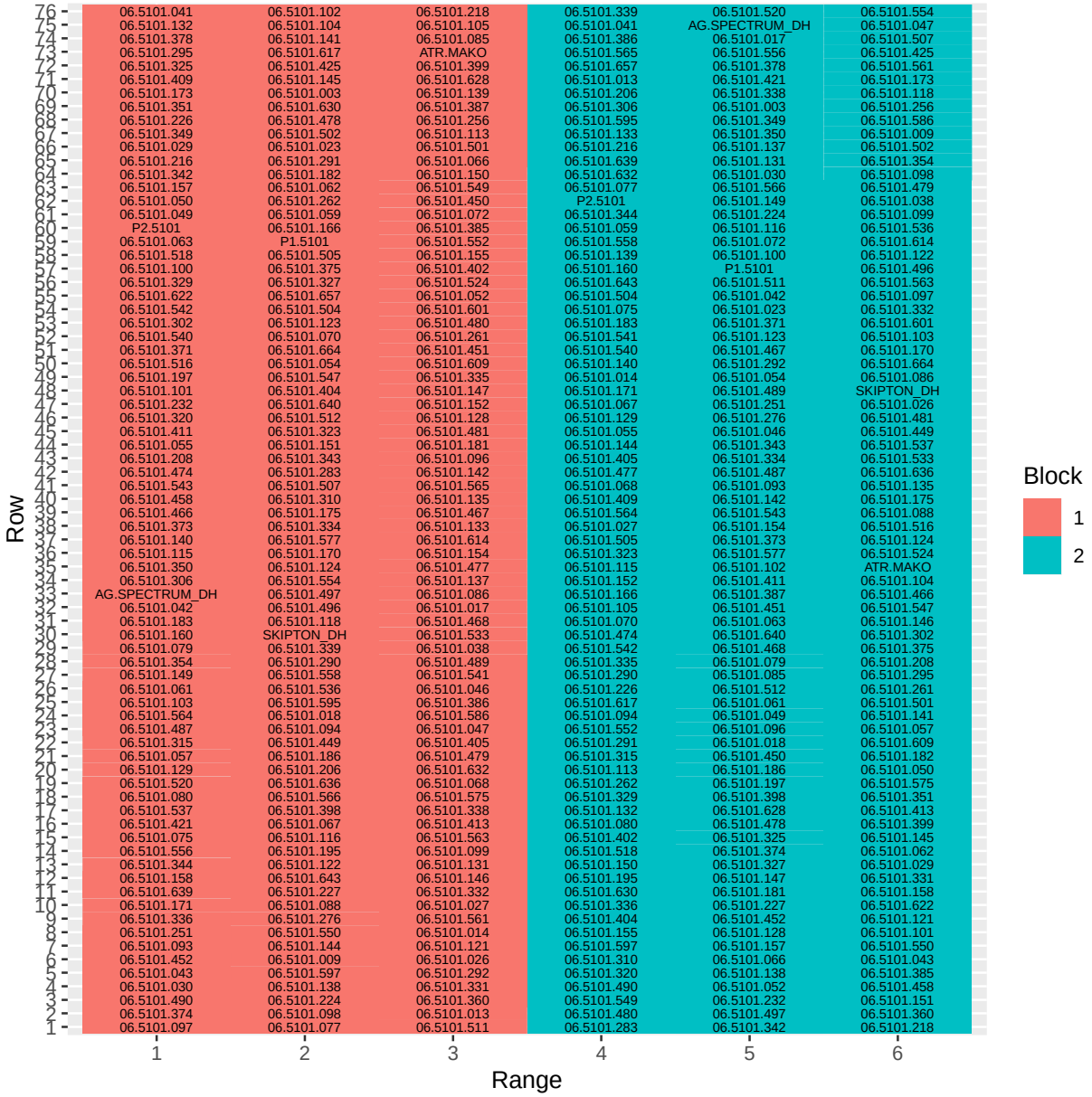


Fig. S2: Layout of experiments conducted under field in 2017 (a, Experiment 1), 2018 (b, Experiment 2), and under rain-out shelter conditions in 2017 (c, Experiment 3) and 2019 (d, Experiment 4).

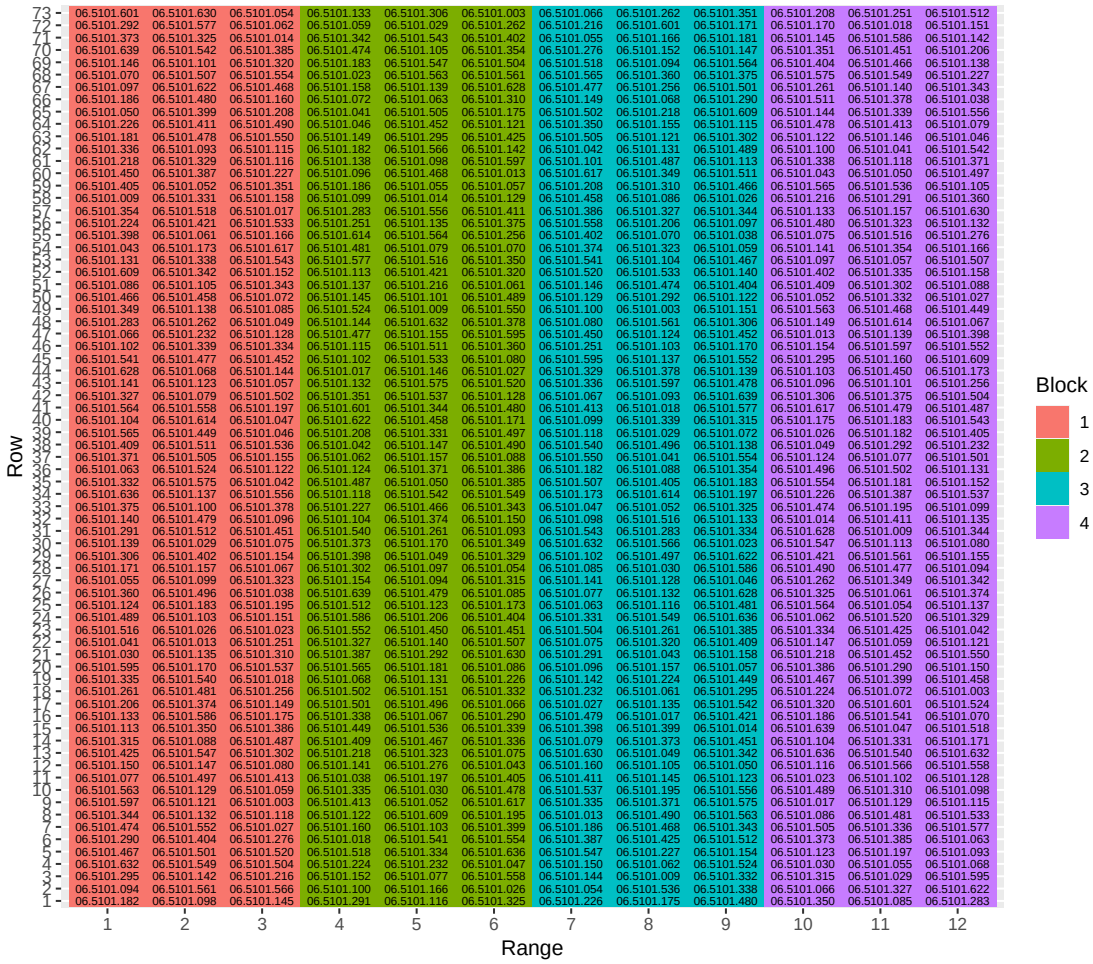
a



b


Row
Range
Block
1
2
1 2 3 4 5 6
76 06.5101.041 06.5101.102 06.5101.218 06.5101.339 06.5101.520 06.5101.554
75 06.5101.132 06.5101.104 06.5101.105 06.5101.041 AG.SPECTRUM_DH 06.5101.047
74 06.5101.378 06.5101.141 06.5101.085 06.5101.386 06.5101.017 06.5101.507
73 06.5101.295 06.5101.617 ATR.MAKO 06.5101.565 06.5101.556 06.5101.425
72 06.5101.325 06.5101.425 06.5101.399 06.5101.657 06.5101.378 06.5101.561
71 06.5101.409 06.5101.145 06.5101.628 06.5101.013 06.5101.421 06.5101.173
70 06.5101.173 06.5101.003 06.5101.139 06.5101.206 06.5101.338 06.5101.118
69 06.5101.351 06.5101.630 06.5101.387 06.5101.306 06.5101.003 06.5101.256
68 06.5101.226 06.5101.478 06.5101.256 06.5101.595 06.5101.349 06.5101.586
67 06.5101.349 06.5101.502 06.5101.113 06.5101.133 06.5101.350 06.5101.009
66 06.5101.029 06.5101.023 06.5101.501 06.5101.216 06.5101.137 06.5101.502
65 06.5101.216 06.5101.291 06.5101.066 06.5101.639 06.5101.131 06.5101.354
64 06.5101.342 06.5101.182 06.5101.150 06.5101.632 06.5101.030 06.5101.098
63 06.5101.157 06.5101.062 06.5101.549 06.5101.077 06.5101.566 06.5101.479
62 06.5101.050 06.5101.262 06.5101.450 P2.5101 06.5101.149 06.5101.038
61 06.5101.049 06.5101.059 06.5101.072 06.5101.344 06.5101.224 06.5101.099
60 P2.5101 06.5101.166 06.5101.385 06.5101.059 06.5101.116 06.5101.536
59 06.5101.063 P1.5101 06.5101.552 06.5101.558 06.5101.072 06.5101.614
58 06.5101.518 06.5101.505 06.5101.155 06.5101.139 06.5101.100 06.5101.122
57 06.5101.100 06.5101.375 06.5101.402 06.5101.160 P1.5101 06.5101.496
56 06.5101.329 06.5101.327 06.5101.524 06.5101.643 06.5101.511 06.5101.563
55 06.5101.622 06.5101.657 06.5101.052 06.5101.504 06.5101.042 06.5101.097
54 06.5101.542 06.5101.504 06.5101.601 06.5101.075 06.5101.023 06.5101.332
53 06.5101.302 06.5101.123 06.5101.480 06.5101.183 06.5101.371 06.5101.601
52 06.5101.540 06.5101.070 06.5101.261 06.5101.541 06.5101.123 06.5101.103
51 06.5101.371 06.5101.664 06.5101.451 06.5101.540 06.5101.467 06.5101.170
50 06.5101.516 06.5101.054 06.5101.609 06.5101.140 06.5101.292 06.5101.664
49 06.5101.197 06.5101.547 06.5101.335 06.5101.014 06.5101.054 06.5101.086
48 06.5101.101 06.5101.404 06.5101.147 06.5101.171 06.5101.489 SKIPTON_DH
47 06.5101.232 06.5101.640 06.5101.152 06.5101.067 06.5101.251 06.5101.026
46 06.5101.320 06.5101.512 06.5101.128 06.5101.129 06.5101.276 06.5101.481
45 06.5101.411 06.5101.323 06.5101.481 06.5101.055 06.5101.046 06.5101.449
44 06.5101.055 06.5101.151 06.5101.181 06.5101.144 06.5101.343 06.5101.537
43 06.5101.208 06.5101.343 06.5101.096 06.5101.405 06.5101.334 06.5101.533
42 06.5101.474 06.5101.283 06.5101.142 06.5101.477 06.5101.487 06.5101.636
41 06.5101.543 06.5101.507 06.5101.565 06.5101.068 06.5101.093 06.5101.135
40 06.5101.458 06.5101.310 06.5101.135 06.5101.409 06.5101.142 06.5101.175
39 06.5101.466 06.5101.175 06.5101.467 06.5101.564 06.5101.543 06.5101.088
38 06.5101.373 06.5101.334 06.5101.133 06.5101.027 06.5101.154 06.5101.516
37 06.5101.140 06.5101.577 06.5101.614 06.5101.505 06.5101.373 06.5101.124
36 06.5101.115 06.5101.170 06.5101.154 06.5101.323 06.5101.577 06.5101.524
35 06.5101.350 06.5101.124 06.5101.477 06.5101.115 06.5101.102 ATR.MAKO
34 06.5101.306 06.5101.554 06.5101.137 06.5101.152 06.5101.411 06.5101.104
33 AG.SPECTRUM_DH 06.5101.497 06.5101.086 06.5101.166 06.5101.387 06.5101.466
32 06.5101.042 06.5101.496 06.5101.017 06.5101.105 06.5101.451 06.5101.547
31 06.5101.183 06.5101.118 06.5101.468 06.5101.070 06.5101.063 06.5101.146
30 06.5101.160 SKIPTON_DH 06.5101.533 06.5101.474 06.5101.640 06.5101.302
29 06.5101.079 06.5101.339 06.5101.038 06.5101.542 06.5101.468 06.5101.375
28 06.5101.354 06.5101.290 06.5101.489 06.5101.335 06.5101.079 06.5101.208
27 06.5101.149 06.5101.558 06.5101.541 06.5101.290 06.5101.085 06.5101.295
26 06.5101.061 06.5101.536 06.5101.046 06.5101.226 06.5101.512 06.5101.261
25 06.5101.103 06.5101.595 06.5101.386 06.5101.617 06.5101.061 06.5101.501
24 06.5101.564 06.5101.018 06.5101.586 06.5101.094 06.5101.049 06.5101.141
23 06.5101.487 06.5101.094 06.5101.047 06.5101.552 06.5101.096 06.5101.057
22 06.5101.315 06.5101.449 06.5101.405 06.5101.291 06.5101.018 06.5101.609
21 06.5101.057 06.5101.186 06.5101.479 06.5101.315 06.5101.450 06.5101.182
20 06.5101.129 06.5101.206 06.5101.632 06.5101.113 06.5101.186 06.5101.050
19 06.5101.520 06.5101.636 06.5101.068 06.5101.262 06.5101.197 06.5101.575
18 06.5101.080 06.5101.566 06.5101.575 06.5101.329 06.5101.398 06.5101.351
17 06.5101.537 06.5101.398 06.5101.338 06.5101.132 06.5101.628 06.5101.413
16 06.5101.421 06.5101.067 06.5101.413 06.5101.080 06.5101.478 06.5101.399
15 06.5101.075 06.5101.116 06.5101.563 06.5101.402 06.5101.325 06.5101.145
14 06.5101.556 06.5101.195 06.5101.099 06.5101.518 06.5101.374 06.5101.062
13 06.5101.344 06.5101.122 06.5101.131 06.5101.150 06.5101.327 06.5101.029
12 06.5101.158 06.5101.643 06.5101.146 06.5101.195 06.5101.147 06.5101.331
11 06.5101.639 06.5101.227 06.5101.332 06.5101.630 06.5101.181 06.5101.158
10 06.5101.171 06.5101.088 06.5101.027 06.5101.336 06.5101.227 06.5101.622
9 06.5101.336 06.5101.276 06.5101.561 06.5101.404 06.5101.452 06.5101.121
8 06.5101.251 06.5101.550 06.5101.014 06.5101.155 06.5101.128 06.5101.101
7 06.5101.093 06.5101.144 06.5101.121 06.5101.597 06.5101.157 06.5101.550
6 06.5101.452 06.5101.009 06.5101.026 06.5101.310 06.5101.066 06.5101.043
5 06.5101.043 06.5101.597 06.5101.292 06.5101.320 06.5101.138 06.5101.385
4 06.5101.030 06.5101.138 06.5101.331 06.5101.490 06.5101.052 06.5101.458
3 06.5101.490 06.5101.224 06.5101.360 06.5101.549 06.5101.232 06.5101.151
2 06.5101.374 06.5101.098 06.5101.013 06.5101.480 06.5101.497 06.5101.360
1 06.5101.097 06.5101.077 06.5101.511 06.5101.283 06.5101.342 06.5101.218

c



d

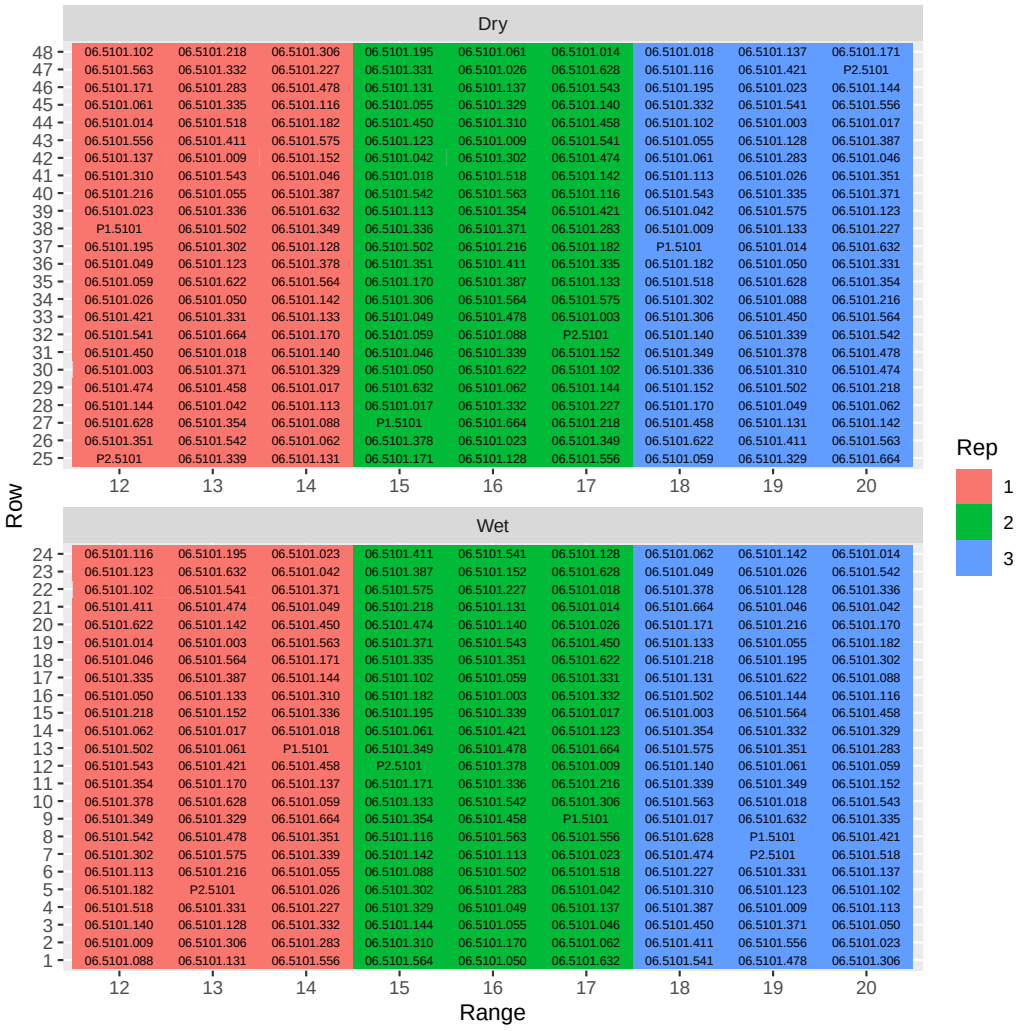
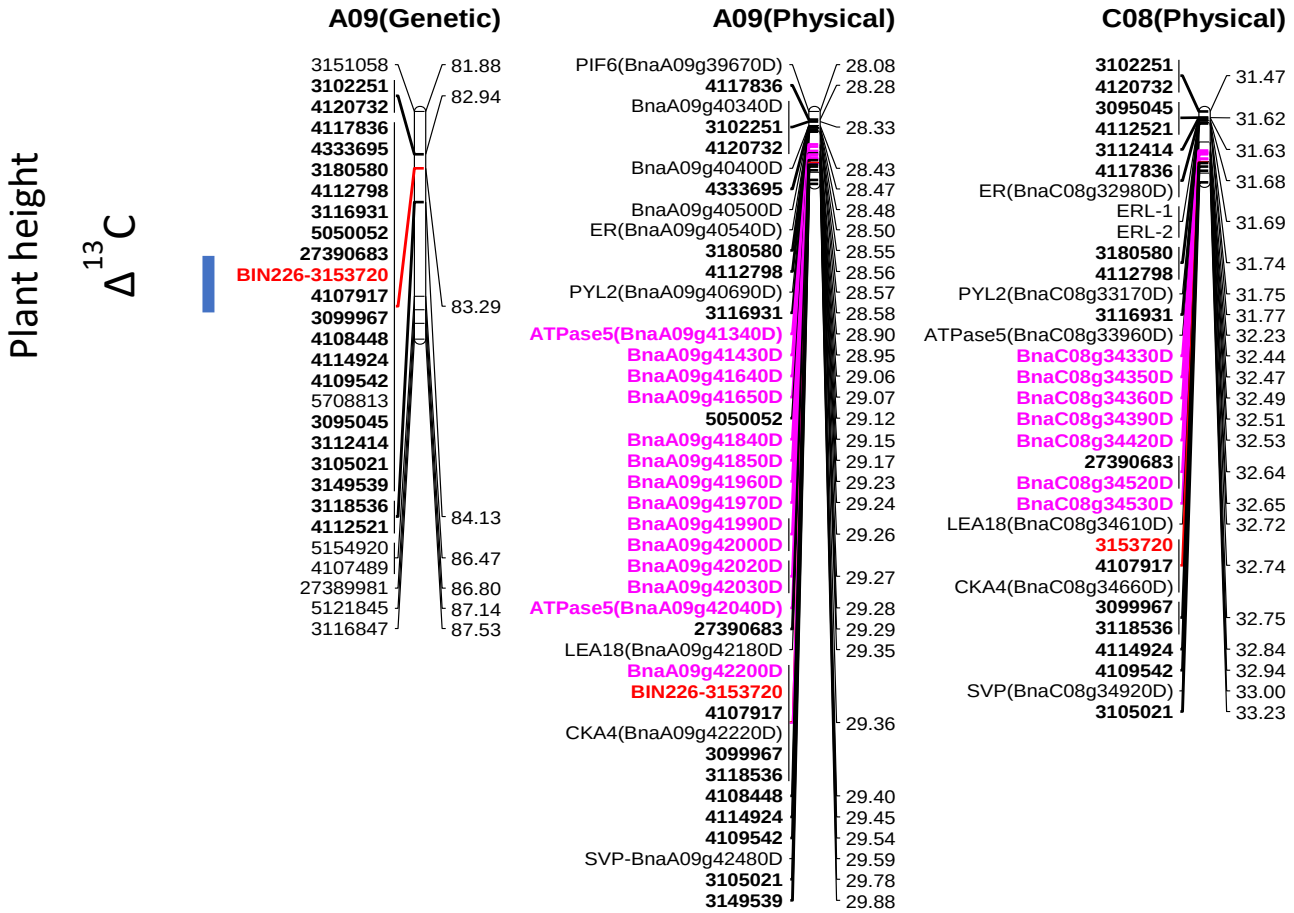
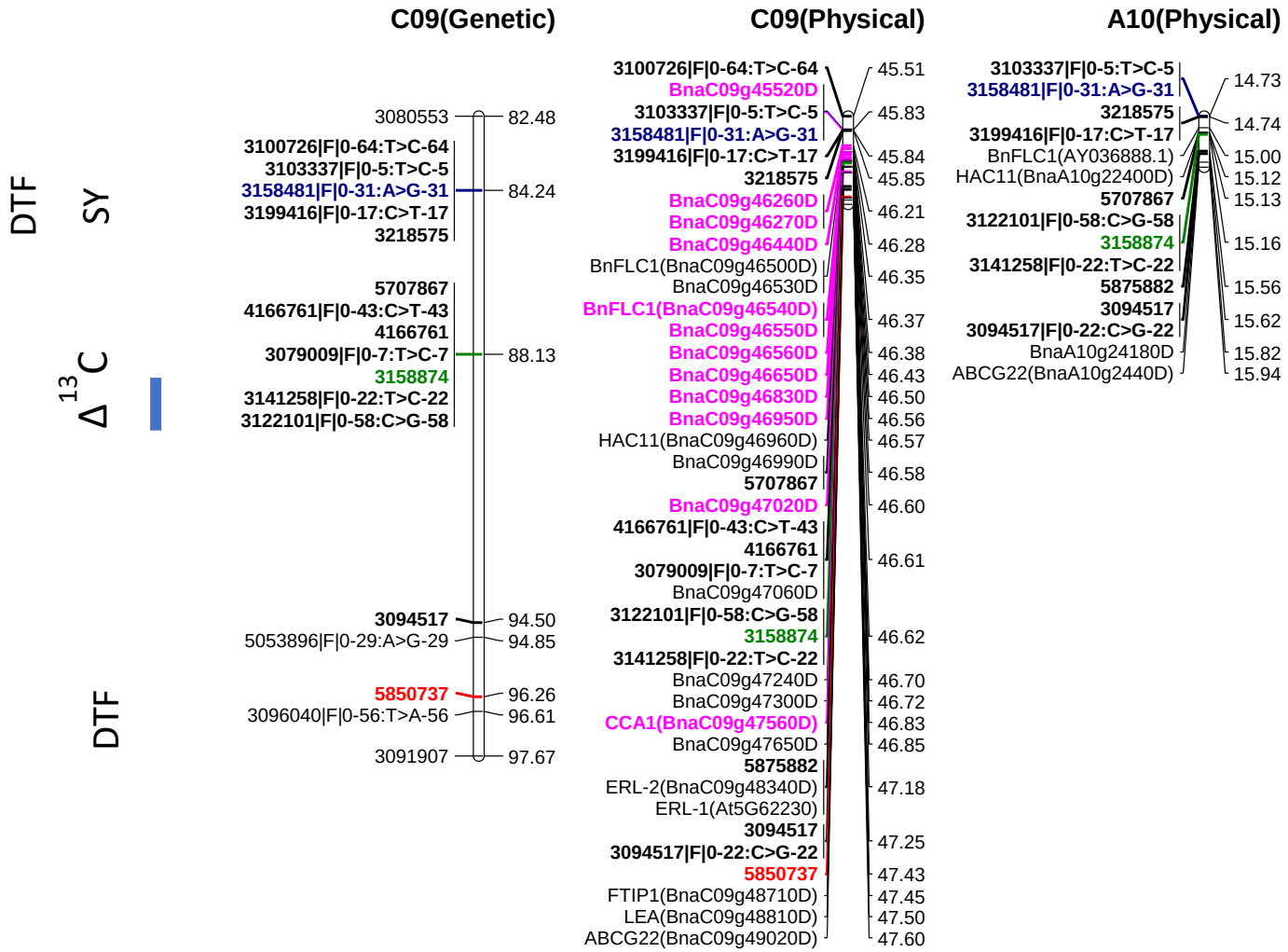


Fig. S3: Graphical representation showing localisation of multi-trait QTL associated with $\Delta^{13}C$ (‰), flowering time (days to flower, DTF); plant height (PHT) and seed yield (SY) on chromosomes A09 (a) and C09 (b) in a DH population from the BC1329/BC9102. DArTseq markers and their genetic map positions are shown on *right*- and *left*-hand side, respectively. Solid lines (in blue and red colour) represent markers that showed significant associations with traits of interest. Map distances are given in cM and displayed using the MapChart.

a



b


DTF
SY
Δ13C
DTF
C09(Genetic)
C09(Physical)
A10(Physical)
3080553 82.48
3100726|F|0-64:T>C-64
3103337|F|0-5:T>C-5
3158481|F|0-31:A>G-31 84.24
3199416|F|0-17:C>T-17
3218575
5707867
4166761|F|0-43:C>T-43
4166761
3079009|F|0-7:T>C-7 88.13
3158874
3141258|F|0-22:T>C-22
3122101|F|0-58:C>G-58
3094517 94.50
5053896|F|0-29:A>G-29 94.85
5850737 96.26
3096040|F|0-56:T>A-56 96.61
3091907 97.67
3100726|F|0-64:T>C-64 45.51
BnaC09g45520D
3103337|F|0-5:T>C-5 45.83
3158481|F|0-31:A>G-31
3199416|F|0-17:C>T-17 45.84
3218575 45.85
BnaC09g46260D
BnaC09g46270D 46.21
BnaC09g46440D 46.28
BnFLC1(BnaC09g46500D) 46.35
BnaC09g46530D
BnFLC1(BnaC09g46540D) 46.37
BnaC09g46550D
BnaC09g46560D 46.38
BnaC09g46650D 46.43
BnaC09g46830D 46.50
BnaC09g46950D 46.56
HAC11(BnaC09g46960D) 46.57
BnaC09g46990D
5707867 46.58
BnaC09g47020D 46.60
4166761|F|0-43:C>T-43
4166761 46.61
3079009|F|0-7:T>C-7
BnaC09g47060D
3122101|F|0-58:C>G-58
3158874 46.62
3141258|F|0-22:T>C-22
BnaC09g47240D 46.70
BnaC09g47300D 46.72
CCA1(BnaC09g47560D) 46.83
BnaC09g47650D 46.85
5875882
ERL-2(BnaC09g48340D) 47.18
ERL-1(At5G62230)
3094517 47.25
3094517|F|0-22:C>G-22
5850737 47.43
FTIP1(BnaC09g48710D) 47.45
LEA(BnaC09g48810D) 47.50
ABCG22(BnaC09g49020D) 47.60
3103337|F|0-5:T>C-5 14.73
3158481|F|0-31:A>G-31
3218575 14.74
3199416|F|0-17:C>T-17
BnFLC1(AY036888.1) 15.00
HAC11(BnaA10g22400D) 15.12
5707867 15.13
3122101|F|0-58:C>G-58
3158874 15.16
3141258|F|0-22:T>C-22
5875882 15.56
3094517
3094517|F|0-22:C>G-22 15.62
BnaA10g24180D 15.82
ABCG22(BnaA10g2440D) 15.94