nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-021-01486-0

Check for updates

Ecological factors influence balancing selection on leaf chemical profiles of a wildflower

Lauren N. Carley[1,2,3,4], Julius P. Mojica[2,5], Baosheng Wang[2,6], Chia-Yu Chen[7,8], Ya-Ping Lin[7,9], Kasavajhala V. S. K. Prasad[10], Emily Chan[2], Che-Wei Hsu[7,11,12], Rose Keith[2,13], Chase L. Nuñez[1,14], Carrie F. Olson-Manning[2,15], Catherine A. Rushworth[2,4,16,17], Maggie R. Wagner[2,18,19], Jing Wang[2,6], Pei-Min Yeh[7], Michael Reichelt[20], Kathryn Ghattas[2], Jonathan Gershenzon[20], Cheng-Ruei Lee[7,21,22] ✉ and Thomas Mitchell-Olds[2,3] ✉

Balancing selection is frequently invoked as a mechanism that maintains variation within and across populations. However, there are few examples of balancing selection operating on loci underpinning complex traits, which frequently display high levels of variation. We investigated mechanisms that may maintain variation in a focal polymorphism—leaf chemical profiles of a perennial wildflower (*Boechera stricta*, Brassicaceae)—explicitly interrogating multiple ecological and genetic processes including spatial variation in selection, antagonistic pleiotropy and frequency-dependent selection. A suite of common garden and greenhouse experiments showed that the alleles underlying variation in chemical profile have contrasting fitness effects across environments, implicating two ecological drivers of selection on chemical profile: herbivory and drought. Phenotype–environment associations and molecular genetic analyses revealed additional evidence of past selection by these drivers. Together, these data are consistent with balancing selection on chemical profile, probably caused by pleiotropic effects of secondary chemical biosynthesis genes on herbivore defence and drought response.

How genetic variation is maintained despite persistent natural selection is a central question in evolutionary biology. Theory demonstrates that both directional and stabilizing selection should reduce genetic variation within populations over time[1]. Since methods for quantifying phenotypic selection were standardized[2], hundreds of studies and thousands of measurements of selection have revealed that directional selection is common in nature[3]. Nevertheless, high levels of additive genetic variance are frequently documented within and among natural populations[4,5]. In addition to posing a fundamental puzzle, the persistence of widespread genetic diversity despite selection directly influences future adaptation: response to selection depends on the amount of genetic variation that exists for adaptive traits[1,4]. Understanding whether and how populations can adapt to changing selective pressures is of critical importance[6–8], especially as environmental change accelerates in the Anthropocene[9].

Some models for persistent genetic diversity (for example, ref. [10]) suggest that a balance between new mutations and purifying selection may explain much of complex trait variation. However, recent experiments provide limited support for mutation–selection balance as a general explanation for standing genetic variation[11,12]. An alternative explanation in these cases is balancing selection, in which natural selection actively maintains genetic diversity within or among populations[13,14]. This process can maintain trait variation with simple genetic bases, including flower colour[15] and self-incompatibility[16] in plants and some immune traits in humans[17]. However, evidence that balancing selection maintains complex trait variation in nature is limited (but see ref. [18]). Furthermore, balancing selection can be generated by a variety of mechanisms, including negative frequency-dependent selection, overdominance, selection that is temporally variable within populations or spatially variable among populations, and antagonistic pleiotropy[19]. The relative importance of these mechanisms remains understudied, and ecological interactions generating them are often unclear[20]. Predicting evolutionary outcomes in heterogeneous environments requires a mechanistic understanding of how ecological interactions influence fitness[21] and genetic diversity.

Here, we explicitly test multiple ecological and genetic mechanisms that may contribute to balancing selection on a locus influencing complex trait variation. We focus on anti-herbivore defence,

[1]Duke University Program in Ecology, Durham, NC, USA. [2]Biology Department, Duke University, Durham, NC, USA. [3]Rocky Mountain Biological Laboratory, Gothic, CO, USA. [4]Department of Plant and Microbial Biology, University of Minnesota Twin Cities, St Paul, MN, USA. [5]Pairwise Plants, Durham, NC, USA. [6]Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China. [7]Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei, Taiwan. [8]Experimental and Clinical Research Center (ECRC) of the MDC and Charité Berlin, Berlin, Germany. [9]World Vegetable Center Headquarters, Tainan, Taiwan. [10]Department of Biology and Cell and Molecular Biology Program, Colorado State University, Fort Collins, CO, USA. [11]Department of Biology, Humboldt Universität zu Berlin, Berlin, Germany. [12]The Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Berlin, Germany. [13]Biology Department, DePauw University, Greencastle, IN, USA. [14]Department for the Ecology of Animal Societies, Max Planck Institute of Animal Behavior, Baden-Württemberg, Germany. [15]Augustana University, Sioux Falls, SD, USA. [16]Evolution and Ecology Department, University of California Davis, Davis, CA, USA. [17]University and Jepson Herbaria, University of California Berkeley, Berkeley, CA, USA. [18]Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, USA. [19]Kansas Biological Survey, Lawrence, KS, USA. [20]Department of Biochemistry, Max Planck Institute for Chemical Ecology, Jena, Germany. [21]Institute of Plant Biology, National Taiwan University, Taipei, Taiwan. [22]Genome and Systems Biology Degree Program, National Taiwan University, Taipei, Taiwan. ✉e-mail: chengrueilee@ntu.edu.tw; tmo1@duke.edu

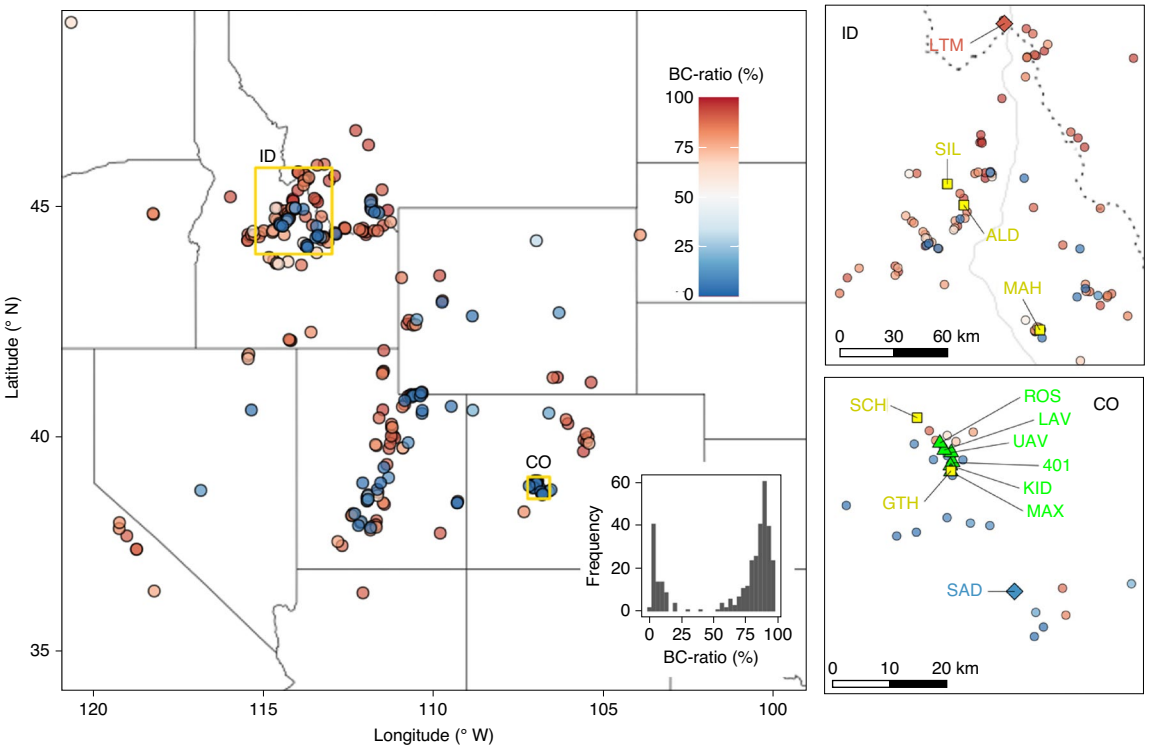


Fig. 1 | GS variation in *B. stricta* is highly polymorphic and widespread. BC-ratio is a quantitative trait that is bimodally distributed (inset) and intermixed geographically across the species range, as shown by this survey of 337 accessions derived from natural populations. The boxes show close-range views of the northern (ID, Idaho) and southern (CO, Colorado) focal regions, showing the locations of the common gardens (filled yellow squares), temporary arrays (green triangles) and parental genotypes used for crosses and genome assembly (red and blue diamonds) in this study.

a complex plant trait comprising many constituent phenotypes—for example, leaf toughness and hairiness[22], life history, size and architecture[23], and primary and secondary metabolites[24,25]—all of which may have their own complex genetic architectures. In *Boechera stricta* (Brassicaceae), a wild relative of *Arabidopsis*, glucosinolate (GS) secondary metabolites contribute to insect resistance and fitness; furthermore, the proportion of aliphatic GS derived from branched-chain amino acid (Val and Ile) precursors, called BC-ratio, is an important axis of GS variation controlled by a known biosynthetic locus[26]. BC-ratio in *B. stricta* is thus an intermediary physiological trait linking complex trait variation to a tractable genetic basis. Past work has documented the molecular evolution that allowed for the diversification of BC-ratio in *B. stricta*[26], but fine-scale patterns of chemical variation, as well as the specific mechanisms contributing to selection on this trait, remained unknown. Here, we report field experiments, greenhouse experiments and molecular genetic analyses that test for balancing selection on this biochemical polymorphism, interrogating multiple ecological and genetic mechanisms that may drive it, including spatial and temporal variation in selection, antagonistic pleiotropy and frequency-dependent selection.

Results

GS variation is widespread in nature. To assess variation in BC-ratio across the species range, we used high-performance liquid chromatography (HPLC) to characterize the GS profiles of accessions collected from 337 wild *B. stricta* populations grown in a common greenhouse environment. These new data provide the most detailed picture of GS variation in *B. stricta* to date. While past research has shown that BC-ratio varies spatially[26], this fine-scaled geographic survey shows that BC-ratio is highly polymorphic across the species range, often at small geographic scales (Fig. 1). Because of this, we investigated a variety of evolutionary processes[13,27–29] that may contribute to balancing selection on this trait.

***BCMA1/3* alleles confer contrasting fitness effects across environments.** The tandemly duplicated *BCMA1/3* genes control BC-ratio in *B. stricta* by modulating the first step in the core aliphatic GS biosynthesis pathway[26]. To determine which processes may influence balancing selection on BC-ratio, we generated near-isogenic lines (NILs) derived from a largely homozygous F4 individual that was heterozygous for a narrow genomic region containing *BCMA1/3* (refs. [26,30]). We screened nearby PCR markers on F5 progeny, and two homozygous F6 closest-flanking recombinants (CFRs) were crossed together, yielding F1 and F2 CFR-NILs. Multiple independent F3 families provide replicated, homozygous CFR-NILs differing at ten loci adjacent to *BCMA1/3* (Methods). We used these CFR-NILs to test the effects of contrasting homozygous *BCMA1/3* haplotypes conferring methionine-derived GS or branched-chain-amino-acid-derived GS (homozygous *MM* and *BB* genotypes, respectively) on components of fitness in the field and laboratory.

Across 15 common garden field environments spanning 780 km, 11 sites (Fig. 1, inset) and four years, there was significant variation in the effects of *BCMA1/3* alleles on insect resistance (Supplementary Table 1). Contrasting alleles showed changes in rank resistance across environments, with the *MM* and *BB* genotypes each conferring greater protection in some environments, and also differentially influenced resistance to a model herbivore (*Trichoplusia ni* (Lepidoptera: Noctuidae)) in the laboratory (Fig. 2a). In *B. stricta*, herbivore damage decreases fitness by reducing reproductive output (Fig. 2b and Supplementary Table 2); thus, GS traits that influence insect resistance are subject to variable selection by insect herbivores across environments.

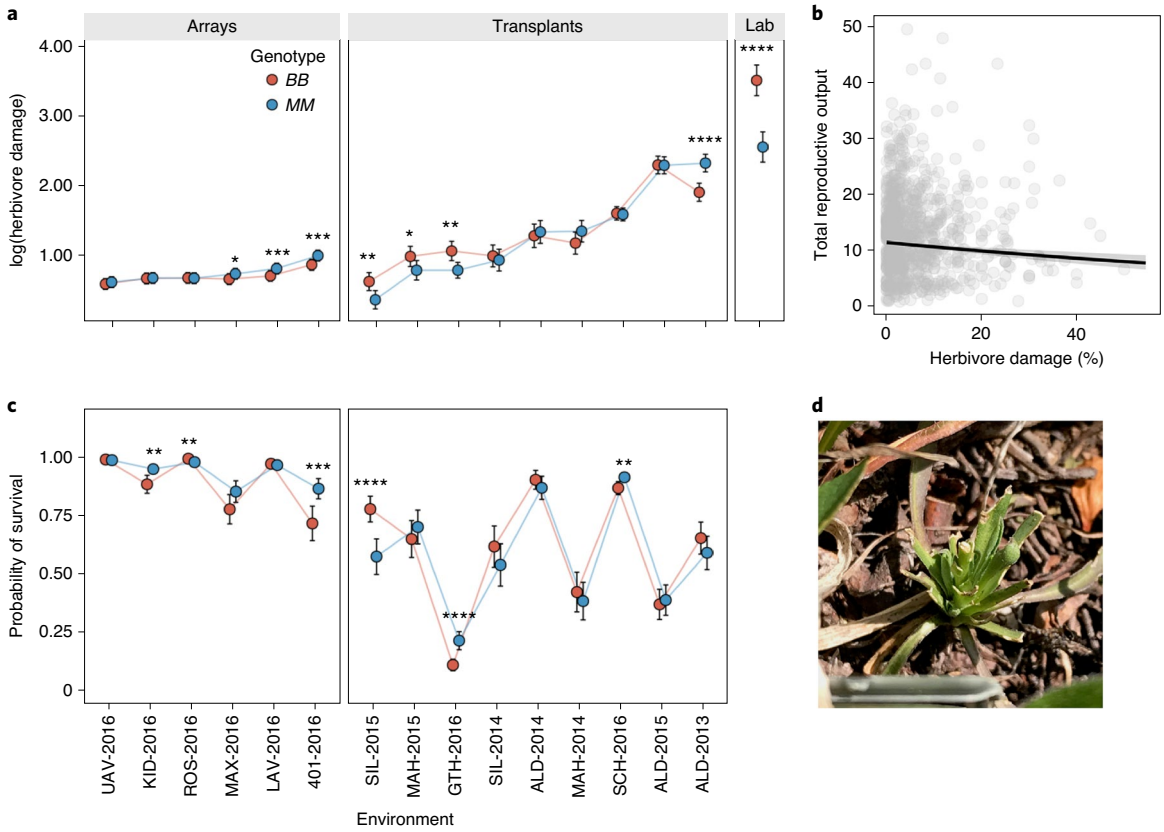


Fig. 2 | Environmental variation in the effects of GS on fitness components. a, Across 16 environments, contrasting *BCMA1/3* alleles conferred variable effects on insect resistance (Supplementary Table 1; *BCMA1/3* × environment: $F_{5,5150} = 2.3623$, $P = 0.0376$; $F_{8,2840.39} = 4.1884$, $P < 0.0001$; $F_{1,6.8801} = 35.772$, $P = 0.0006$ in arrays, transplants and laboratory, respectively). **b**, Herbivore damage reduces fitness by decreasing fruit size among plants that reproduce (Supplementary Table 2; damage: $F = 5.5318$, $P = 0.01867$). **c**, In field environments (presented in order of herbivore damage, as in **a**), contrasting *BCMA1/3* alleles also show variable effects on survival (Supplementary Table 3; *BCMA1/3* × environment: $\chi^2 = 27.839$, $P = 0.0010$ and $\chi^2 = 61.769$, $P = 0.0010$ in arrays and transplants, respectively). **d**, Photo of a field transplant showing substantial herbivore damage to leaf tissue and the flowering stalk. In **a** and **c**, different experiments are separated in boxes; left to right, temporary field arrays, permanent field transplants and controlled laboratory conditions following a challenge with model herbivore *T. ni*. Within environments, the circles show the least-squares means of each fitness component after accounting for other model effects (Supplementary Tables 1–3), and the error bars show ±1 standard error. The asterisks indicate pairwise significant differences among *BCMA1/3* genotypes within each environment as follows: $^*P < 0.10$; $^{**}P < 0.05$; $^{***}P < 0.01$; $^{****}P < 0.0001$. In **b**, the black curve represents the best-fit line from linear regression, and the grey shaded areas represent the 95% confidence interval along the curve; herbivore damage was log-transformed in the linear model, and here the linear predictions are back-transformed to the raw data scale.

BCMA1/3 alleles also influenced survival in the field (Supplementary Table 3), with the *MM* and *BB* genotypes changing in rank survival across environments (Fig. 2c). However, herbivore resistance did not directly influence survival in these experiments (Supplementary Table 4). Survival trade-offs across environments thus implicate other ecological forces, along with herbivory, in shaping selection on BC-ratio.

Finally, by manipulating *BCMA1/3* genotype frequencies in six field environments, we tested for frequency-dependent effects of *BCMA1/3* alleles. Genotype frequency did not affect either insect resistance or survival (Extended Data Fig. 1 and Supplementary Tables 1 and 3). Thus, spatial variation in selection, rather than frequency-dependent selection, may maintain GS polymorphisms in this species.

Drought influences selection on *BCMA1/3*. Recent work on *Arabidopsis* has revealed that GSs can modulate drought response by regulating stomatal aperture[31,32]. In a subset of our field experiments showing evidence of drought stress (Fig. 3a), we tested whether drought also shapes selection on BC-ratio. Field arrays experiencing stronger viability selection via drought showed significant increases in the frequency of the *MM* genotype (Supplementary Table 5 and Fig. 3b). Thus, in addition to herbivory, drought influences selection on BC-ratio, and *MM* alleles seem to confer higher fitness under drought stress.

In a controlled dry-down greenhouse experiment, we found that variation in drought response among CFR-NIL genotypes is underpinned by differing morphological responses to drought. Specifically, the *MM* genotype decreases its size under drought compared with controls, while *BB* does not (Fig. 3c,d, Extended Data Fig. 2 and Supplementary Table 6). Because size drives water use (Fig. 3e,f and Supplementary Table 7), this suggests that *MM* alleles reduce water use in response to drought, which may yield higher fitness under drought stress in the field.

To determine whether GSs also influence drought response in a broad range of genetic backgrounds, we performed a second greenhouse dry-down experiment with a panel of 350 accessions from across the species range (Fig. 1). BC-ratio was genetically correlated with leaf water content under drought (Supplementary Table 8): low-BC-ratio accessions maintained higher leaf water content under drought (Fig. 3g). Finally, in this diverse panel of accessions, precipitation-related climate variables in home

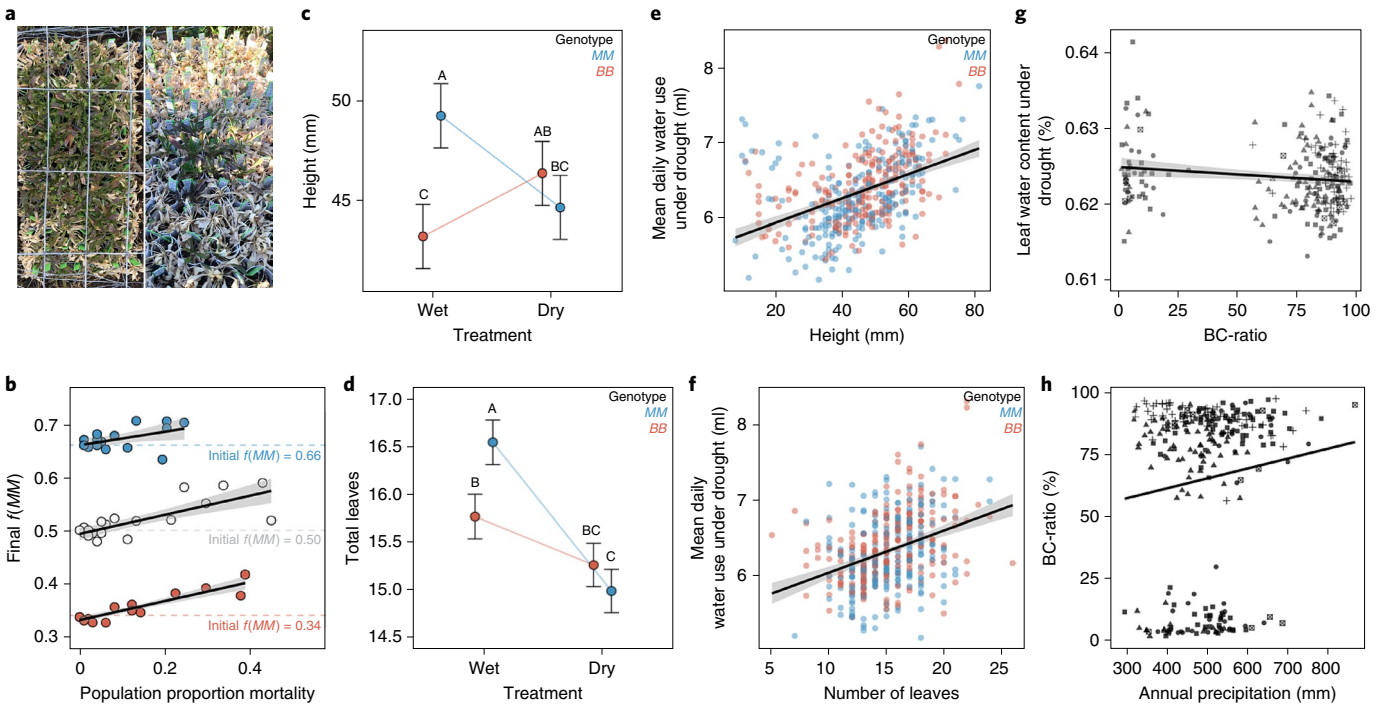


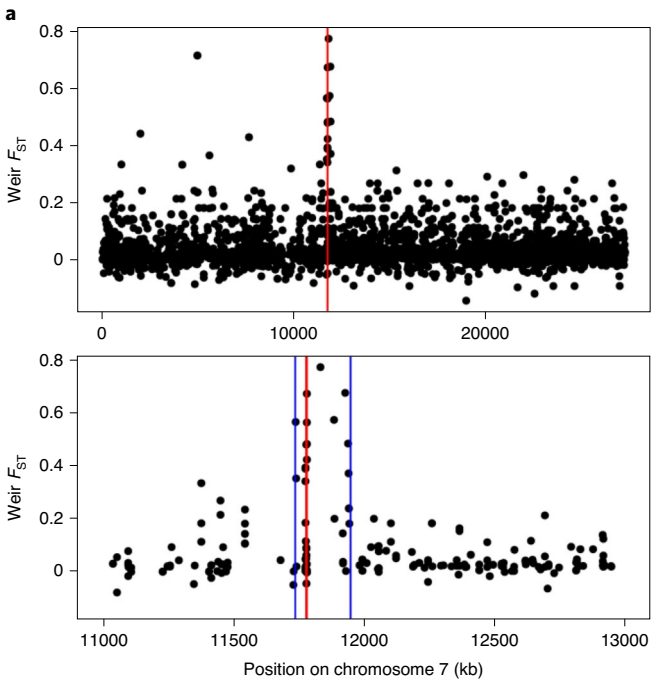
Fig. 3 | Drought influences selection on *BCMA1/3*. **a**, Drought stress in the field varies across sites (left, low; right, high). **b**, In the Colorado field arrays, the frequency of the *MM* genotype (*f*(*MM*)) increased significantly in arrays that experienced stronger mortality following drought (effect of the proportion mortality on *f*(*MM*): $F = 21.8870$, $P < 0.0001$; Supplementary Table 5). **c**,**d**, The *MM* CFR-NIL genotype suppresses growth (height (**c**) and leaf number (**d**)) under drought, relative to well-watered conditions (*BCMA1/3* × environment: $\chi^2 = 27.32$, $P < 0.0001$ and $\chi^2 = 8.52$, $P = 0.0035$ for height and leaf number, respectively; Supplementary Table 6). The *BB* genotype does not. **e**,**f**, Phenotypes that *MM* genotypes suppress under drought (height (**e**) and leaf number (**f**)) are positively correlated with water use under drought ($\chi^2 = 12.8290$, $P = 0.0003$, $\beta = 0.0074$ and $\chi^2 = 44.0913$, $P < 0.0001$, $\beta = 0.0480$ for height and leaf number, respectively; Supplementary Table 7). **g**, Among 237 accessions, BC-ratio is genetically correlated with leaf water content under drought ($F_{1,231} = 8.17$, $P = 0.0046$; Supplementary Table 8). **h**, Precipitation in 283 home environments is positively correlated with BC-ratio ($F = 12.68$, $P = 0.0008$, $\beta = 0.074$; Supplementary Table 11)—that is, Met-GS are more common in drier habitats. In **b** and **e**–**h**, the solid lines and grey shaded areas indicate the linear predictions and 95% confidence intervals, respectively; in **h**, the linear prediction reflects coefficient estimates modelled using genetic groups to control for population structure. In **b**, the points represent summary statistics from arrays with genotype frequency shown by the colour of the points (blue indicates high, or 66% *MM;* white indicates medium, or 50% *MM*; and red indicates low, or 34% *MM*). The horizontal dashed lines show the starting genotype frequency in each treatment. In **c**,**d**, the points represent least-squares mean estimates of traits in wet and dry greenhouse environments following a two-week progressive dry-down; the error bars show ±1 standard error. The uppercase letters denote pairwise significant differences among means (Tukey's honestly significant difference); means not connected by the same letter are significantly different ($P < 0.05$). In **e**,**f**, the least-squares mean daily water use under drought for each individual plant during dry-down is shown, plotted against individual size (height (**e**) and leaf number (**f**)). The points are colour-coded by *BCMA3* genotype (blue indicates *MM*; red indicates *BB*). In **g**,**h**, the shapes denote broad genetic groups (circles indicate COL, triangles indicate NOR, squares indicate UTA, crosses indicate WES and x-boxes indicate Admixed).

environments predict the distribution of BC-ratio phenotypes across the landscape (Supplementary Tables 9–11 and Extended Data Figs. 3 and 4). The best climatic predictor of BC-ratio was annual precipitation (Supplementary Table 11); natural accessions producing low-BC-ratio GS phenotypes were found in drier environments (Fig. 3h). After controlling for population structure, this genotype–environment association is evidence for large-scale geographically heterogeneous selection on BC-ratio by climate[33].

Our findings build on recent evidence for the involvement of GS in drought response[31,32], suggesting that *BCMA1/3* alleles altering BC-ratio confer contrasting morphological responses to drought stress. Thus, environmental variation in drought stress as well as herbivory may favour diverse GS profiles across the landscape.

Linkage disequilibrium decays rapidly around *BCMA3*. Tightly linked loci near *BCMA1/3* might covary with *BCMA1/3* alleles, possibly contributing to the observed phenotypic patterns. Using a de novo long-read assembly around *BCMA1/3* (Methods), we scanned patterns of genetic differentiation (F_{ST}) along chromosome 7 to identify flanking regions in which the *MM* and *BB* CFR-NIL haplotypes have divergent single nucleotide polymorphisms (SNPs). This revealed the *BCMA1/3* haplotype to be 212 kilobases (kb) in length, spanning positions 11,737 to 11,949 kb on chromosome 7, extending 42 kb in the 5′ direction and 167 kb in the 3′ direction of *BCMA3* (Fig. 4a). A total of 41 SNPs fall within this haplotype.

The version_2 SAD12 reference genome shows a total of 11 genes, including *BCMA1/3*, in this CFR-NIL interval. Prior studies have shown that, besides *BCMA1/3*, several of these genes may contribute to drought response or change expression in response to drought (Fig. 4b). Other loci within the nonrecombinant interval could thus influence the phenotypes observed in our common garden experiments. To determine whether flanking genes covary with *BCMA3* in nature, we used sequence data from a range-wide panel of natural accessions[34] to characterize linkage disequilibrium (LD) around *BCMA3*. In geographic areas with high genetic diversity (COL and UTA[34]), LD surrounding *BCMA3* is low (Fig. 4c). Thus, while nearby regions cosegregate with *BCMA3* in the CFR-NILs, low LD between *BCMA3* and surrounding regions suggests that



b

kb from *BCMA3*	*At* homologue (alias)	Annotation (reference)
–41	ATG14000	Rapid alkalinization factor (RALF)-like protein (refs. [57,58])
0	**AT1G16410 (*CYP79F1*)**	**Cytochrome P450 family 79, subfamily C (ref. [31])**
49	ATG14010	Rapid alkalinization factor (RALF)-like protein (refs. [57,58])
98	AT4G14020	Rapid alkalinization factor (RALF)-like protein (refs. [57,58])
104	AT4G14030 (*SBP1*)	Selenium-binding protein 1 (ref. [59])
106	AT4G14040 (*EDA38*)	Selenium-binding protein 2 (ref. [59])
109	**AT4G14050 (*MEF35*)**	**Pentatricopeptide repeat (PPR) superfamily protein (ref. [60])**
138	AT4G14060	Pathogenesis-related protein Bet v I family (ref. [61])
161	AT4G14070 (*AAE15*)	Long-chain-fatty-acid ligase/acyl-synthetase (ref. [62])
166	AT4G14080 (*MEE48*)	Beta-1,3-glucanase related (ref. [63])
168	AT4G14100	None

Fig. 4 | The *BCMA1/3* CFR-NIL interval contains ten flanking loci, but they show little correlation (LD) in natural populations. a, Comparisons of genotyping by sequencing (GBS) data between *BB* and *MM* CFR-NIL genotypes show an F_{ST} peak around the *BCMA3* gene (red line), revealing a ~212-kb non-recombinant interval (delimited by blue lines) among the CFR-NILs. **b**, Eleven *B. stricta* genes occur in the non-recombinant CFR-NIL haplotype, including *BCMA3*[31,57–63]. The shaded rows indicate that homologues in *A. thaliana* (*At*) have variable expression in response to drought. The shaded row with bold text indicates that *At* homologues have impacts on drought response, which have been validated with functional genetic studies. The shaded row with bold red text indicates that *At* homologues have functionally verified effects on both insect resistance and drought response. **c**, Heat maps showing pairwise LD (r^2) in a 712-kb interval surrounding *BCMA3*, for three groups of accessions: 157 COL accessions, 126 UTA accessions and 233 pooled COL + UTA accessions. The blue numbers on the diagonals indicate the position on chromosome 7 (in base pairs) of the *BCMA3* gene and the closest SNP to the limits of the surrounding non-recombinant region in the CFR-NILs. In **b**, all *B. stricta* gene positions reflect the version_2 SAD12 reference genome and fall on chromosome 7.

linked loci are unlikely to explain genetic correlations between BC-ratio and drought response (Fig. 3g) or phenotype–environment associations between BC-ratio and drought in home environments (Fig. 3h).

Molecular signatures are consistent with balancing selection on *BCMA*. Prior research identified functional amino acid substitutions in the BCMA3 enzyme that influence BC-ratio[26]. To characterize genetic variants underlying BC-ratio phenotypes, we performed Sanger sequencing on *BCMA3* in a panel of 110 accessions. In addition to the previously described amino acid substitutions altering enzyme activity[26], at least two structural variants are present in *BCMA3* (Supplementary Table 12). An eight-base-pair deletion causes a frameshift mutation upstream of both amino acid substitutions of interest, disrupting the BCMA3 enzyme after 137 amino acids. Additionally, several accessions showed premature stop codons later in the amino acid sequence caused by nonsense mutations. All premature stop codons are associated with low BC-ratio (Fig. 5a and Supplementary Table 12), presumably because they disrupt the function of the BCMA3 enzyme. These derived mutations are geographically widespread; among the accessions we assayed, we found stop codon only and stop codon + deletion variants spanning 328 km and 729 km of the species range, respectively (Fig. 5b). Given low seed dispersal[35] and outcrossing[36] in *Boechera*, widespread variants in natural habitats suggest that the functional genetic bases of variation in GS phenotypes have been maintained over long periods and are not ephemeral novel mutants.

Finally, a molecular hallmark of balancing selection is elevated nucleotide variation within genes under selection[13]. We characterized nucleotide diversity (π) within *BCMA3* and genes of similar

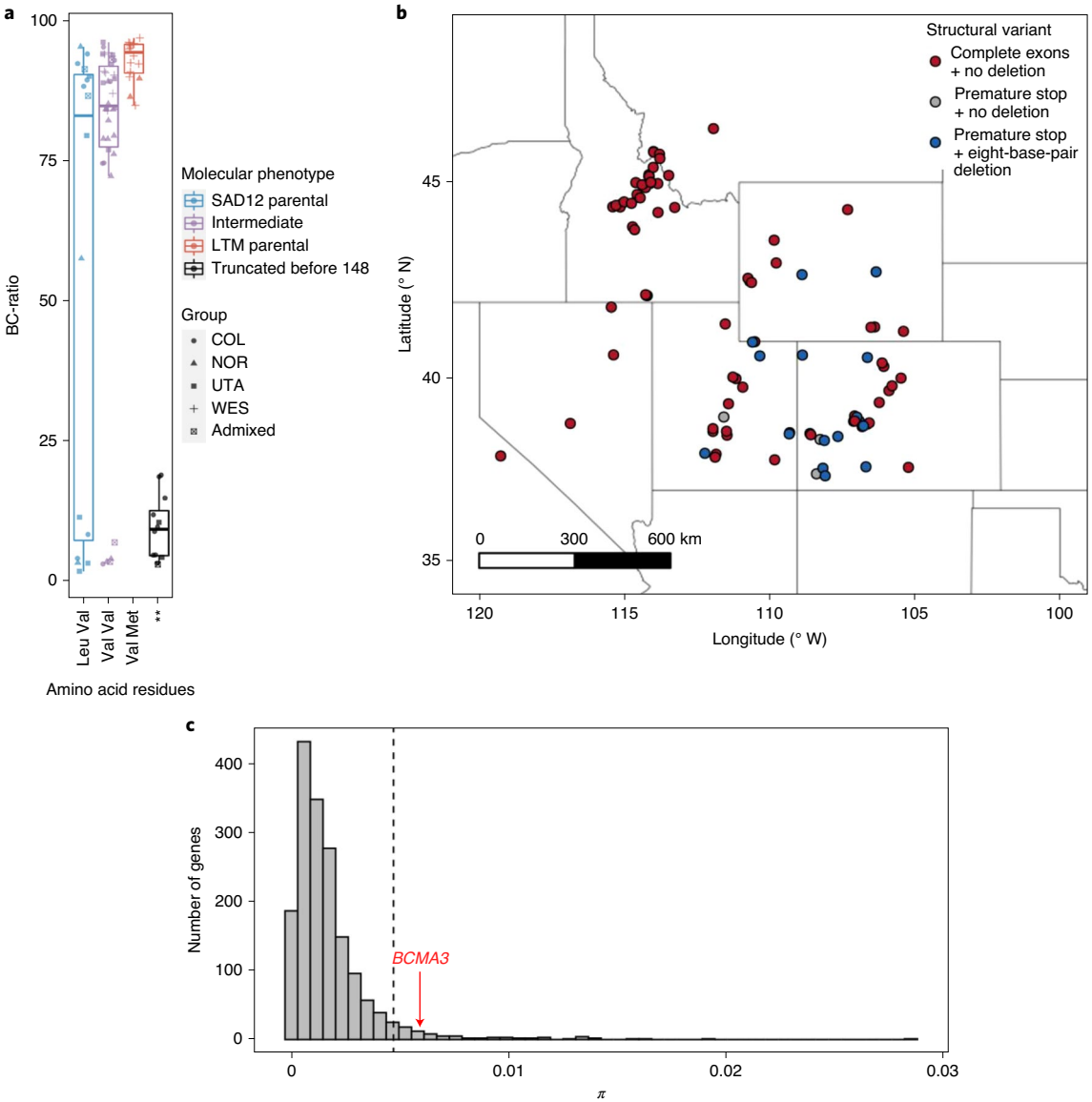


Fig. 5 | Molecular evidence of balancing selection on *BCMA3*. **a**, Sanger sequencing of *BCMA3* alleles reveals structural genetic variation influencing BC-ratio; the presence of a premature stop codon disrupting BCMA3 expression significantly predicts some low-BC-ratio phenotypes ($F = 6.8141$, $P = 0.0110$; Supplementary Table 12), in addition to critical amino acid residues in positions 148 and 268 of BCMA3 (ref. [26]). The data points represent the genetic mean BC-ratio for each accession, and the shapes denote genetic groups[37]. The box plots delineate the 25th, 50th and 75th percentiles with boxes and 1.5× the interquartile range with whiskers. The colours show the BCMA3 enzyme molecular phenotype, with Leu148/Val268 (blue) matching the SAD12 parental genotype and Val148/Met268 (red) matching the LTM parental genotype. Purple accessions have critical amino acid residues at 148/268 that are intermediate to the CFR-NIL parental genotypes. Black accessions have premature stop codons (asterisks) upstream of position 148. **b**, Genetic variants underlying BC-ratio are geographically widespread across the species range. **c**, *BCMA3* has high nucleotide diversity ($\pi_{BCMA3} = 0.00591$) relative to comparable genes in the *B. stricta* genome ($\bar{\pi} = 0.00169$); π_{BCMA3} (red arrow) exceeds the 5% tail (dashed line) of 1,689 comparable genes in the *B. stricta* genome ($P = 0.0290$).

size across a panel of 54 accessions. Consistent with predictions for balancing selection, we found that *BCMA3* is more polymorphic than 97.1% of 1,689 comparable loci across the genome (Fig. 5c).

Effect size of *BCMA1/3*. Since we aim to understand selection on genes that influence complex trait variation, we estimated the average effect of the *BCMA1/3* polymorphism on herbivore damage in nature. The mean effect size of *BCMA1/3* on herbivore damage was 0.172 standard deviations, indicating that it is a small-effect quantitative trait locus[37] influencing herbivore damage and fitness in nature.

Discussion

The degree to which different evolutionary processes may explain the maintenance of polymorphism has been the subject of debate for decades[38,39]. Our experiments document spatial variation in selection driving evolution of the *BCMA1/3* polymorphism in *B. stricta*. Specifically, functional links between genetic variation in *BCMA1/3*, chemical (GS) and physiological (drought tolerance) functional traits, and fitness in nature reveal trade-offs of *BCMA1/3* alleles across environments. Changes in the sign of allelic effects on fitness across environments are hallmarks of balancing selection[13,40,41]. Our data thus show that balancing selection has

maintained complex trait variation in GS profiles across the landscape. This conclusion is corroborated by molecular analyses, which reveal that nucleotide diversity within *BCMA3* is elevated compared with other genes in this species and that GS phenotypes in common garden are genetically correlated with drought conditions in home environments. Finally, while the *BCMA1/3* polymorphism evolved after *B. stricta* diverged from its close congener (Extended Data Fig. 5), genetic variants underlying GS profiles are geographically widespread, suggesting that selection has maintained multiple variants over time.

While we see clear evidence of functional trait and fitness trade-offs conferred by *BCMA1/3* alleles across environments, alleles also had synergistic effects on fitness components in some environments; conducting our experiments in a subset of sites or years could have led to the erroneous conclusion that only one allele was favoured by selection. For example, in the GTH-2016 and 401-2016 environments, the *MM* allele confers both higher survival and reduced herbivore damage (Fig. 2a,c). Conditional neutrality was also common; in 7 of 15 tested field environments, contrasting *BCMA1/3* alleles conferred no detectable differences in either insect resistance or survival. Furthermore, in many environments in which fitness trade-offs were detectable, the magnitude of allelic effects was small. These findings suggest that experiments must deploy large sample sizes in many environments to detect changes in allelic effects when they do occur[42,43] and that the relative lack of empirical evidence for balancing selection may be influenced by limited statistical power. Beyond this practical concern, these findings reveal biologically relevant environmental variation in the expression of trade-offs, emphasizing the critical importance of environmental context in understanding complex patterns of natural selection.

Our results are consistent with examples of quantitative trait loci expressing contrasting effects on fitness under variable laboratory and field conditions[44,45], as well as pleiotropic effects of SNPs across environments contributing to polygenic trait variation[18]. Our work builds on these studies by offering explanations of functional consequences for allelic variation across environments on complex phenotypes that underlie fitness. Such mechanistic understanding has been achieved previously in now-classic studies focusing on traits with simple genetic architecture, including flower colour[15] and self-incompatibility[16] in plants, as well as *HLA* loci in humans[17]. While we have emphasized the effects of a gene of interest, the ecological and genetic signatures of selection on *BCMA1/3* show that tractable genes with large physiological effects may have small effects on ecologically important traits, thus influencing genetic variation of complex traits in nature. Similarly, approaches emphasizing tractable intermediary phenotypes have helped identify genetic bases of human metabolic variation[46].

In addition to revealing the genetic basis of cross-environment fitness trade-offs that drive balancing selection on *BCMA1/3*, we identify two key drivers of selection on GS evolution in *B. stricta*: herbivore pressure and drought stress. The magnitude of selection exerted by each of these probably varies and/or covaries across the landscape. For example, drought stress may alter herbivore pressure directly by changing arthropod abundance or activity, or indirectly by influencing plant growth and secondary metabolites. This could yield complex, non-additive fitness landscapes influencing trait variation across environments. We encourage further research into the implications of ecological interactions among selective drivers for the maintenance of polymorphism.

Finally, our results build on evidence that GSs are multifunctional biomolecules that, in conjunction with hormones such as auxin and IAA, help coordinate organismal responses to both biotic and abiotic stressors[31,47–49]. While classically known for mediating biotic interactions with herbivores and microbes[50], recent functional genetics studies in *Arabidopsis* have shown that mutant lines with knockouts of *BCMA1/3* orthologues have reduced drought tolerance[31] due to the effects of GS degradation products on stomatal aperture[32]. This study suggests that allelic variation in GS molecular structure, in addition to the presence/absence or up/down-regulation of GS, may elicit differential responses to drought as well as to herbivores. Ultimately, traits that mediate biotic and abiotic interactions, such as GS, are likely to experience complex patterns of natural selection that may decrease the likelihood of a single variant rising to fixation. These results thus highlight how multiple ecological drivers of selection can influence balancing selection on complex traits.

Methods

Characterizing GS variation across the landscape. We characterized the chemical profiles of 337 natural accessions collected from across the species range[34]. Progeny of wild-collected seeds were self-pollinated under controlled greenhouse conditions to minimize maternal effects. We then measured GS profiles using HPLC following ref. [51] and estimated the least-squares mean value of BC-ratio per accession by fitting a restricted maximum likelihood model with accession ID and greenhouse block as random effects (Supplementary Methods, 'HPLC').

Generation of CFR-NILs. We chose wild-collected accessions with high- and low-BC-ratio phenotypes (LTM and SAD12 from Montana and Colorado, respectively; Fig. 1) to generate the experimental genotypes used here via a crossing pedigree described in Extended Data Fig. 6. Previously[26], we identified a near-isogenic F4 derived from the SAD12 × LTM cross, screened 5,213 F5s for recombination near *BCMA1/3* and then scored 13 tightly linked PCR markers on 205 homozygous F6 recombinants. Here, we identified two F6 homozygous CFRs, which were crossed together, generating a heterozygous double-recombinant F1. In the F2, we self-pollinated multiple *MM* and *BB* homozygotes, yielding F3 families representing replicated homozygous *MM* and *BB* CFR-NILs (hereafter, families). Using GBS (Supplementary Methods, 'Genotyping-by-sequencing'), we determined the extent of the non-recombinant *BCMA1/3* haplotype.

Laboratory herbivory experiment. We grew 225 juvenile CFR-NILs in a randomized complete block design under controlled greenhouse conditions (18–21 °C day / 13–16 °C night; 16 h day length; watering to saturation daily; fertilization at 300 ppm N weekly) for six weeks before challenging them with a model herbivore. Five blocks each contained 45 individuals, with five replicates per independent CFR-NIL family (four *BB* families and five *MM* families). We confined trays containing full statistical blocks in plexiglass chambers with mesh panels for ventilation. In these chambers, we applied a single second-instar cabbage looper larva (*T. ni*) to each plant. The larvae roamed within the chambers, feeding freely within a block for five days. We then scored each plant for insect damage as in ref. [52].

We used a restricted maximum likelihood mixed-effects model to test for the effect of genotype on resistance to *T. ni*; we fit herbivore damage in response to the fixed effect of *BCMA1/3* genotype, the random effect of block and the random effect of family nested within genotype. We tested the significance of the fixed effect using both Satterthwaite's method and Kenward-Roger's method, which yielded similar results (Supplementary Table 1), and we tested the significance of the random effects using likelihood ratio tests.

Field experiments: common gardens. We transplanted 6,860 CFR-NILs into three experimental gardens in central Idaho and two experimental gardens in southwest Colorado, near the source populations for the parental accessions (Fig. 1). We transplanted cohorts containing 360–1,350 individuals between 2013 and 2015, using a randomized complete block design. The transplants were spaced at constant density and planted directly into the surrounding vegetation. Each cohort contained replicates using at least eight CFR-NIL families to control for possible effects of unlinked loci outside of the *BCMA1/3* region. We measured herbivore damage, survival and reproduction of these individuals to estimate the effects of *BCMA1/3* alleles on fitness-related traits in different environments and to test the effects of herbivore damage on reproductive fitness.

To assess variation in insect resistance across environments ($N = 3{,}674$), we used restricted maximum likelihood mixed-effects models to test for *BCMA1/3* × environment effects on herbivore damage. We modelled herbivory in response to fixed effects of *BCMA1/3*, environment and the *BCMA1/3* × environment interaction, and random effects of statistical block and CFR-NIL line nested within *BCMA1/3* genotype. We tested the significance of the fixed effects using both Satterthwaite's method and Kenward-Roger's method, which yielded similar results (Supplementary Table 1). We tested the significance of the random effects using likelihood ratio tests.

To assess variation in survival across environments ($N = 6{,}860$), we used a generalized linear mixed-effects model (binomial distribution; logit link) to fit survival in response to the same fixed and random effects used in the herbivory model. We tested the significance of the fixed effects using both Wald tests and parametric bootstrapping, which yielded qualitatively similar

results (Supplementary Table 3), and we tested the significance of the random effects using likelihood ratio tests. In both models, if we detected a significant genotype × environment interaction, we used pairwise contrasts to compare allelic effects within each environment.

We tested for natural selection by herbivory across all transplant environments ($N = 3{,}094$) using a generalized linear mixed-effects regression with a binomial distribution and a logit link function; we fit reproductive success (0/1) in response to fixed effects of herbivore damage, environment and the damage × environment interaction, and a random effect of block. A significant negative effect of herbivory on reproduction indicates that selection favours increased herbivore resistance (Supplementary Table 2). Furthermore, in one environment with high sample sizes (SCH-2016; $N = 848$), we used a linear mixed-effects model to fit log-transformed total reproductive output of reproductive plants (number of fruits × average fruit length) in response to the fixed effect of herbivore damage and the random effect of block (Supplementary Table 2). Finally, in the same environment (SCH-2016; $N = 1{,}219$), we tested for effects of herbivore damage on survival; we used a generalized linear mixed-effects model with a binomial distribution and logit link to fit survival in 2017 in response to the fixed effect herbivore damage in 2016 and the random effect of block (Supplementary Table 4). In all models testing for natural selection by herbivory, we tested the significance of the fixed effects using Wald tests with Type III sums of squares, and of the random effects using likelihood ratio tests.

Full details about the experimental conditions and statistical analyses are provided in the Supplementary Methods under 'Common garden experiments'.

Field experiments: temporary arrays. In 2016, we deployed 5,880 CFR-NILs into six environments in Colorado (Fig. 1) to test for frequency-dependent effects of the *BCMA1/3* haplotype. We grew juvenile plants in 98-cell 'cone-tainer' racks and assigned racks to three treatments with different starting frequencies of the *MM* genotype: high (66% *MM*), medium (50% *MM*) and low (34% *MM*). The plants were randomized within each treatment rack, and we used ten CFR-NIL families to account for possible effects of unlinked loci outside the *BCMA1/3* region. We arranged ten racks (three high, four medium and three low *f*(*MM*)) in a random configuration in each of the six array sites in Colorado, and we sunk each rack flush with the soil and neighbouring vegetation. We watered the arrays twice per week for the first two weeks of the growing season and then every other day. After eight weeks in field conditions, we censused for survival and herbivore damage as described above. We used these data to test for variation across environments in the effects of *BCMA1/3* on herbivore defence and survival, as well as effects of genotype frequency on fitness components. Statistical models for herbivory ($N = 5{,}193$) and survival ($N = 5{,}880$) in the temporary arrays were identical to those described above for permanent field transplants, but they included additional fixed effects of *BCMA1/3* allele frequency and a genotype × frequency interaction (Supplementary Tables 1 and 3).

Finally, we tested for array-level response to selection. Using least-squares analysis of covariance, we regressed the array-level final *f*(*MM*) onto the proportion mortality of the arrays, fixed effects of environment and starting *f*(*MM*), and the proportion mortality × starting *f*(*MM*) interaction (Supplementary Table 5). Full details about the experimental design and statistical analyses are provided in the Supplementary Methods under 'Temporary array experiments'.

Greenhouse dry-down experiments. We performed two controlled progressive dry-down experiments in the greenhouse to test for variation in drought response among CFR-NILs (2 genotypes × 5 families per genotype × 100 replicates = 1,000 individuals) and a broad panel of genotypes (350 genotypes × 6 replicates = 2,100 individuals). In both experiments, we withheld water incrementally to simulate drought during the growing season. Importantly, we controlled for individual variation in drought stress by weighing all pots daily during the drought treatments and watering pots with the specific volumes of water required to maintain a uniform volumetric water content across replicates (Supplementary Methods, 'CFR-NIL dry-down'). We compared the morphological and physiological responses of genotypes across drought and well-watered treatments to test the effects of the *BCMA1/3* alleles (CFR-NIL experiment) and of BC-ratio (broad accession panel) on drought response.

Specifically, in the CFR-NIL experiment, we used linear mixed-effects models to fit each phenotypic response variable (Supplementary Table 6) in response to the fixed effects of genotype, drought treatment and the genotype × drought interaction, the fixed covariate of initial height, and the random effects of block and family line. We tested the significance of the fixed effects using Wald tests with Type III sums of squares, and of the random effects using likelihood ratio tests. If we detected a significant genotype × treatment effect, we assessed pairwise group differences using a Tukey's honestly significant difference post-hoc test. Furthermore, for individuals in the drought treatment, we tested whether genotype influences water use under drought using a suite of linear mixed-effects models; we fit least-squares mean daily water use (averaged over the duration of the drought treatment) in response to the fixed effects of a single morphological variable, genotype and the genotype × morphology interaction, and the random effects of block and family line (Supplementary Table 7).

In the experiment using a broad panel of accessions, we tested for genetic covariance between drought response and GS profile by regressing least-squares mean leaf water content on least-squares mean BC-ratio for each accession (see above) in a linear model. We used three different statistical approaches to control for population structure (Supplementary Methods, 'Phenotype–environment correlations'), all of which yielded similar results (Supplementary Table 8).

Full details of both experimental designs, drought treatment standardization methods and statistical analyses are provided in the Supplementary Methods.

Phenotype–environment correlations. To test for past selection on GS profile in *B. stricta*, we assessed phenotype–environment associations between BC-ratio and climatic variables from the locations of origin for a suite of accessions across the species range. First, we accounted for covariance among climate variables using principal components analysis. We then used a linear model to regress genetic mean BC-ratio for each accession on latitude, longitude and elevation of origin as well as the first five principal component axes for climate variation (Supplementary Table 9). We identified a targeted subset of climate variables that loaded highly onto the principal components axis that significantly predicted BC-ratio (Supplementary Table 10). Finally, we tested for correlations among each of these target climate variables and BC-ratio using linear models. Due to non-normal residuals, we validated the *P* values using permutation tests (Supplementary Methods, 'Phenotype–environment correlations', Supplementary Table 11 and Extended Data Fig. 4). In all of our phenotype–environment association models, we took three different statistical approaches to controlling for population structure (Supplementary Methods, 'Dry-down with broad panel of accessions: Plant measurements and data analysis'). These approaches all consistently identified climate principal component axis 1 as the best multivariate predictor of BC-ratio (Supplementary Table 9), but they sometimes differed qualitatively in identifying which raw climate variables loading onto principal component axis 1 predicted BC-ratio (Supplementary Table 11). However, all modelling approaches consistently detected significant positive correlations between BC-ratio and the following climate variables (Supplementary Table 11): annual precipitation (Fig. 3h), precipitation in the wettest month (Extended Data Fig. 3b) and precipitation in the wettest quarter (Extended Data Fig. 3b).

Dissecting the *BCMA1/3* region. To characterize genetic variation in the segregating CFR-NIL region, we assembled version_2 reference genomes for the LTM (BC-GS) and SAD12 (Met-GS) parental accessions, providing high-quality long-read coverage in the *BCMA1/3* region (Extended Data Fig. 7). We then aligned GBS reads (Supplementary Information) from 65 replicate CFR-NIL families to the SAD12 reference to identify SNPs segregating among CFR-NILs. We calculated F_{ST}[34] between *MM* and *BB* homozygotes in 20-kb non-overlapping windows along chromosome 7 and used high-F_{ST} SNPs to identify the extent of the segregating *BCMA1/3* locus in the CFR-NILs.

In addition, we aligned published Illumina sequence data[34] to the de novo version_2 SAD12 reference genome for 233 accessions from the COL and UTA genetic groups, which show high genetic diversity[34] and are polymorphic in BC-ratio. We estimated pairwise LD (r^2)[53] between each pair of SNPs within a 712-kb interval around *BCMA1/3* (spanning the 212-kb CFR region, plus 250 kb on each side; Supplementary Methods, 'Linkage disequilibrium near *BCMA1/3*').

Polymorphism in *BCMA3*. To explore the extent of genetic variation underlying the BC-ratio polymorphism, we Sanger-sequenced a subset of 110 accessions from the panel of 337 described above. After trimming and aligning the sequences (Supplementary Methods, 'Polymorphism in *BCMA3*'), we assigned the accessions to three structural variant categories: complete exons, premature stop and eight-base-pair deletion, and premature stop and no deletion. We also predicted amino acid sequences of the BCMA3 enzyme for each accession by translating the gene sequence data using the Biostrings package in R[54], and we identified the amino acid variants at positions 148 and 268, which are hypothesized to cause differential GS biosynthesis between the SAD12 and LTM genotypes[26]. We used linear models to test how structural variants and amino acid variants influence phenotypic variation in BC-ratio (Supplementary Methods, 'Polymorphism in *BCMA3*').

Molecular population genetic signatures of selection. To determine whether *BCMA3* showed molecular signatures of balancing selection, we compared the observed level of nucleotide diversity (π) in *BCMA3* with the distribution of π in a subset of comparable genes in the *B. stricta* genome across 54 accessions from the COL and UTA genetic groups, all of which have complete *BCMA3* exons (that is, they are not pseudogenes). Following ref. [34], we calculated the observed π among silent sites in *BCMA3* and other genes of similar length (1,689 genes), and we compared the value for *BCMA3* with the distribution of π across other genes (Supplementary Methods, 'Molecular signatures of selection').

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The new reference genome assemblies and raw Nanopore reads for the SAD12 and LTM genotypes have been submitted to NCBI (BioProject number PRJNA609209).

The short reads of the GBS data for the CFR-NIL families have been submitted to NCBI (BioProject number PRJNA659863). Previously published genomic data are archived with ref. [55]. All other data reported in this manuscript are archived in the Dryad digital data repository (https://doi.org/10.5061/dryad.7h44j0zsr)[56]. All biological materials are available from the Arabidopsis Biological Resource Center (ABRC) or from the authors.

Code availability

The code used for this manuscript is archived in the Dryad digital repository (https://doi.org/10.5061/dryad.7h44j0zsr)[56].

Received: 24 November 2020; Accepted: 7 May 2021;
Published online: 17 June 2021

References

1. Falconer, D. S. & Mackay, T. F. C. *Introduction to Quantitative Genetics* (Longman, 1996).
2. Lande, R. & Arnold, S. J. The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226 (1983).
3. Kingsolver, J. G., Diamond, S. E., Siepielski, A. M. & Carlson, S. M. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evol. Ecol.* **26**, 1101–1118 (2012).
4. Barrett, R. D. H. & Schluter, D. Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44 (2008).
5. Kulbaba, M. W., Sheth, S. N., Pain, R. E., Eckhart, V. M. & Shaw, R. G. Additive genetic variance for lifetime fitness and the capacity for adaptation in an annual plant. *Evolution* **73**, 1746–1758 (2019).
6. Lande, R. & Shannon, S. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* **50**, 434–437 (1996).
7. Etterson, J. R. & Shaw, R. G. Constraint to adaptive evolution in response to global warming. *Science* **294**, 151–154 (2001).
8. Anderson, J. T., Inouye, D. W., McKinney, A. M., Colautti, R. I. & Mitchell-Olds, T. Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proc. R. Soc. B* **279**, 3843–3852 (2012).
9. Steffen, W., Crutzen, P. J. & McNeil, J. R. The Anthropocene: are humans now overwhelming the great forces of nature? *Ambio* **36**, 614–621 (2007).
10. Zhang, X.-S. & Hill, W. G. Genetic variability under mutation selection balance. *Trends Ecol. Evol.* **20**, 468–470 (2005).
11. McGuigan, K., Aguirre, J. D. & Blows, M. W. Simultaneous estimation of additive and mutational genetic variance in an outbred population of *Drosophila serrata*. *Genetics* **201**, 1239–1251 (2015).
12. Huang, W. et al. Spontaneous mutations and the origin and maintenance of quantitative genetic variation. *eLife* **5**, e14625 (2016).
13. Mitchell-Olds, T., Willis, J. H. & Goldstein, D. B. Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nat. Rev. Genet.* **8**, 845–856 (2007).
14. Charlesworth, B. Causes of natural variation in fitness: evidence from studies of *Drosophila* populations. *Proc. Natl Acad. Sci. USA* **112**, 1662–1669 (2015).
15. Subramaniam, B. & Rausher, M. D. Balancing selection on a floral polymorphism. *Evolution* **54**, 691–695 (2000).
16. Charlesworth, D. Balancing selection and its effects on sequences in nearby genome regions. *PLoS Genet.* **2**, e64 (2006).
17. Hedrick, P. W. & Thomson, G. Evidence for balancing selection at HLA. *Genetics* **104**, 449–456 (1983).
18. Troth, A., Puzey, J. R., Kim, R. S., Willis, J. H. & Kelly, J. K. Selective trade-offs maintain alleles underpinning complex trait variation in plants. *Science* **361**, 475–478 (2018).
19. Delph, L. F. & Kelly, J. K. On the importance of balancing selection in plants. *N. Phytol.* **201**, 45–56 (2014).
20. Anderson, J. T., Wagner, M. R., Rushworth, C. A., Prasad, K. V. S. K. & Mitchell-Olds, T. The evolution of quantitative traits in complex environments. *Heredity* **112**, 4–12 (2014).
21. Anderson, J. T. & Wadgymar, S. M. Climate change disrupts local adaptation and favours upslope migration. *Ecol. Lett.* **23**, 181–192 (2020).
22. Agrawal, A. A. & Fishbein, M. Plant defense syndromes. *Ecology* **87**, S132–S149 (2006).
23. Carmona, D., Lajeunesse, M. J. & Johnson, M. T. Plant traits that predict resistance to herbivores. *Funct. Ecol.* **25**, 358–367 (2011).
24. DeLucia, E. H., Nabity, P. D., Zavala, J. A. & Berenbaum, M. R. Climate change: resetting plant–insect interactions. *Plant Physiol.* **160**, 1677–1685 (2012).
25. Mithöfer, A. & Boland, W. Plant defense against herbivores: chemical aspects. *Annu. Rev. Plant Biol.* **63**, 431–450 (2012).
26. Prasad, K. V. S. K. et al. A gain-of-function polymorphism controlling complex traits and fitness in nature. *Science* **337**, 1081–1084 (2012).
27. Bergelson, J., Dwyer, G. & Emerson, J. J. Models and data on plant–enemy coevolution. *Annu. Rev. Genet.* **35**, 469–499 (2001).
28. Hodgins, K. A. & Barrett, S. C. H. Female reproductive success and the evolution of mating-type frequencies in tristylous populations. *N. Phytol.* **171**, 569–580 (2006).
29. Trotter, M. V. & Spencer, H. G. Complex dynamics occur in a single-locus, multiallelic model of general frequency-dependent selection. *Theor. Popul. Biol.* **76**, 292–298 (2009).
30. Tuinstra, M. R., Ejeta, G. & Goldsbrough, P. B. Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic loci that differ at quantitative traits. *Theor. Appl. Genet.* **95**, 1005–1011 (1997).
31. Salehin, M. et al. Auxin-sensitive Aux/IAA proteins mediate drought tolerance in *Arabidopsis* by regulating glucosinolate levels. *Nat. Commun.* **10**, 4021 (2019).
32. Hossain, M. S. et al. Glucosinolate degradation products, isothiocyanates, nitriles, and thiocyanates, induce stomatal closure accompanied by peroxidase-mediated reactive oxygen species production in *Arabidopsis thaliana*. *Biosci. Biotechnol. Biochem.* **77**, 977–983 (2013).
33. Mitchell-Olds, T. & Schmitt, J. Genetic mechanisms and evolutionary significance of natural variation in *Arabidopsis*. *Nature* **441**, 947–952 (2006).
34. Wang, B. et al. Ancient polymorphisms contribute to genome-wide variation by long-term balancing selection and divergent sorting in *Boechera stricta*. *Genome Biol.* **20**, 126 (2019).
35. Bloom, T. C., Baskin, J. M. & Baskin, C. C. Ecological life history of the facultative woodland biennial *Arabis laevigata* variety *laevigata* (Brassicaceae): seed dispersal. *J. Torrey Bot. Soc.* **129**, 21–28 (2002).
36. Song, B.-H. et al. Multilocus patterns of nucleotide diversity, population structure, and linkage disequilibrium in *Boechera stricta*, a wild relative of *Arabidopsis*. *Genetics* **181**, 1021–1033 (2009).
37. Mackay, T., Stone, E. & Ayroles, J. The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* **10**, 565–577 (2009).
38. Hedrick, P. W. Genetic polymorphism in heterogeneous environments: a decade later. *Annu. Rev. Ecol. Syst.* **17**, 535–566 (1986).
39. Hedrick, P. W. Antagonistic pleiotropy and genetic polymorphism: a perspective. *Heredity* **82**, 126–133 (1999).
40. Turelli, M. & Barton, N. H. Polygenic variation maintained by balancing selection: pleiotropy, sex-dependent allelic effects and G×E interactions. *Genetics* **166**, 1053–1079 (2004).
41. Gillespie, J. H. & Langley, C. H. A general model to account for enzyme variation in natural populations. *Genetics* **76**, 837–848 (1974).
42. Anderson, J. T., Willis, J. H. & Mitchell-Olds, T. Evolutionary genetics of plant adaptation. *Trends Genet.* **27**, 258–266 (2011).
43. Anderson, J. T., Lee, C.-R., Rushworth, C. A., Colautti, R. I. & Mitchell-Olds, T. Genetic trade-offs and conditional neutrality contribute to local adaptation. *Mol. Ecol.* **22**, 699–708 (2013).
44. Oakley, C. G., Ågren, J., Atchison, R. A. & Schemske, D. W. QTL mapping of freezing tolerance: links to fitness and adaptive trade-offs. *Mol. Ecol.* **23**, 4304–4315 (2014).
45. Price, N. et al. Combining population genomics and fitness QTLs to identify the genetics of local adaptation in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **115**, 5028–5033 (2018).
46. Kettunen, J. et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat. Genet.* **44**, 269–276 (2012).
47. Abuelsoud, W., Hirschmann, F. & Papenbrock, J. in *Drought Stress in Plants* Vol. 1 (eds Hossain, M. A. et al.) 227–248 (Springer, 2016).
48. Nguyen, D., Rieu, I., Mariani, C. & van Dam, N. M. How plants handle multiple stresses: hormonal interactions underlying responses to abiotic stress and insect herbivory. *Plant Mol. Biol.* **91**, 727–740 (2016).
49. Shani, E. M. et al. Plant stress tolerance requires auxin-sensitive Aux/IAA transcriptional repressors. *Curr. Biol.* **27**, 437–444 (2017).
50. Hopkins, R. J., van Dam, N. M. & van Loon, J. J. A. Role of glucosinolates in insect–plant relationships and multitrophic interactions. *Annu. Rev. Entomol.* **54**, 57–83 (2009).
51. Burow, M., Müller, R., Gershenzon, J. & Wittstock, U. Altered glucosinolate hydrolysis in genetically engineered *Arabidopsis thaliana* and its influence on the larval development of *Spodoptera littoralis*. *J. Chem. Ecol.* **32**, 2333–2349 (2006).
52. Wagner, M. R. & Mitchell-Olds, T. Plasticity of plant defense and its evolutionary implications in wild populations of *Boechera stricta*. *Evolution* **72**, 1034–1049 (2018).
53. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
54. Pagès, H., Aboyoun, P., Gentleman, R. & DebRoy, S. Biostrings: Efficient manipulation of biological strings. R package version 2.56.0 (2020).
55. Wang et al. Correction to: Ancient polymorphisms contribute to genome-wide variation by long-term balancing selection and divergent sorting in *Boechera stricta*. *Genome Biol.* **20**, 16 (2019).
56. Carley, L. et al. Data to accompany: Ecological factors influence balancing selection on leaf chemical profiles of a wildflower. *Dryad Data* https://doi.org/10.5061/dryad.7h44j0zsr (2021).

57. Atkinson, N. J., Lilley, C. J. & Urwin, P. E. Identification of genes involved in the response of *Arabidopsis* to simultaneous biotic and abiotic stresses. *Plant Physiol.* **162**, 2028–2041 (2013).
58. Sharma, A. et al. Comprehensive analysis of plant rapid alkalization factor (RALF) genes. *Plant Physiol. Biochem.* **106**, 82–90 (2016).
59. Dutilleul, C., Jourdain, A., Bourguignon, J. & Hugouvieux, V. The *Arabidopsis* putative selenium-binding protein family: expression study and characterization of *SBP1* as a potential new player in cadmium detoxification processes. *Plant Physiol.* **147**, 239–251 (2008).
60. Jiang, S.-C. et al. Crucial roles of the pentatricopeptide repeat protein SOAR1 in *Arabidopsis* response to drought, salt and cold stresses. *Plant Mol. Biol.* **88**, 369–385 (2015).
61. Wen, J., Vanek-Krebitz, M., Hoffmann-Sommergruber, K., Scheiner, O. & Breitender, H. The potential of *Betv1* homologues, a nuclear multigene family, as phylogenetic markers in flowering plants. *Mol. Phylogenet. Evol.* **8**, 317–333 (1997).
62. Koo, A. J., Fulda, M., Browse, J. & Ohlrogge, J. B. Identification of a plastid acyl-acyl carrier protein synthetase in *Arabidopsis* and its role in the activation and elongation of exogenous fatty acids. *Plant J.* **44**, 620–632 (2005).
63. Henrissat, B. et al. Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc. Natl Acad. Sci. USA* **92**, 7090–7094 (1995).

Acknowledgements

We thank the Duke University greenhouse staff, E. Hornstein, S. Mahanes, L. Schumm, R. Bingham, N. Niezgoda, C. Ried, A. Simha, W. de Vries, E. Cousins and K. Stinson for assistance with fieldwork in Colorado; E. Raskin, C. Strock, B. Guyton, W. Mitchell, J. Lessing, M. Olszack, A. Zemenick, M. McMunn, K. Stiff, S. Clemens, T. Park, S. Shriber-Olds and R. Colautti for assistance with fieldwork in Idaho; and J. Reithel, S. Sprott and C. Heald for logistical support that facilitated fieldwork. We thank the Rocky Mountain Biological Laboratory, the Crested Butte Land Trust, the US Forest Service, D. Finlayson, R. Capps, A. Mears and P. Lehr for permission to conduct field experiments. We thank V. Grant and A. Zhao for assistance with laboratory experiments, E. Iversen at the Duke University Statistical Consulting Center for guidance regarding some analyses, and K. Donohue, M. Rausher and W. Morris for feedback that greatly improved this manuscript. We thank the Computer and Information Networking Center at National Taiwan University for high-performance computing facilities and the Technology Commons in the College of Life Science at National Taiwan University for molecular biology equipment. This work was supported by the Ministry of Science and Technology of Taiwan (grant no. 108-2636-B-002-004 to C.-R.L.), the Guangdong Natural Science Funds for Distinguished Young Scholar (grant no. 2018B030306040 to B.W.), the National Institutes of Health (grant no. R01 GM086496 to T.M.-O.) and the Rocky Mountain Biological Laboratory (Snyder Endowment graduate fellowship to L.N.C.).

Author contributions

L.N.C., J.P.M., C.-R.L., J.W., C.L.N. and T.M.-O. designed the project. L.N.C., J.P.M., C.-Y.C., K.V.S.K.P., E.C., R.K., C.L.N., C.F.O.-M., C.A.R., M.R.W., J.W., P.-M.Y., K.G., C.-R.L. and T.M.-O. collected the data. L.N.C., J.P.M., B.W., C.-Y.C., Y.-P.L., M.R., J.G., C.-W.H., C.-R.L. and T.M.-O. analysed the data. L.N.C., C.-R.L. and T.M.-O. wrote the paper. All authors read and approved the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41559-021-01486-0.

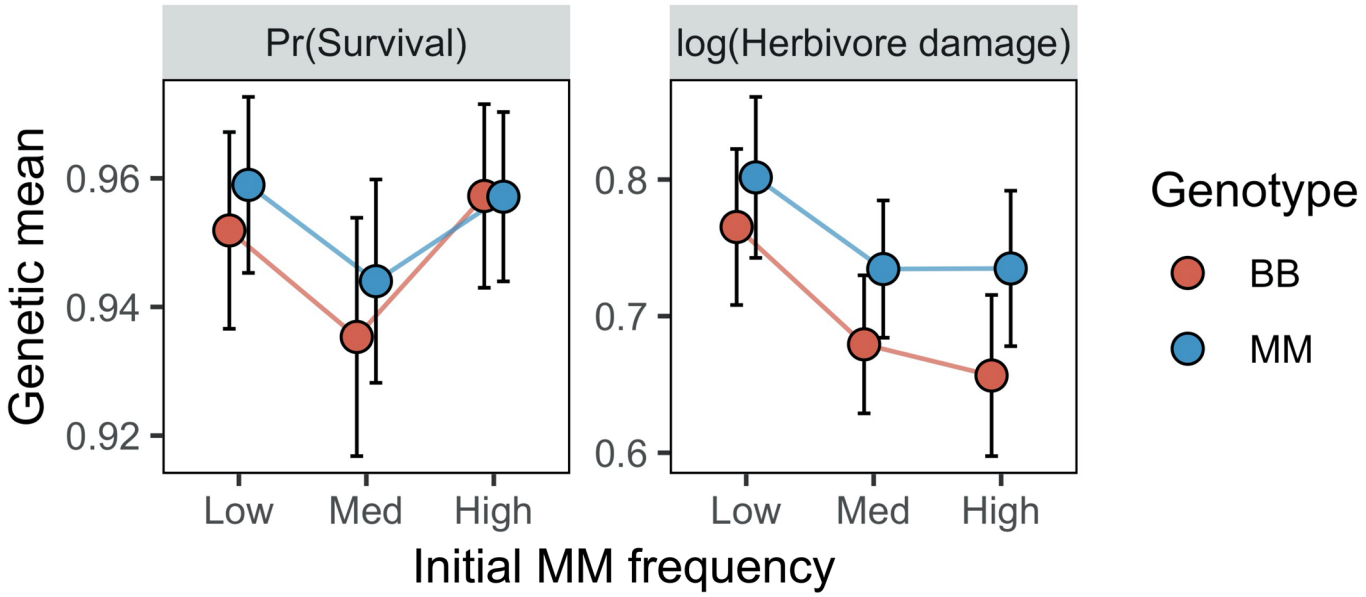
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41559-021-01486-0.

Correspondence and requests for materials should be addressed to C.-R.L. or T.M.-O.

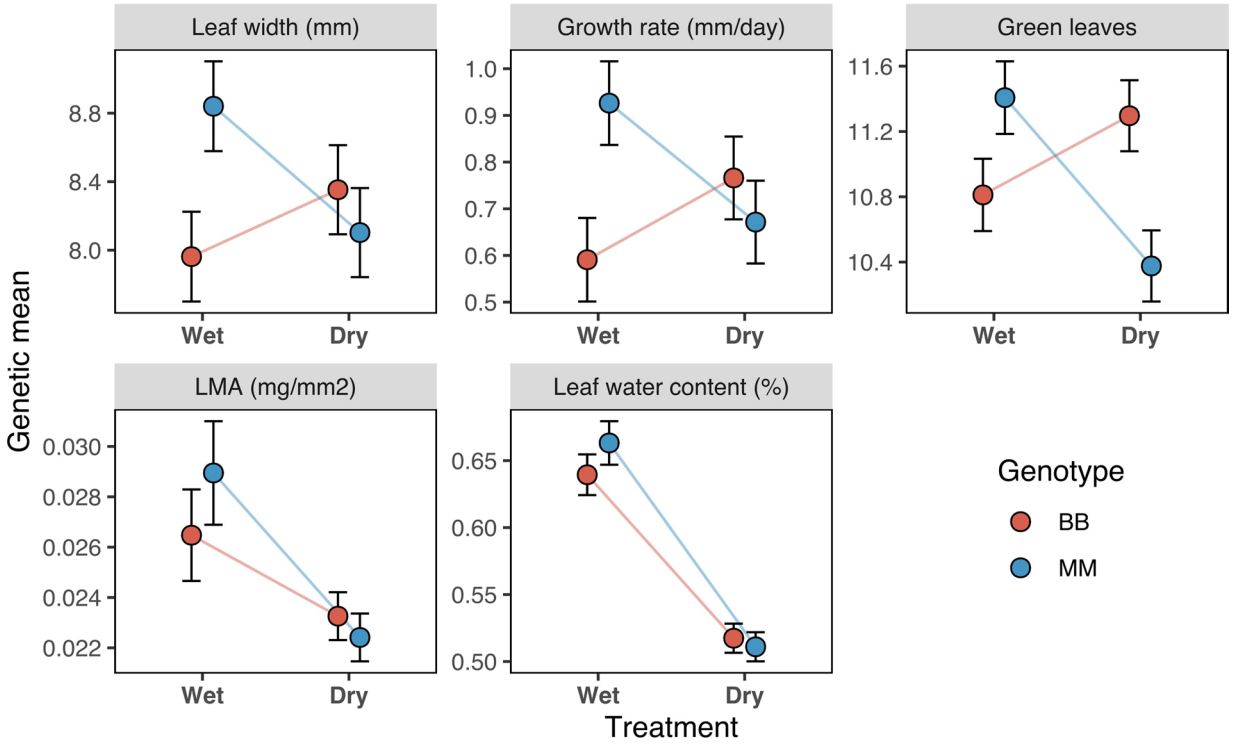
Peer review information *Nature Ecology & Evolution* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

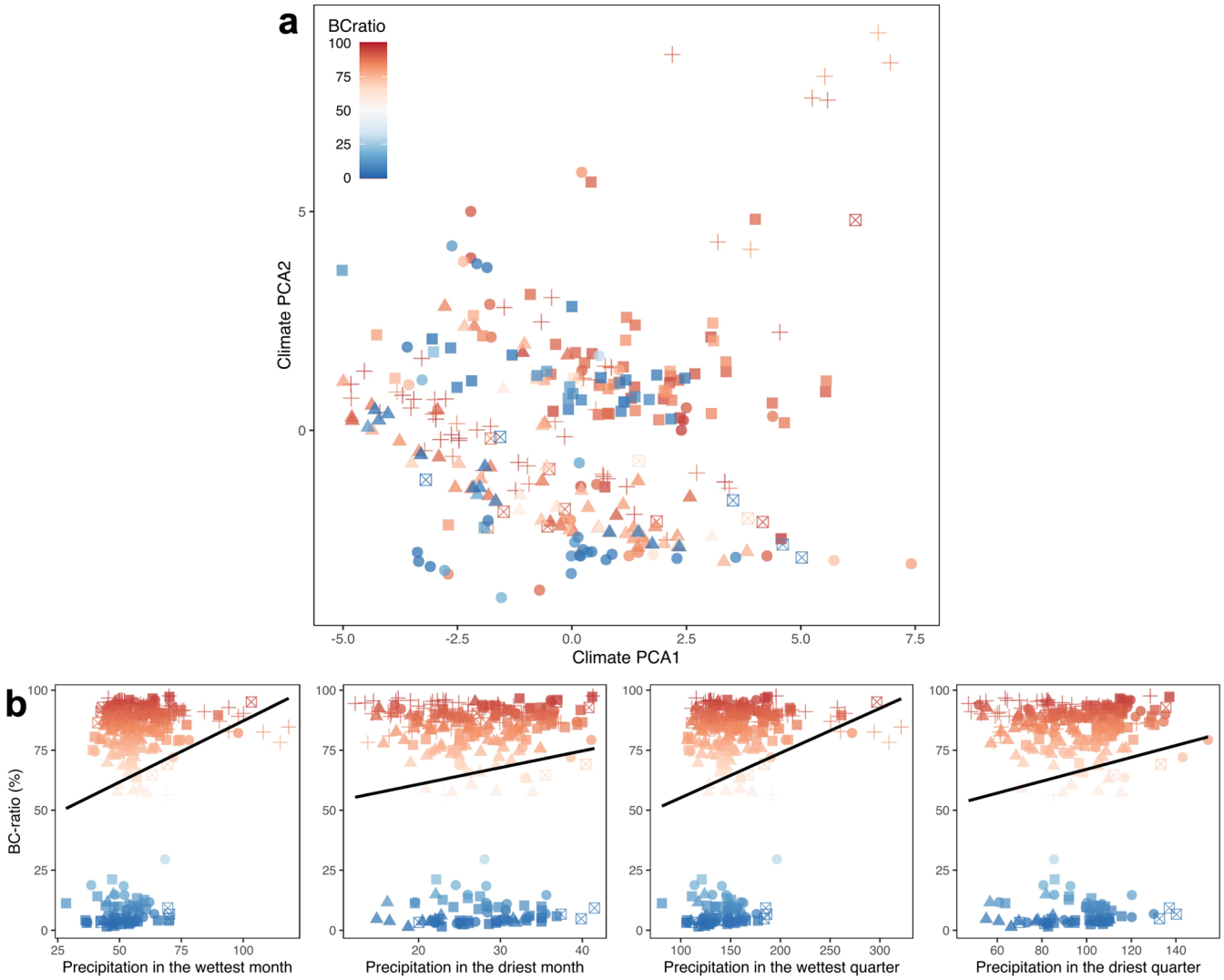
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



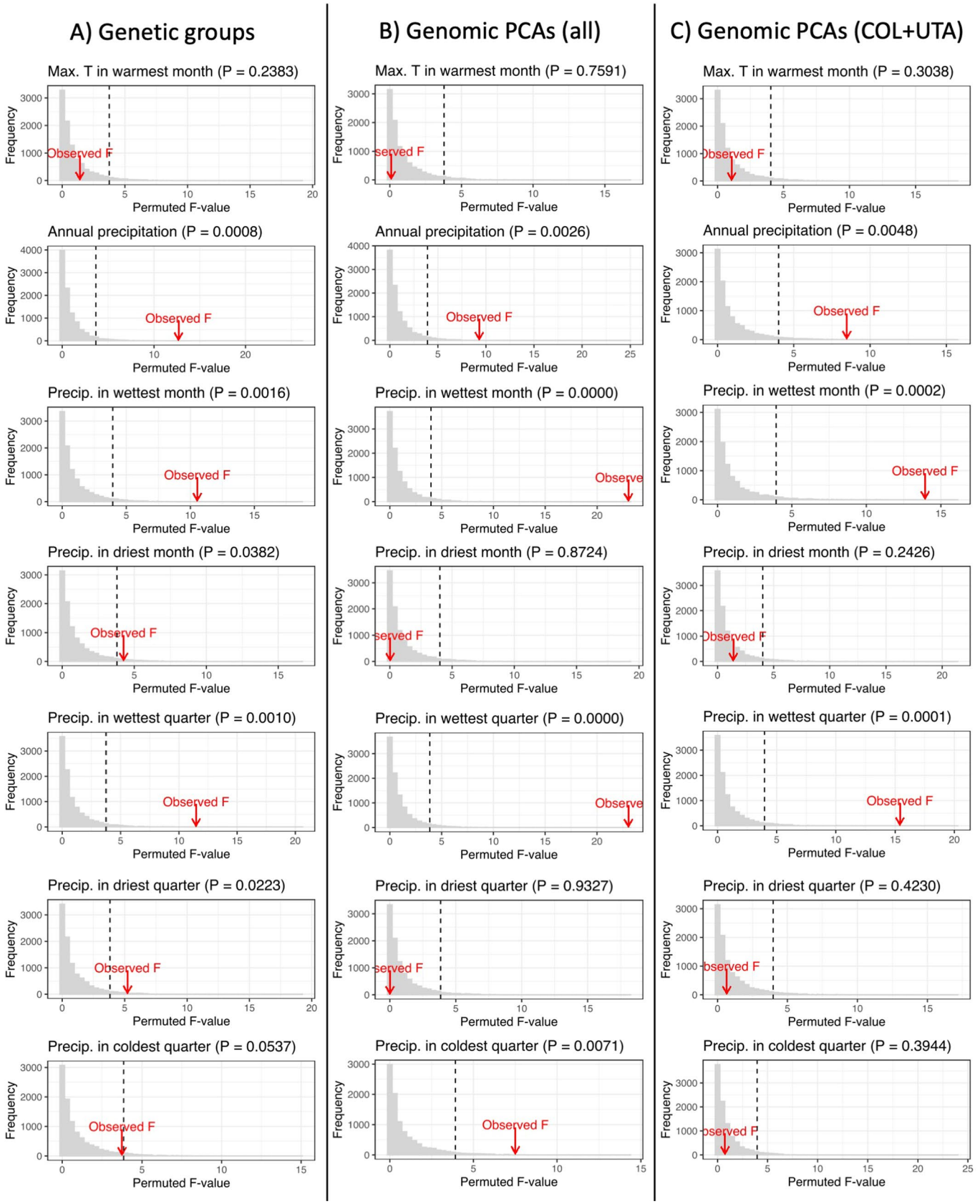
Extended Data Fig. 1 | Genotype frequency does not alter the effect of *BCMA1/3* on herbivore resistance or survival. In experimental arrays in which we manipulated the starting genotype frequency of the *BCMA1/3* homozygotes, there was no effect of genotype frequency on herbivore damage or survival. In each panel, points represent least-squares means estimates of the response variable for each genotype in each *BCMA1/3* frequency treatment, and error bars represent ± one standard error.



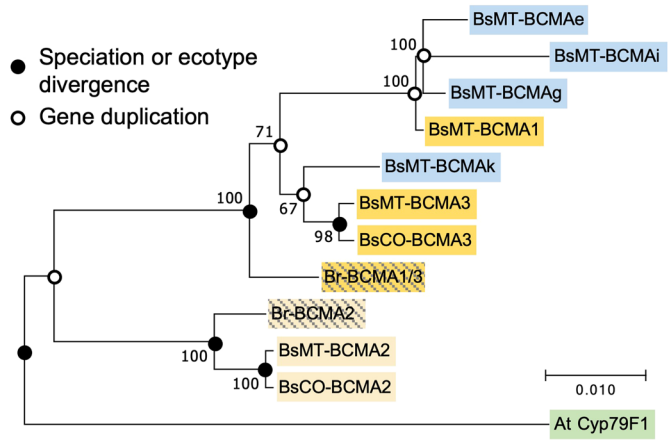
Extended Data Fig. 2 | Genetic variation in norms of reaction to drought stress. *BCMA1/3* alleles in the CFR-NIL background confer contrasting response to drought by altering morphological traits such as leaf size and number as well as physiological traits such as growth. Both genotypes reduce leaf water content under drought, but genetic differences in this response were only marginally significant. In each panel, points represent least-squares means estimates of the response variable for each genotype in each *BCMA1/3* frequency treatment, and error bars represent ± one standard error.



Extended Data Fig. 3 | Drought-related climate variables are correlated with multivariate climatic predictors of BC-ratio. a BC-ratio varies across climate space, with PC1 the strongest predictor (Supplementary Table 9). **b** Linear models and permutation tests reveal that low BC-ratio phenotypes are significantly correlated with drier environments of origin. Points in all panels represent phenotypic (LS mean BC-ratio) and environmental variation (WorldClim data from the location of origin) of a broad panel of accessions. Colors represent BC-ratio, ranging from 0% (blue) to 100% (red). Shapes denote genetic groups as described in Wang et al. (2019). COL: circles; NOR: triangles; UTA: squares; WES: crosses. Black lines represent lines of best fit estimated using linear models using discrete groups to control for population structure ('approach A') as described in Supplementary Methods.

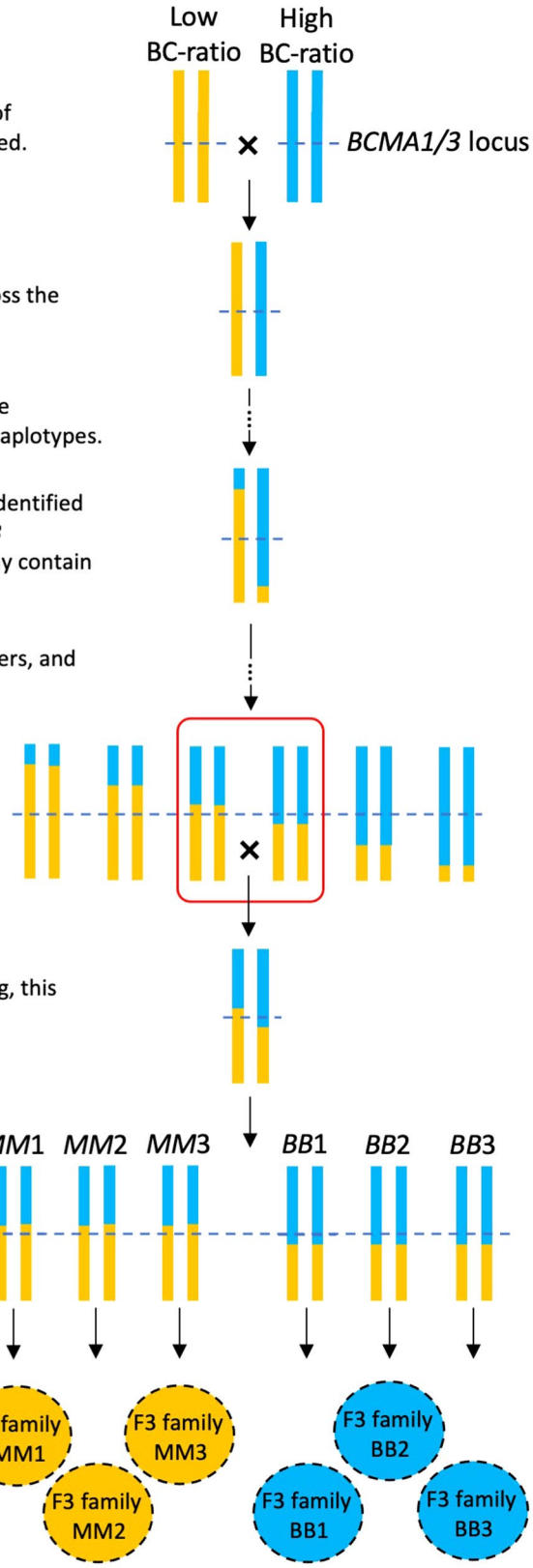


Extended Data Fig. 4 | Permuted vs. observed *F*-statistics relating BC-ratio to climate variables. Panes correspond to linear models presented in Supplementary Table 11. In each pane, gray bars show the frequency distribution of the test statistic relating each climate variable to BC-ratio from 10,000 permutations shuffling BC-ratio values without replacement (Supplementary Methods), red arrows show the observed *F*-statistic from each true model (Supplementary Table 11), and dashed lines mark the location of the extreme 95% tail in the empirical cumulative distribution function of permuted *F*-statistics, using three different methods to control for population structure (columns A-C; Supplementary Methods).



Extended Data Fig. 5 | Functional and copy number variation in *BCMA* evolved recently within *B. stricta*. Maximum likelihood phylogenetic reconstruction of *BCMA* copy sequences (excluding severely truncated copies) elucidates the evolutionary history of *BCMA* duplications in *Boechera*. Colored boxes behind (pseudo)gene names categorize features as follows: blue boxes contain nonfunctional *BCMA* pseudogenes on chromosome 7, light yellow boxes contain functional copies of *BCMA2* on chromosome 2, and dark yellow boxes contain functional copies of *BCMA3* and *BCMA1* on chromosome 7. Shaded boxes indicate paralogs in *B. retrofracta*, and green box indicates the *A. thaliana* ortholog (*CYP79F1*). Scale bar shows genetic distance in nucleotide differences per base pair.

BCMA1/3 CFR-NILs: Chromosome 7 Pedigree



In the **parental generation,** wild accessions homozygous for production of methionine and branched-chain glucosinolates were identified and crossed.

An **F1** heterozygous for parental alleles at *BCMA1/3* (and elsewhere across the genome) was generated.

F2, **F3**, and **F4** progeny were produced by single-seed descent to increase homozygosity across the genome and allow recombination of parental haplotypes.

In the **F4** generation, an individual heterozygous around *BCMA1/3* was identified using PCR markers. Due to previous generations of selfing, this *BCMA1/3* heterozygote was largely homozygous elsewhere in the genome, but may contain other segregating loci.

F5s that had recombined near *BCMA1/3* were identified using PCR markers, and selfed to generate **F6s** by single seed descent.

Homozygous **F6** recombinants were screened using 13 markers closely linked to *BCMA1/3* to identify closest flanking recombinants (red box).

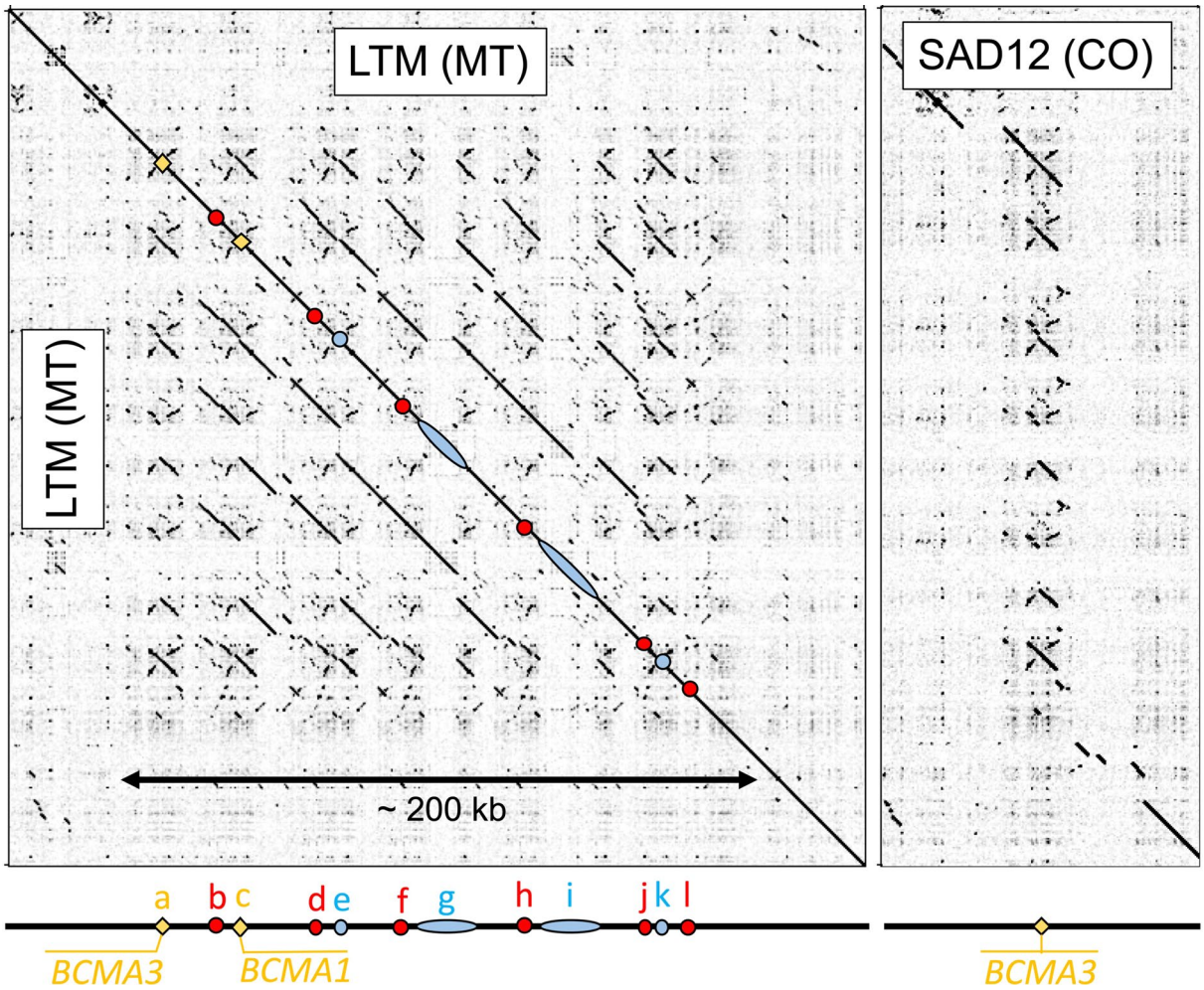
Crossing the F6 CFRs yielded an **F1 CFR heterozygote**. Due to prior selfing, this was largely homozygous elsewhere in the genome, but some unlinked polymorphism is expected.

Selfing a single F1 CFR heterozygote yielded multiple independent **F2 CFRs homozygous for each *BCMA1/3* haplotype.** Some rare unlinked polymorphism is still expected elsewhere in the genome.

Selfing F2 CFR homozygotes gave independent **homozygous F3 CFR families** representing methionine-derived *(MM)* and branched-chain amino acid-derived *(BB)* genotypes differing at *BCMA1/3*. Replicated family lines within each genotype control for unlinked polymorphisms. These F3 families were used in laboratory and field experiments, and in sequencing to determine the extent of the non-recombinant interval around *BCMA1/3*.

Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | *BCMA1/3* CFR-NILs: Chromosome 7 Pedigree. Chromosomal pedigree showing how closest flanking recombinant near-isogenic lines (CFR-NILs) were generated for use in laboratory and field experiments. See Methods and Supplementary Information for details. Within each step, diploid homologous pairs of Chromosome 7 are shown.



Extended Data Fig. 7 | Long-read assemblies of the LTM and SAD12 parents reveal substantial variation in tandem repeats and *BCMA* copy number in a 200 kb region on chromosome 7. Functional *BCMA* gene copies are indicated in yellow; red circles show severely truncated, non-functional *BCMA* copies; blue ellipses indicate close-to-full-length copies of *BCMA* containing frameshift deletions or transposon insertions. Blue and yellow elements match those shown in Extended Data Figures 5 and Supplementary Figure 2.

nature research

Corresponding author(s): Thomas Mitchell-Olds
Cheng-Ruei Lee

Last updated by author(s): 2021/04/14

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used in data collection.
Data analysis	All analyses of plant traits (herbivore resistance, survival, reproduction, drought respose, etc.) were conducted using JMP Pro v. 12 or higher and/or R v. 3.6.0 or higher. R code used for analyses is archived along with the supporting data in the Dryad Digital Repository (DOI: 10.5061/dryad.7h44j0zsr). Custom Python and R scripts were also used in analyzing genomic data; those that have not been previously published are archived along with the supporting data in the Dryad Digital Repository (DOI: 10.5061/dryad.7h44j0zsr).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

New reference genome assemblies and raw Nanopore reads for SAD12 and LTM genotypes have been submitted to NCBI (BioProject number PRJNA609209). The short reads of the GBS data for CFR-NIL families have been submitted to NCBI (Bioproject number PRJNA659863). Previously published genomic data are archived with Wang et al. 2019 (DOI: 10.1186/s13059-019-1729-9). Code and all other data reported in this manuscript are archived in the Dryad Digital Repository (DOI: 10.5061/dryad.7h44j0zsr).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	The studies reported here include: analytical chemistry, field transplant experiments, field array experiments, a lab herbivory assay, two drought tolerance dry-down experiments, Sanger sequencing analysis, and analyses of genomic data to characterize linkage disequilibrium, flanking genes, and FST. Full details regarding the study design for each of these components are provided in the Methods, Results, and Supplementary Information of the manuscript.
Research sample	This study focuses on wild-derived accessions and greenhouse crosses of the wildflower Boechera stricta (Brassicaceae). The wild-derived accessions are meant to represent a broad range of populations across the species range. To minimize maternal environmental effects, wild-derived accessions were self-pollinated for a minimum of 1-2 generations under greenhouse conditions prior to experimental use. Genotypes generated by crossing were created to target a gene of interest (BCMA3). The parental accessions used in the cross were chosen from the broad panel of naturally-derived accessions because they showed strong differences in foliar chemistry, the focal trait in this study.
Sampling strategy	Seeds from wild plants in the field were collected to deliberately span a broad geographic area. Because self-pollination is common in our study species, seeds were not collected from wild parent plants within ~1km of other existing collections. For experimental work, the sample sizes of experiments were determined primarily by logistical feasibility, maximizing replication within the means of the researchers managing each experiment.
Data collection	Leaf chemical phenotypes were determined using high-performance liquid chromatography. These data were collected at the Department of Biochemistry in the Max Planck Institute for Chemical Ecology. All other plant phenotypic data was collected in the laboratory and field by the authors, using standard measuring methods (rulers, balances). Researchers phenotyping plants were blind to the genotype of the plants, as the plants were tagged using a unique ID number rather than a descriptive label.
Timing and spatial scale	Field data were collected during the short growing season in high-elevation Rocky Mountain sites (June-September), which is constrained by snowfall on both ends, from 2013-2017. We focus on late-season census data from the field because herbivore damage remains visible on leaves throughout the season, so later measurements reflect cumulative damage across the growing season. Late-season censuses also allow us to measure survival through the growing season, and reproduction among plants that survived and had experienced vernalization. The laboratory herbivory assay was conducted in 2016. Data were collected at the end of the herbivory treatment. Chemical phenotyping of the broad panel of accessions was performed in 2016. Genomic data on the broad panel of accessions and on the experimental crosses was generated incrementally from 2015-2019, as additional DNA extractions and library preparations could be performed depending on the availability of personnel. The dry-down experiment using the broad panel of accessions was conducted in 2015. The dry-down experiment using the experimental crosses was conducted in 2017. In both of these experiments, data were collected from the beginning of the drought treatment to the end of the drought treatment. The plants were destructively harvested, so no further data collection was possible.
Data exclusions	In our field array experiment, one array experienced extremely high mortality. Analyses of the effect of mortality rate on genotype frequencies are presented both including and excluding this outlier. This is described in Supplementary Information, and results from both approaches are included in Supplementary Table 5. In the dry-down experiment with CFR-NILs, three individual plants were extreme outliers for estimated leaf mass per area (LMA), possibly reflecting measurement error. We excluded these three outliers from our analysis of the effect of genotype and drought treatment on LMA. This exclusion is retained and annotated in the analysis code archived with this manuscript (DOI: 10.5061/dryad.7h44j0zsr). Following genotyping-by-sequencing, two samples with few or low quality reads were removed. For other analyses, we included all available data. In multivariate models, missing data from one or more variables precluded the inclusion of some individual plants/genotypes.
Reproducibility	Because we anticipated the effects of our focal gene to vary depending on ecological context, we reproduced our field herbivory and survival experiments in 15 environments. All laboratory experiments were conducted once unless otherwise noted.
Randomization	All field and laboratory experiments were conducted using either randomized complete block designs or completely randomized designs. Occasionally, completely randomized designs were used instead of randomized complete blocks if the number of replicates

	across genotypes and treatments was not evenly divisible by the functional blocking unit (e.g. greenhouse rack that can hold a fixed number of pots). Plant positions within blocks were determined using a random number generator.
Blinding	In all experiments with human-collected phenotype data, plants were labeled using a unique plant ID representing its position in the randomization scheme. While the accessions and genotypes used in this study differ in foliar chemistry, they are not distinguishable from one another with the naked eye.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	Fieldwork was conducted at 11 sites in the Rocky Mountains, USA. The field sites were grouped in two focal regions, central Idaho and west-central Colorado. Sites in the northern and southern Rocky Mountains ranged from 1812-2531 and 2888-3145 m in elevation, respectively. Accordingly, they differed in mean temperature and precipitation.
Location	The latitude, longitude, and elevation of each wild accession is included in the data archived along with this manuscript (DOI: 10.5061/dryad.7h44j0zsr) in the file "BCMA-final-RefPops-AllMerged.csv". The latitude, longitude, and elevation of all field sites where experiments were conducted is included in the data archived along with this manuscript (DOI: 10.5061/dryad.7h44j0zsr) in the file "site_locations.csv". All greenhouse and laboratory experiments were conducted at Duke University in Durham, NC, USA.
Access & import/export	Permits for fieldwork were obtained from United States Forest Service Regions 1, 2, 4, 5, and 6, and Grand Canyon, Sequoia, and Kings Canyon National Parks between 1999 and 2019.
Disturbance	Our field experiments displaced a small amount of soil for each transplanted individual or array. At the conclusion of all experiments, transplants were destroyed and any removed soil was replaced. For fenced transplant gardens, there was also disruption to soil and vegetation at the time of fence construction. Fences were removed and soil replaced at all sites that are no longer being actively used for research.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	One laboratory experiment utilized live, second-instar insect herbivores (Trichoplusia ni [Lepidoptera: Noctuidae]) purchased from a commercial supplier under USDA APHIS permit number P526P-15-00202.
Wild animals	This study did not involve wild animals.
Field-collected samples	Field-collected accessions were grown under greenhouse conditions (18-21°C day/13-16°C night; 16h day length; watering to saturation daily; fertilization at 300 ppm N weekly; unless otherwise noted in the Methods, e.g. during drought experiments). We clipped the fruits off of experimental transplants before ripening to prevent the dispersal of non-local seed into nearby populations. At the end of field experiments, remaining plants were uprooted, infrastructure (e.g. plant tags, fences and cages to prevent browsing) was removed, and any displaced soil was replaced.
Ethics oversight	All fieldwork conducted in Colorado was reviewed and approved by the Science Committee and Science Director at the Rocky Mountain Biological Laboratory, and included consideration of environmental impacts. Garden and array sites in Colorado were permitted by the United States Forest Service under special use permit GUN 1120 through coordination with the Rocky Mountain

Biological Laboratory, or by private property owners. Garden sites in the northern Rocky Mountains were permitted by the United States Forest Service or by private property owners.

Note that full information on the approval of the study protocol must also be provided in the manuscript.