Received: 5 January 2023 | Accepted: 1 August 2023
DOI: 10.1111/pce.14691

DATA INSIGHT

Plant, Cell & Environment WILEY

Combined genomics to discover genes associated with tolerance to soil carbonate

Silvia Busoms[1] | Laura Pérez-Martín[1] | Joana Terés[1] | Xin-Yuan Huang[2] | Levi Yant[3] | Roser Tolrà[1] | David E. Salt[4] | Charlotte Poschenrieder[1]

[1]Department of Animal Biology, Plant Biology, and Ecology, Plant Physiology Laboratory, Bioscience Faculty, Universitat Autònoma de Barcelona, Barcelona, Spain

[2]State Key Laboratory of Crop Genetics and Germplasm Enhancement, College of Resources and Environmental Sciences, Nanjing Agricultural University, Nanjing, China

[3]Future Food Beacon of Excellence & School of Life Sciences, University of Nottingham, Nottingham, UK

[4]Future Food Beacon of Excellence & School of Biosciences, University of Nottingham, Sutton, UK

Correspondence
Charlotte Poschenrieder
Email: Charlotte.Poschenrieder@uab.es

Funding information
SPANISH MINISTRY OF SCIENCE AND INNOVATION, Grant/Award Number: PID2019-1040000RB-I00

Abstract

Carbonate-rich soils limit plant performance and crop production. Previously, local adaptation to carbonated soils was detected in wild *Arabidopsis thaliana* accessions, allowing the selection of two demes with contrasting phenotypes: A1 (carbonate tolerant, c+) and T6 (carbonate sensitive, c−). Here, $A1_{(c+)}$ and $T6_{(c-)}$ seedlings were grown hydroponically under control (pH 5.9) and bicarbonate conditions (10 mM $NaHCO_3$, pH 8.3) to obtain ionomic profiles and conduct transcriptomic analysis. In parallel, $A1_{(c+)}$ and $T6_{(c-)}$ parental lines and their progeny were cultivated on carbonated soil to evaluate fitness and segregation patterns. To understand the genetic architecture beyond the contrasted phenotypes, a bulk segregant analysis sequencing (BSA-Seq) was performed. Transcriptomics revealed 208 root and 2503 leaf differentially expressed genes in $A1_{(c+)}$ versus $T6_{(c-)}$ comparison under bicarbonate stress, mainly involved in iron, nitrogen and carbon metabolism, hormones and glycosylates biosynthesis. Based on $A1_{(c+)}$ and $T6_{(c-)}$ genome contrasts and BSA-Seq analysis, 69 genes were associated with carbonate tolerance. Comparative analysis of genomics and transcriptomics discovered a final set of 18 genes involved in bicarbonate stress responses that may have relevant roles in soil carbonate tolerance.

KEYWORDS
Arabidopsis, bicarbonate stress, BSA-Seq: transcriptomics, calcareous soil

1 | INTRODUCTION

Soil alkalinity is a highly stressful environmental factor that limits plant growth and crop yield (Filippi et al., 2019; Li et al., 2018). High pH soils are present in 30% of the Earth surface specially located in areas with arid and semiarid climate. The pH values of most calcareous soils are within the range of 7.5–8.5 (Loeppert & Suarez, 1996) with bicarbonate concentrations between 5 and 35 mmol L^{-1}. The main anions present in excess in alkaline soils are HCO_3^- and CO_3^{2-} (Al-Busaidi & Cookson, 2003). Soil carbonates act as pH buffers and play an important role in rhizosphere processes, hampering nutrient availability to plants. The low availability of nitrogen (N), phosphorus (P) and micronutrients such as iron (Fe), zinc (Zn) and manganese (Mn) produce nutrient deficiencies in many plant species cultivated on carbonated soils (Dreyer et al., 2020; Hui et al., 2019; Terés et al., 2019). Moreover, the high pH surrounding plant roots can alter the membrane potential (Felle, 2001) and inhibit the uptake of essential ions, and thus further contribute to nutrient deficiency in sensitive plants.

The uptake of bicarbonate can cause important metabolic disturbances due to both alteration of cell pH homoeostasis and dark fixation of inorganic carbon. Bicarbonate can be quickly incorporated into organic acids, mainly malate, causing inhibition of mineral nutrient transport from roots to shoot, reduction of root growth, and oxidative stress in sensitive calcifuges, but not in tolerant calcicole species (Lee & Woolhouse, 1969; Sagervanshi et al., 2021). Due to these multiple direct and indirect injuries, sensitive species exhibit a complex syndrome of stress symptoms, including morpho-anatomical changes in roots, disturbance of water relations, signs of iron and/or zinc deficiency, reduction of total photosynthetic pigments and photosynthetic activity; accumulation of osmo-protectants, soluble sugars, and organic acids; and activation of the biosynthesis of antioxidant enzymes (Ahmad et al., 2014; Sarkar et al., 2022; Zou et al., 2021).

A better understanding of plant alkaline tolerance mechanisms and cultivation of new varieties of alkali-tolerant crops is needed to improve carbonated soils and increase food production (Cao et al., 2022). However, current knowledge of the bicarbonate stress response of plants is limited. Most studies have been conducted on roots of crop species such as *Glycine max* or *Oryza sativa* which are species with moderate bicarbonate tolerance (Waters et al., 2018; Zhang et al., 2017). Nonetheless, after several decades of effort still a proper model is missing to understand the adaptive mechanism to carbonate stress and the tolerance mechanism at the molecular-genetic level. Early experiments revealed that organic acid accumulation in response to bicarbonate occurred both in sensitive and tolerant species and it was speculated that differential compartmentation of organic acids may play a role in the tolerance (Sagervanshi et al., 2021). In fact, several anion transporters, and gene families such as ALMT, NRT/POT and SLAHs were related to bicarbonate response in *G. max* (Duan et al., 2018; Waters et al., 2018).

During recent years, combined omics are applied to successfully characterize plant tolerance responses to complex stress factors (e.g., Mehari et al., 2021). Resequencing-based bulk segregation analysis (BSA) utilizes a strategy of pooling individuals with extreme phenotypes to conduct rapidly linked marker screening (Li & Xu, 2022). Whole-genome studies are useful to detect signature of selection, quantitative trait locus, and single nucleotide polymorphism (SNP) markers. However, some important changes affecting the regulation of key genes can remain hidden. The combination of genomic and transcriptomic studies allows a more accurate screening of the candidate genes involved in specific physiological processes. Here, we took advantage of the natural variation present in two *Arabidopsis thaliana* populations with contrasting phenotypes of soil alkalinity tolerance (Terés et al., 2019) to highlight the loci involved in bicarbonate stress responses and adaptation to carbonated soils by combining BSA-Seq and RNA-Seq technologies.

2 | MATERIALS AND METHODS

2.1 | Plant material

In previous studies, natural variation of *A. thaliana* populations from Catalonia (Busoms et al., 2015) were tested in a multiyear small-scale common garden under carbonated and noncarbonated soils (Terés et al., 2019). Seeds of six extreme phenotype plants from $A1_{(c+)}$, a soil carbonate tolerant deme, and six extreme phenotype plants from $T6_{(c-)}$, a soil carbonate sensitive deme, were collected from the last reciprocal transplant experiment performed in 2015 and stored under fresh (4°C) and dry conditions until the beginning of hydroponic and soil cultivation experiments. A1 × T6 and T6 × A1 crosses were generated using the six $A1_{(c+)}$ and six $T6_{(c-)}$ lines (generating 12 crossing lines) and self-crossed for three generations (F1, F2, F3). Col-0 seeds and seeds of 18 T-DNA mutants with Col-0 background were purchased from NASC/ABRC (Scholl et al., 2000). Segregating lines were genotyped using specific primers and multiplied for T2 generations. The nonexpression of the genes was validated through qPCR (Supporting Information: Data set SD1).

2.2 | Growth conditions

Selected seeds were surface sterilized by soaking in 70% (v/v) ethanol for 1 min, suspended in 30% (v/v) commercial Clorox bleach and 1 drop of Tween-20 for 5 min and rinsed five times in sterile 18 MΩmilli-Q water. Seeds were kept in water, at 4°C, in the dark conditions for 48 h to synchronize germination.

Hydroponic culture: *A. thaliana* seeds from six $A1_{(c+)}$ and six $T6_{(c-)}$ lines were sown in 0.2 mL tubes containing 0.6% phyto agar (Duchefa) prepared in nutrient solution 1/2 Hoagland solution (HS), pH 5.9 in a growth chamber (12 light/12 dark hours, 150 µmol $cm^{-2}·s^{-1}$, 40% humidity and 25°C). The bottom of the tubes containing seedlings was cut off and the tubes were placed in 150 mL hydroponic container with aerated nutrient solution 1/2 HS, pH 5.9. After 15 days postgermination, treatment was applied, and seedlings were separated in different sets. The treatments consisted of control (½ HS at pH 5.9), high pH (½ HS at pH 8.3) and bicarbonate (½ HS at pH 8.3 and 10 mM $NaHCO_3$). Solutions were buffered with different proportions of 2-(N-morpholino)ethanesulfonic acid and BTP (BIS-TRIS propane) depending on the final pH. Plants of $A1_{(c+)}$, $T6_{(c-)}$ used for transcriptomic analysis were collected after 48 h under treatment. Plants of $A1_{(c+)}$, $T6_{(c-)}$, Col-0 and the 18 T-DNA knockout mutants were collected after 10 days under treatment. Image analysis for root length measurements were performed with Image J, software16 (Schneider et al., 2012).

Greenhouse experiments: four cultivations were conducted consecutively under the same conditions: (1) six $A1_{(c+)}$, six $T6_{(c-)}$, six A1 × T6 and six T6 × A1 F1 lines; (2) five $A1_{(c+)}$, two $T6_{(c-)}$, 10 A1 × T6 and four T6 × A1 F2 families; (3) five $A1_{(c+)}$, two $T6_{(c-)}$, 66 A1 × T6 and 28 T6 × A1 F3 families; (4) Col-0 and the 18 T-DNA mutants of candidate genes. Plants were cultivated in carbonated soil (LP) mixed with perlite (3:1) (Table 1) in 6 × 6 cm square pots under

TABLE 1 Physical and chemical properties of the natural carbonated soil excavated from Les Planes d'Hostoles (LP, 42° 03′ 45.1″ N; 2° 32′ 46.6″ E).

	Geology Limestones		Texture Clay-loam		O.M. 47.3		CaCO3 33.25			pH 7.9	
%	Zn	Cu	Mo	P	S	Mg	Fe	Mn	Na	Ca	K
Mean (μg/g)	4.9	2.1	0.02	40.3	25.5	33.9	40.8	23.3	39.2	353.8	60
SD	1.9	0.7	0.01	10.5	7.2	9.6	10.7	7	6.1	56.2	20.1

Abbreviations: O.M., organic matter; SD, standard deviation.

semicontrolled conditions (12-h light/12-h dark photoperiod, temperature between 15°C and 25°C). Five seeds of each line were sown in pots and distributed randomly in the greenhouse. Two weeks after germination, seedlings were thinned out so that only one plant per pot was left. Irrigation with distilled water was applied twice a week at the bottom of the trays. For the third cultivation, two leaves per plant were collected 20 days after germination and stored at −80°C for subsequent DNA extraction and sequencing of the selected plants. Rosette diameter and bolting time were monitored, and silique number was counted at maturity in all the experiments.

2.3 | Ionomic analysis

Root ionome was assessed in plants submitted for 10 days to treatment conditions. Roots were carefully washed with 18 MΩ to remove nutrients from solution. Plant material was dried for 4 days at 60°C. Approximately 0.1 g was used to perform an open-air digestion in Pyrex tubes using 0.7 mL concentrated HNO_3 at 110°C for 5 h in a hot-block digestion system (SC154-54-Well Hot Block™, Environmental Express) placed inside a flow cabinet. The concentrations of selected elements (Ca, K, Mg, Na, P, S, B, Mo, Cu, Fe, Mn, Zn) were determined by inductively coupled plasma optical emission spectroscopy ICP-OES (Thermo Jarrell-Ash, model 61E Polyscan).

2.4 | Genomic analysis

Leaf material from two $A1_{(c+)}$ individuals (F0-tolerant pool), two $T6_{(c-)}$ individuals (F0-sensitive pool) and 19 F3 individuals from seven tolerant families (two T6 × A1 and five A1 × T6) (F3-tolerant pool) were used for DNA extraction (Supporting Information: Data set SD5). DNA was isolated with a Qiagen DNeasy kit and libraries were prepared using a TruSeq DNA PCR-Free Sample Preparation Kit from Illumina. Whole-genome sequencing was performed on Illumina HiSeq. 2000 by Novogene. Read depth was measured after processing and alignment to the TAIR10 reference (below) at 15× coverage for each individual sample and 180× for the pooled sample (19 *F3* individuals).

Raw sequence data was processed as follows: (1) read trimming by default parameters with Sickle (Joshi & Fass, 2011); (2) alignment to the TAIR10 reference with bwa mem 0.7.12 (processing with SAMtools 1.3); (3) removal of duplicate reads using Picard (MarkDuplicates); (4) application to read group name correction to the bam files using Picard (AddOrReplaceReadGroups); (5) realigning Indels using the GATK 'GenomeAnalysis' Toolkit (McKenna et al., 2010).

Bi-allelic SNPs were identified using 'HaplotypeCaller' and genotyped using 'GenotypeGVCF' (both tools in GATK). Data were filtered using GATK SelectVariants using these 'GATK best practices' guidelines: QD < 2.0||MQ < 40.00||FS > 60.0||SOR > 4.0 || MQRankSum ≤ 8.0 || ReadPosRankSum ≤ 8.0 and a minimum coverage of 10 per sample. Genotyping of the pooled sample (19 *A1* × *T6* individuals) was performed with LoFreq, (Wilm et al., 2012) as GATK failed with ploidy levels required (ploidy-38 for 19 diploid samples). LoFreq was run with default settings. VCFs were then merged with BCFtools (Li, 2011). The allele frequency (AF) of all variants in the F3-tolerant pool versus the F0-sensitive pool was determined using the Multipool programme (Edwards & Gifford, 2012). Several intervals with bi-polar peaks were identified across the genome. Intervals where the AF between the two pools differed considerably and had a positive LOD score were further explored for candidate genes.

The whole genome of four $A1_{(c+)}$ and four $T6_{(c-)}$ individuals was sequenced in a previous study (Busoms et al., 2018) and were merged with the two parental samples sequenced for the BSA-seq experiment. To obtain a consensus sequence for each deme, private SNPs of T6 and A1 (shared for the 6 samples) were selected using GATK and Col-0 TAIR10 sequence as a reference. Once we obtained the consensus sequence, we selected the sites that differ between demes using GATK 'concordance' command obtaining one vcf file with T16 AF < 0.1 and A1 AF > 0.9 variants and a second vcf file with T16 AF > 0.9 and A1 AF < 0.1 variants. We merged the files using VCFtools (Danecek et al., 2011) and the amino acid changes between A1 and T6 were obtained and quantified using SNPeff (Cingolani, 2022). Genes with 3–10 genetic modifiers were selected as candidates (Supporting Information: Data set SD3).

2.5 | RNA isolation and microarray scanning

Root tissue from 15-day-old plants cultivated under hydroponic conditions was recollected at 48 h after initiating the treatments. At least 12 plants per line and treatment were pooled to perform three biological replicates. Leaves were immersed in liquid nitrogen, homogenized to a fine powder, and stored at −80°C. The total RNA of about 100 mg of leaf powder for each biological replicate was extracted using the Maxwell® plant RNA kit (Promega Corporation) following the manufacturer's instructions. To ensure a high quality and quantity RNA material, Experion RNA Analysis Kit (Biorad) was performed. After quality control, total RNA was tagged, hybridized and cleaned using GeneChip® 3′ IVT Express Kit. Microarray expression analysis were performed by Affymetrix GeneChip™ Arabidopsis Genome ATH1 Array Agilent 2100 Bioanalyzer.

Differential expression analysis: The function *gcrma* (package gcrma) was used to adjust background intensities in Affymetrix array data, which include optical noise and nonspecific binding. The R package 'limma' was used to test the expression data and search for differentially expressed genes (DEGs) when pairs of experimental groups were compared. A model matrix was defined by specifying the groups that belonged to each sample. Then, limma's functions *lmFit* and *contrasts. fit*, together with the model matrix, were used to adjust the expression data to a linear model to extract fold changes and confidence statistics associated to the comparisons of interest. Function *eBayes* was then used to rank genes in order of evidence for differential expression. The R library *ath1121501.db* was used to include names and descriptions of the genes associated to the microarray's features probed. The resulting *p*-values were adjusted using the Benjamini and Hochberg's approach (Benjamini & Yekutieli, 2001) for controlling the false discovery rate. To perform a list of DEGs, genes were filtered by adj p-value (adjusted *p*-value) < 0.05 and log fold change (LFC) > 1 and LFC < −1.

Gene ontology (GO), KEGG pathway and functional protein association networks analysis: GO enrichment analysis of DEGs was implemented by AgriGO V2 (Tian et al., 2017) and FLAME (Thanati et al., 2021). Significant GO terms ($p < 0.05$) were classified into three categories: biological function, molecular process and cellular component. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and STRING version 11.0 were used to understand high-level function and gene interaction network from differential expressed genes (Szklarczyk et al., 2019).

2.6 | Statistical analysis of phenomics and ionomics

Normal distribution of data was confirmed by Levene's test and nonnormal data were transformed before applying any parametric tests. Mean-standardized values (1 < value > 1) of elemental contents of root material were used to represent the radar plots and compare each accession profile. One-way or multivariate analysis of variance was used to test for significant differences ($p < 0.05$) between means of data with respect to physiological parameters, ionome, fitness and gene expression. To test for correlations between two variables, a bivariate fit was applied. To perform multiple comparisons of group means, Tukey's HSD test was conducted. To perform multiple comparisons to a control (wild type [WT]), Dunnett's test was applied. Statistical analyses and plots were performed using SAS Software JMP v.16.0 and are available at Supporting Information: Data sets SD2 and SD5.

3 | RESULTS AND DISCUSSION

Natural variation of wild species provides a relevant complementary resource to discover novel gene functions. It has been proven that *A. thaliana* demes from the north-east of Spain harbour high genetic diversity (Castilla et al., 2020) and they have a great potential for studying traits related to adaptation (e.g., Pérez-Martín et al., 2022) and allelic variants that specifically interact with the environment (e.g., Busoms et al., 2021). Terés et al. (2019) selected two demes with contrasting phenotypes in soil carbonate tolerance, $A1_{(c+)}$ and $T6_{(c-)}$, and performed several studies to characterize the physiological mechanisms activated by these demes under carbonate and iron-deficiency stress.

3.1 | Root characterization of *A. thaliana* demes under high pH and bicarbonate stress

The hydroponic cultivation of $A1_{(c+)}$, $T6_{(c-)}$ and Col-0 plants under control pH (pH 5.9), high pH (pH 8.3), or bicarbonate (bic, 10 mM $NaHCO_3$, pH 8.3) conditions confirmed the higher tolerance of $A1_{(c+)}$ to both pH 8.3 and bic treatments. Col-0 exhibited an intermediate response, while $T6_{(c-)}$ was the most sensitive (Figure 1, Pérez-Martín et al., 2021). Root length was hardly affected by exposure to bic in $A1_{(c+)}$, the deme evolved on the calcareous soil. Contrastingly, root growth in $T6_{(c-)}$ was strongly inhibited by bic, but less affected by high pH (Figure 1a). In *A. thaliana* Col-0, used as a well-established reference genotype, root length was more inhibited by bicarbonate than $A1_{(c+)}$, but less than the sensitive deme $T6_{(c-)}$. In contrast to both natural demes, Col-0 did not show difference in root growth inhibition between high pH and the bic treatment (Figure 1a).

The standardized root mineral content revealed clear differences in response to high pH and bic among the three *A. thaliana* variants (Figure 1b). In deme $A1_{(c+)}$ both high pH and bic reduced root Fe, Cu and Zn concentrations in comparison to the control treatment. However, bicarbonate stress enhanced the uptake of macronutrients such as Ca and P in $A1_{(c+)}$ and Col-0 (Supporting Information: Figure S1). In general, high pH and bic reduced Fe contents in both organs in all plants. However, $A1_{(c+)}$ distinctively maintained higher Fe levels under bicarbonate exposure and the Fe translocation from roots to shoots was clearly enhanced in comparison to the control treatment (Figure 1c). Contrastingly, the sensitive $T6_{(c-)}$ had the lowest leaf Fe content despite having higher root Fe levels (Supporting Information: Figure S1). Because of this, $T6_{(c-)}$ exhibited a clear inhibition by bicarbonate of Fe translocation from roots to shoots, while high pH strongly enhanced the Fe translocation in this sensitive deme (Figure 1c).

3.2 | Transcriptional changes in response to bicarbonate stress

Microarray analyses in roots were performed to characterize, at the gene expression level, the mechanisms underlying the differential response to bicarbonate of the two naturally selected accessions. Gene responses were assessed by using samples after 48 h of treatment exposure, when the plants still did not exhibit any foliar symptoms of the stress treatments. We performed pairwise comparison to understand the differences between $A1_{(c+)}$-tolerant and $T6_{(c-)}$ -sensitive lines in response to different treatments (pH 8.3

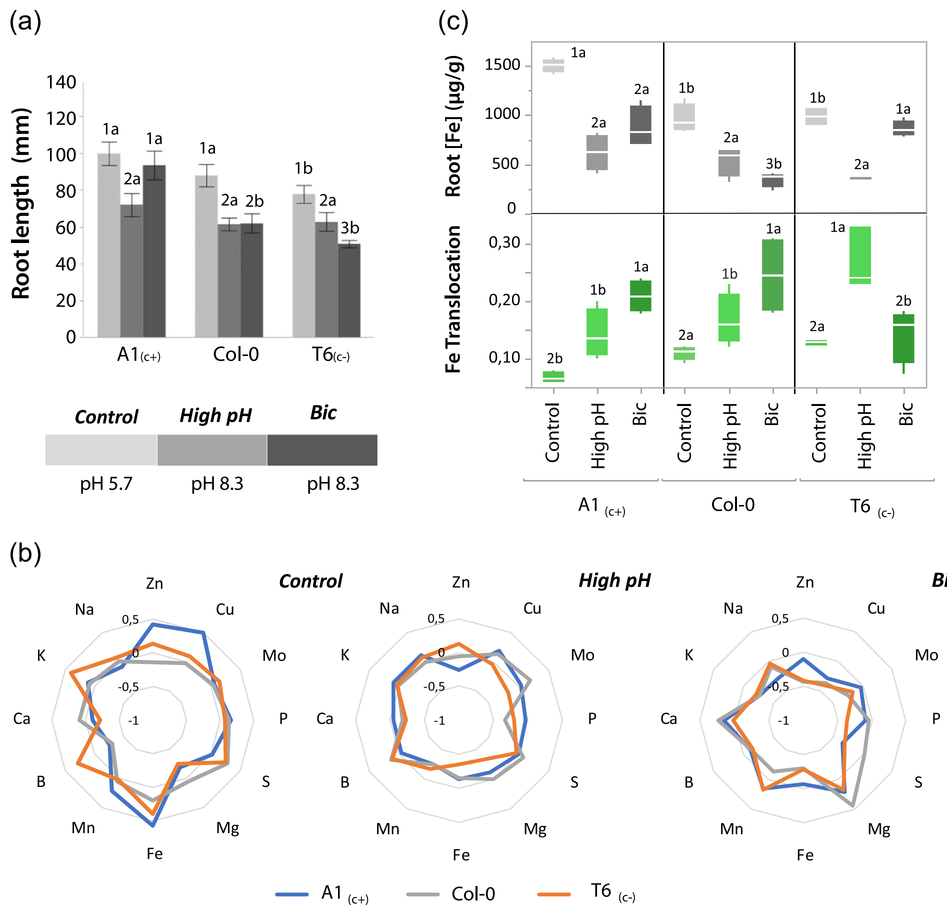


FIGURE 1 Root growth and ionome. (a) Root length (cm), (b) root ionome profile (mean-standardized values of 12 elements) and (c) root Fe concentration (µg/g) and Fe translocation ratio (leaf [Fe]/root [Fe]) of A1(c+), T6(c-) and Col-0 *Arabidopsis thaliana* plants cultivated in hydroponics under control (pH 5.7), high pH (pH 8.3) and bic (10 mM NaHCO3, pH 8.3) conditions for 10 days. Numbers indicate significant differences between treatments per accession and letters indicate significant differences between accessions under the same treatment (*t*-test, $p < 0.05$). [Color figure can be viewed at wileyonlinelibrary.com]

vs pH 5.9 and bic vs control pH 5.9) (Supporting Information: Data set SD3). A total of 4595 DEGs were identified, considering accession and treatment (Figure 2). The bicarbonate treatment caused a higher number of DEGs than the high pH treatment, indicating that bicarbonate stress involves more complex processes than simply the specific responses to alkaline pH (Figure 2a). For that, we further focused our analysis on the pathways found in the plants exposed to the bic treatment.

Previously, RNA-Seq of leaves from $A1_{(c+)}$ and $T6_{(c-)}$ individuals revealed that bicarbonate exposure quickly upregulated Fe-deficiency related genes in the sensitive $T6_{(c-)}$ but not in the tolerant $A1_{(c+)}$ (Pérez-Martín et al., 2021). In leaves, the highest number of DEGs was observed in $T6_{(c-)}$ but in roots the tolerant $A1_{(c+)}$ exhibited a huge response when exposed to bic (Figure 2b). These differences in DEG number may denote contrasted deme strategies toward bicarbonate stress. GO terms indicated that in roots of $A1_{(c+)}$ the main pathways activated after 48 h of exposure were related to activation of biological processes and metabolism located in the extracellular region (Figure 2c, Supporting Information: Data set SD4). Also, enzymatic regulatory activity, transport and transcription factors were altered. In $T6_{(c-)}$ only up and down modifications in catalytic activity were found. The analysis of the KEEG pathways revealed that the sensitive line activated only two principal pathways involved in catalytic activity while in $A1_{(c+)}$ more than 20 pathways were up or downregulated (Figure 2d, Supporting Information: Data set SD4). This analysis indicates that the modifications are being produced in the sugar, lipid and protein metabolism of the tolerant line.

In rice, better tolerance to saline-alkaline conditions was due to superior Fe acquisition and higher root-to-shoot Fe translocation (Li et al., 2016). This was attributed to enhanced root development and higher expression of genes involved in Fe acquisition and transport in the tolerant variety in comparison to the sensitive one. Contrastingly, among the DEGs in our tolerant *A. thaliana* deme, no differential enhancement of main Fe-acquisition-related genes was observed. Moreover, root expression of *bHLH101* was more than 4 times higher in the sensitive $T6_{(c-)}$ than in the tolerant $A1_{(c+)}$ (Supporting Information: Data set S5). bHLH101 is a transcription factor that controls plant Fe homoeostasis in a FIT-independent way (Sivitz et al., 2012). It is required for proper growth under Fe deficiency

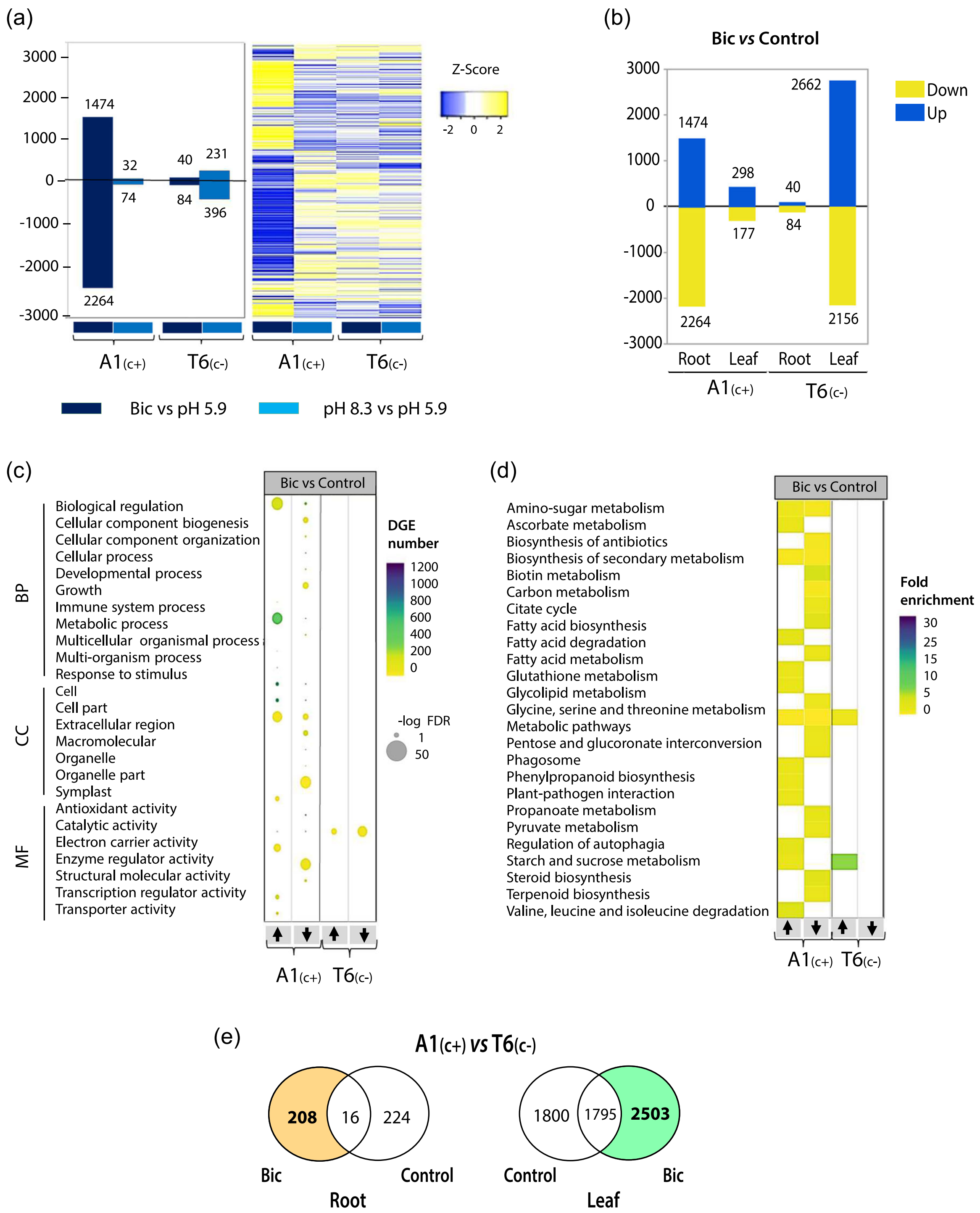


FIGURE 2 Transcriptomics. (a) Total number and heatmap profile of differentially expressed genes (DEGs) from pairwise comparisons bic versus control (dark blue) and high pH versus control (light blue) in A1$_{(c+)}$ and T6$_{(c-)}$ demes. (b) Total number of pairwise comparisons bic versus control in roots and leaves of A1$_{(c+)}$ and T6$_{(c-)}$ demes. Bubble plots indicating (c) significant Gene Ontology (GO) analysis and (d) heatmap of KEGG pathway analysis of A1$_{(c+)}$ and T6$_{(c-)}$ root DEGs in bic versus control comparison. Arrows indicate up or downregulated genes. (e) Venn diagrams of pairwise comparisons A1$_{(c+)}$ bic versus T6$_{(c-)}$ bic and A1$_{(c+)}$ control versus T6$_{(c-)}$ control in roots and leaves. Selected DEGs are highlighted in brown for root and green for leaves. DEGs, GO and KEGG terms were filtered at log fold change (LFC > 1, LFC < −1) and adjusted $p < 0.05$. [Color figure can be viewed at wileyonlinelibrary.com]

conditions. These results further confirm that Fe deficiency is earlier perceived by the sensitive $T6_{(c-)}$.

Focusing on $A1_{(c+)}$ versus $T6_{(c-)}$ comparison under the bic treatment and excluding the genes also up or downregulated in control conditions, 208 DEGs were identified in roots, 2503 DEGs in leaves, and 26 DEGs in both tissues (Figure 2d, Supporting Information: Data set SD3). Two contrasted root strategies imply the possibility of differences in signal perceiving and transduction. To further explore these differential mechanisms, we performed protein–protein interaction network functional enrichment analysis with the STRING database on the 208 genes (Figure 3, Supporting Information: Data set SD4). These genes can be classified in five main keywords according to FLAME: catalytic, binding and ubiquitin activity; phosphorus metabolic process and oxidoreductase activity. We found two main cores of genes with high network connectivity: one lead by genes involved in nucleus binding (TOPII, CDC20.2, MAD2, AUR1, FZR3, CYCB1;3) and microtubule movement regulation (AT5G60930, ATK5, AT3G20150, AT4G15830) that are indispensable for normal plant development and fertility but have not previously been associated with abiotic stress resistance; and another controlled by the NIA1 and Nitrite Reductase 1 (NIR1) genes. NIA1 (or GNR1) encodes the cytosolic minor isoform of nitrate reductase and NIR1 is involved in the second step of nitrate assimilation. Contrastingly, *NRT2.3* was less expressed in $A1_{(c+)}$ than in the sensitive $T6_{(c-)}$. NRT2.3 is a high-affinity nitrate transporter with still poorly defined functions but linked to the signalling of several phytohormones (Kiba et al., 2011). The occurrence of inorganic carbon and nitrogen in karst and carbonated soils affects the carbon/nitrogen metabolism of plant species. Under bicarbonate stress, growth reduction is enhanced -due to the inhibition of the photosynthesis and nitrogen metabolism but water use efficiency is promoted in tolerant plants (Xia & Wu, 2022). This balance is vital for plants to adapt to alkaline environments and the regulation of the 'nitrogen hub genes' found here might be important for C fixation and bicarbonate tolerance.

3.3 | Phenotypic variation for soil carbonate tolerance traits among demes

The natural habitat of $A1_{(c+)}$ and $T6_{(c-)}$ differs mainly in soil pH ($A1_{(c+)}$: 7.4; $T6_{(c-)}$: 6.5) and soil carbonate content ($A1_{(c+)}$: 12%; $T6_{(c-)}$: 0.8%) (Terés et al., 2019). Natural carbonated soil from a location closer to $A1_{(c+)}$ deme (pH = 7.9; $CaCO_3$ = 33%) was excavated for studying the tolerance capability of each deme and their progeny. F1 progeny of both crosses (A1 × T6 & T6 × A1) exhibited similar growth and fitness as the tolerant parental $A1_{(c+)}$ (Figure 4a,c), while F2 progeny showed segregation (Figure 4b,c). In addition to symptoms of iron deficiency and reduced growth, $T6_{(c-)}$ grown on alkaline soil suffers from delayed flowering and infertility that substantially hamper its fitness (Supporting Information: Figure S2, Supporting Information: Data set SD5). Only 35% of F2 progeny plants were able to flower and reproduce (Figure 4b). In consequence, the fitness of both crosses did not strictly fulfill the Mendelian phenotype ratio (3:1) (Figure 4b,c), suggesting that the 'soil carbonate tolerance phenotype' in our demes is partial-dominant or a polygenic trait. Moreover, considering only the plants that were able to flower, seed production was higher in the A1 × T6 than in the T6 × A1 offsprings (Figure 4c) pointing to a potential parental effect.

3.4 | Identification of genes associated with soil carbonate tolerance by BSA-Seq analysis

During meiosis, recombination reasserts the complement of alleles segregating in hybrid progeny (Salomé et al., 2012). Thus, the study of F_2 populations from contrasted phenotype crosses can provide valuable information on genetic mechanism by association (Benowicz et al., 2020). With the purpose of performing a BSA-Seq study, plants from 65 families of the A1 × T6 cross and from 28 families of the T6 × A1 cross were self-pollinated. F3 progeny plants were cultivated and screened under the same soil and growth conditions to select the extreme pools (Supporting Information: Data set SD5). We realized that the most sensitive families had very low germination rates, or the seedlings died after a few days postgermination (NP plants), therefore, we were not able to sequence the sensitive pool. Despite this, considering the families with a higher number of flowered individuals and with an elevated silique production, seven families were selected as the tolerant pool ($F3_{TP}$) (Figure 4d).

The AF comparison of the $F3_{TP}$ with the sensitive parental ($T6_{(c-)}$) enabled the identification of the top 0.5% outlier SNPs from the empirical distribution (Figure 4e). The 5617 SNPs of both tails were associated to 1119 genes, mainly located in chromosome 5. Within this chromosomal group, the higher proportion of SNPs were assigned to the short arm (Figure 4f, Supporting Information: Data set SD6). Salomé et al. (2012) found that crossovers and segregation distortion in F2 populations from different *A. thaliana* accessions were more frequent in chromosomes 1 and 5. Here, 68% of our phenotype-associated divergent outlier genes belong to chromosome 5, suggesting that this chromosome harbours key polymorphisms that facilitate soil carbonate tolerance.

Additionally, we sequenced the whole genome of 6 $A1_{(c+)}$ and 6 $T6_{(c-)}$ individuals collected from their natural habitat. It is well known that amino acid substitutions in a protein can cause a drastic phenotypic effect (Bartlett & Whipple, 2013). The AF comparison between $A1_{(c+)}$ and $T6_{(c-)}$ picked out 977 genes with 3–10 nonsynonymous variants with high or moderate predicted effect (Supporting Information: Data set SD7). There are many examples of genomic variants whose frequencies are correlated with environmental variables and temporal changes consistent with natural selection in *A. thaliana* (Castilla et al., 2020; Monroe et al., 2022). $A1_{(c+)}$ is locally adapted to carbonated soils, thus it was important to consider the genomic differences between $A1_{(c+)}$ and $T6_{(c-)}$ to narrow the BSA results down. The 69 genes in common between these two analyses were selected as the 'genomic candidate genes list' (Figure 4g, Supporting Information: Data set SD8).

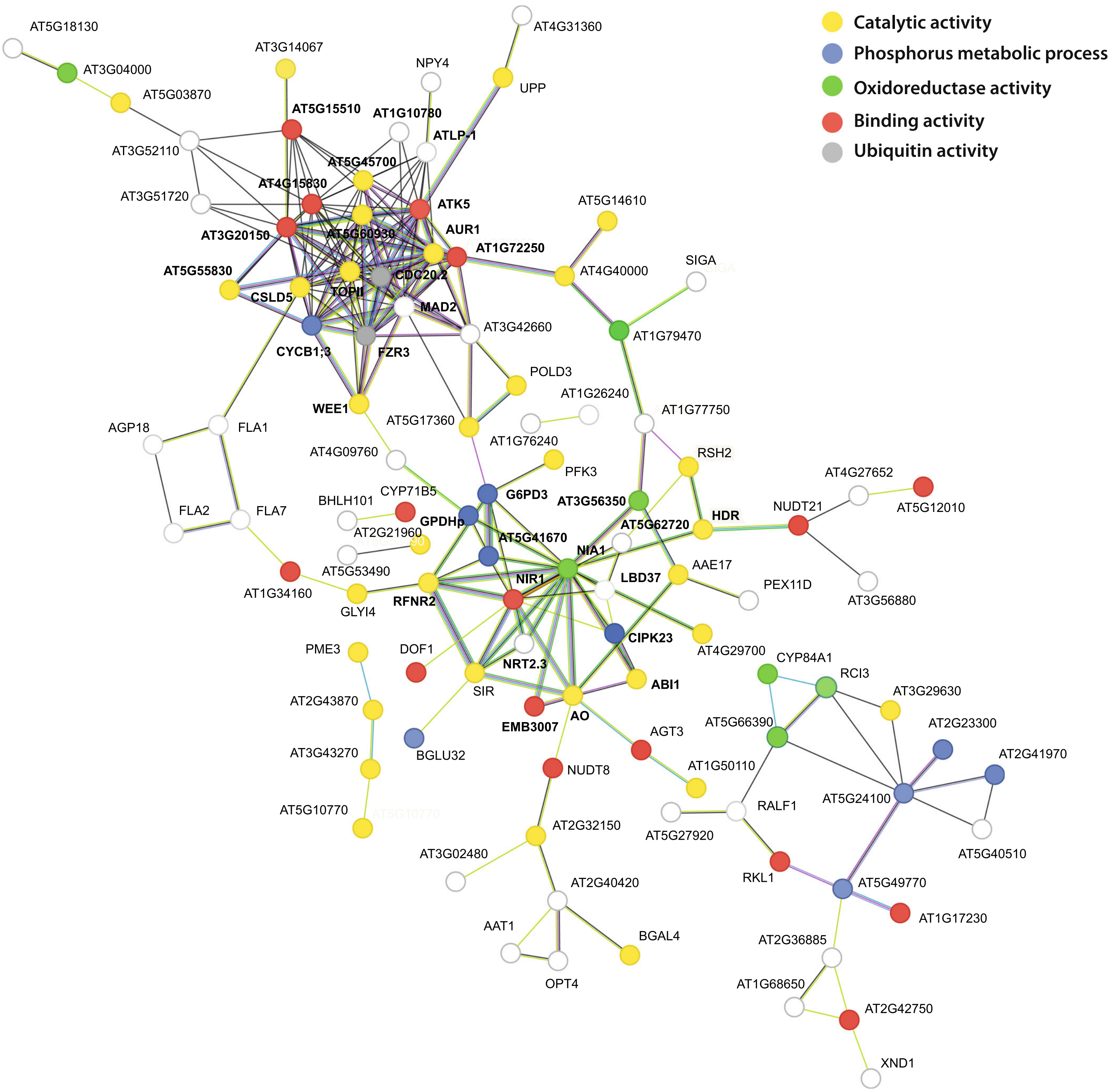


FIGURE 3 Protein–protein interaction network of bicarbonate stress. Gene protein interaction network of the 208 root differentially expressed genes up or downregulated in the tolerant $A1_{(c+)}$ under the bic treatment. Each sphere corresponds to one gene, nodes represent protein interactions (source: STRING. org) and colours represent Gene Ontology terms. *Source*: bib. fleming. gr. [Color figure can be viewed at wileyonlinelibrary.com]

3.5 | Association analysis for prediction of candidate genes

It is not feasible to determine the consequences of the missense variants on gene expression, protein structure and function of an elevated number of putative candidate genes. Alternatively, candidate gene lists can be bounded conducting association analysis such as the successfully applied to identify genes associated with plant architecture in *Brassica napus*, (Ye et al., 2022) water stress responses in wheat (Derakhshani et al., 2020) or chilling stress tolerance in rice (Guo et al., 2020). Here, we performed an association analysis by combining the genomic and transcriptomic data described above. The candidate gene lists obtained from the BSA-Seq and the parental comparison analysis (Supporting Information: Data set SD6–8) was compared with the root and leaf DEGs exclusive of the bic treatment in $A1_{(c+)}$ versus $T6_{(c-)}$ comparison (Supporting Information: Data set SD3), resulting in 18 matching genes (Figure 5a,b). The expression pattern of

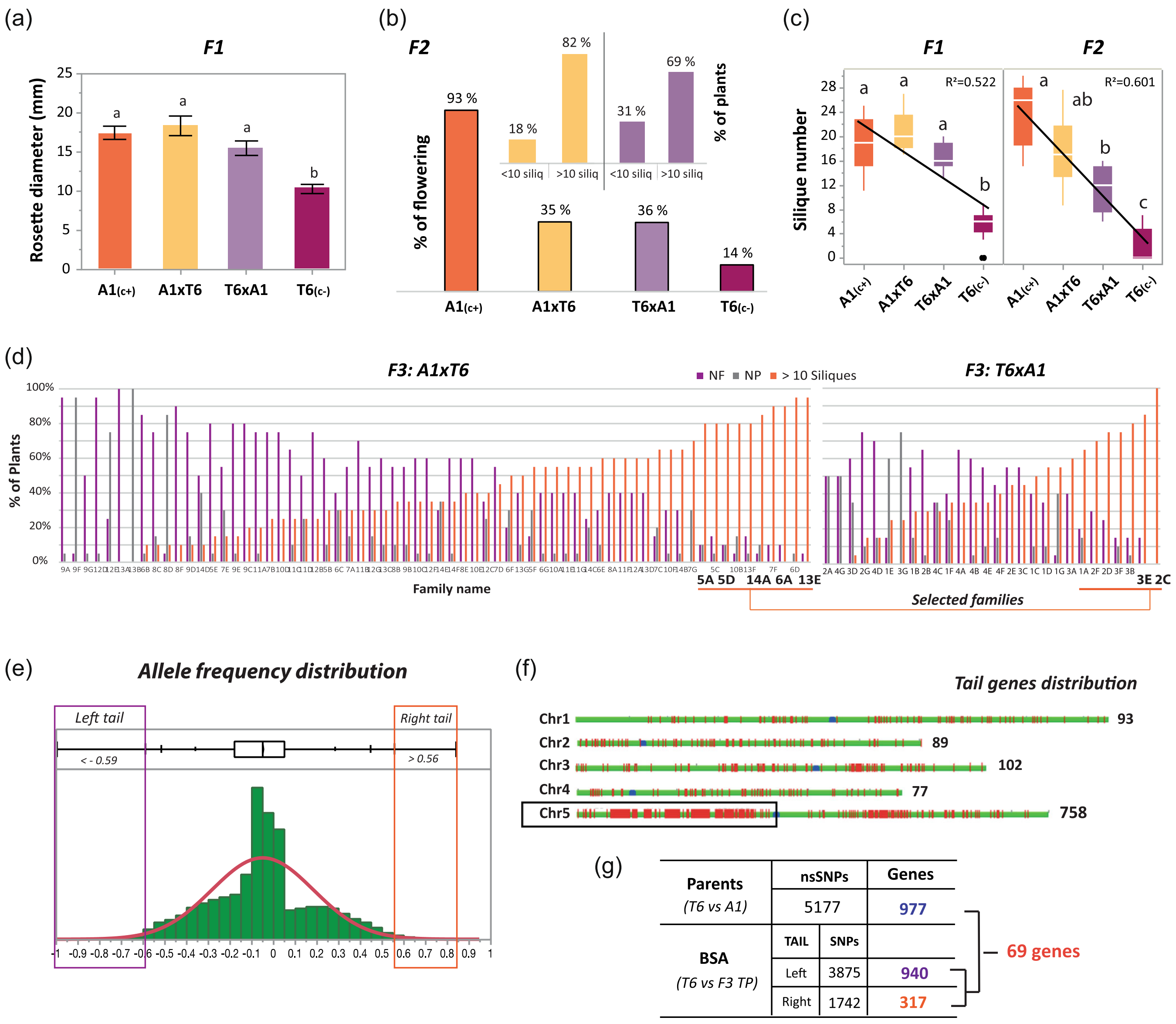
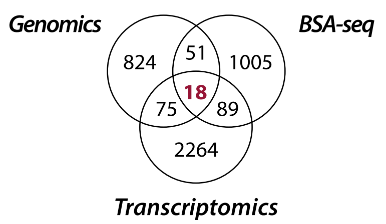


FIGURE 4 *Arabidopsis thaliana* soil carbonate tolerance inheritance and bulk segregant analysis sequencing (BSA-Seq). (a) Rosette diameter (mm); (b) percentage of flowered plants and plants with more or less than 10 siliques and (c) fitness (silique number) of $A1_{(c+)}$, $T6_{(c-)}$ parents, F1 and F2 progeny from reciprocal crosses cultivated in natural carbonated soil. (d) Percentage of plants that did not prosper (no plant, purple bars), plants that did not flower (not flower, grey bars) and plants that produce more than 10 siliques (orange bars) from A1 × T6 F3 and T6 × A1 F3 progeny cultivated in natural carbonated soil. Selected families for BSA analysis are indicated in bold. (e) Allele frequency distribution of SNPs detected between the sensitive parent ($T6_{(c-)}$) and the F3 tolerant pool. (f) Chromosome distribution of genes associated with the SNPs from BSA left and right tails. (g) Summary table of the SNPs and genes detected and shared between parent's comparison ($A1_{(c+)}$ versus $T6_{(c-)}$) and between $T6_{(c-)}$ and the F3 tolerant pool comparison. Letters indicate significant differences (Tukey's HSD, $p < 0.05$). [Color figure can be viewed at wileyonlinelibrary.com]

these genes confirms that generally they are more expressed in the roots of $A1_{(c+)}$ but also in the leaves of the sensitive $T6_{(c-)}$ deme (Figure 5c,d, Supporting Information: Data set SD9). This suggests that roots of the tolerant $A1_{(c+)}$ are reacting quickly to the adverse alkaline condition avoiding secondary effects in the leaves. Contrastingly, in the sensitive $T6_{(c-)}$ leaves are promptly stressed, especially due to the inhibition of Fe translocation (Figure 1c), and leaf gene expression adjusts to these alterations to alleviate further injury.

The SNP changes present in these 18 genes were explored a several missense variants were identified in both demes. The high number of SNPs present in Allantoate Amido Hydrolase (AHH) and PCH1 of $A1_{(c+)}$, and in INV-E of $T6_{(c-)}$ stands out (Supporting Information: Data set SD9, Supporting Information: Figure S4). Some of these variants could be responsible for the observed gene expression differences or be linked to structural changes that affect the regulation of the gene. To evaluate the consequences of the no-expression of these genes under carbonate stress, the

(a)



(b)

Identifier	Primary Gene Symbol	Mutant
AT1G14130	DIOXYGENASE FOR AUXIN OXIDATION 1 **(DAO1)**	m01
AT1G66430	FRUCTOKINASE 6 **(FRK6)**	m02
AT2G03760	SULPHOTRANSFERASE 12 **(SOT12)**	m03
AT2G16365	PHOTOPERIODIC CONTROL OF HYPOCOTYL 1 **(PCH1)**	m04
AT2G19910	RNA-dependent RNA polymerase family protein	m05
AT4G02060	PROLIFERA **(PRL)**	m06
AT4G20070	ALLANTOATE AMIDOHYDROLASE **(AAH)**	m07
AT5G07440	GLUTAMATE DEHYDROGENASE 2 **(GDH2)**	m08
AT5G08010	Hypothetical protein	m09
AT5G13220	JASMONATE-ZIM-DOMAIN PROTEIN 10 **(JAZ10)**	m10
AT5G15900	TRICHOME BIREFRINGENCE-LIKE 19 **(TBL19)**	m11
AT5G15920	STRUCTURAL MAINTENANCE OF CHROMOSOMES 5 **(SMC5)**	m12
AT5G16250	Transmembrane protein	m13
AT5G16390	CHLOROPLASTIC ACETYLCOENZYME A CARBOXYLASE 1 **(CAC1)**	m14
AT5G16910	CELLULOSE-SYNTHASE LIKE D2 **(CSLD2)**	m15
AT5G19940	FIBRILLIN6 **(FBN6)**	m16
AT5G22510	ALKALINE/NEUTRAL INVERTASE E **(INV-E)**	m17
AT5G23920	Transmembrane protein	m18

(c)

(d)

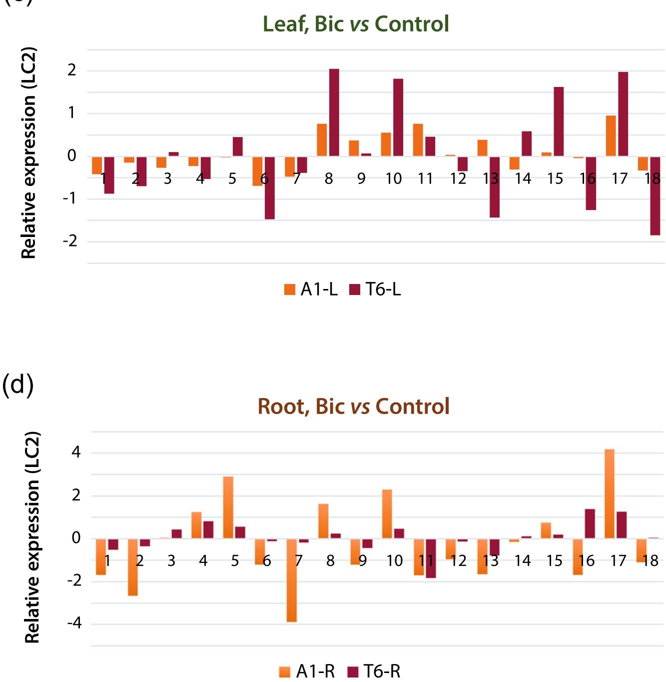


FIGURE 5 Combined analysis for bicarbonate tolerance candidate genes detection. (a) Number and (b) description of putative candidate genes obtained from the genomic and transcriptomic association analysis of $A1_{(c+)}$ and $T6_{(c-)}$ demes. Relative fold change (bic versus control) of the 18 candidate genes in the (c) leaves and (d) roots of $A1_{(c+)}$ (orange bars) and $T6_{(c-)}$ (purple bars) plants submitted to bic stress (10 mM $NaHCO_3$, pH 8.3) for 48 h. [Color figure can be viewed at wileyonlinelibrary.com]

knock-out mutants of the 18 genes (T-DNA lines, Supporting Information: Data set SD1), together with the wild type of Col-0 and $A1_{(c+)}$ and $T6_{(c-)}$, were cultivated on the same carbonated soil (LP) and under the same control (pH 5.9) and bicarbonate hydroponic conditions (10 mM $NaHCO_3$, pH 8.3) used in our previous studies.

Seeds of m03 (Sulphotransferase12) did not germinate in any condition and, therefore, *sot12* could not been tested. Instead, *gdh2* (m08) germinated normally under carbonated soil and we observed an enhanced fitness of this mutant in this condition (Supporting Information: Figure S3A,B). Both mutants *csld2* (m15) and *fbn6* (m16) coding, respectively, for Cellulose Synthase Like D2 and Fibrillin6 died only when treated with bicarbonate stress under hydroponic conditions (Supporting Information: Figure S3A,B), suggesting hypersensitivity to this stress. CSLD2, is required for normal root hair development (Bernal et al., 2008). However, the inhibition of root hair development alone may not cause plant death. The failure to survive is more likely related to the putative role of CSLD2 in collaboration with CSLD3 and CSLD5 in hemicellulose biosynthesis (Yin et al., 2011) affecting early plant development. FBN6 is also required for normal plant development and *fbn6* mutants have stunted root growth. This has been related to alteration in sulphate reduction, enhanced glutathione and cadmium tolerance (Lee et al., 2020). We have previously shown that bicarbonate tolerance in *A. thaliana* was associated with the upregulation of leaf genes related to sulphate acquisition (Pérez-Martín et al., 2021). So far FBN6 has mainly been studied in relation to chloroplasts, while the role in root development is still unknown. Clearly, further characterization of the role of FBN6 in plant development under bicarbonate is required.

On the other hand, *dao1* (m01), *aah* (m07) and *jaz10* (m10) grew significantly better than the WT under bic stress (Supporting Information: Figure S3A) but only *dao1* exhibited higher fitness when it was cultivated on carbonated soil (Supporting Information: Figure S3B). *DAO1*, coding for a 2-oxoglutarate and Fe (II)-dependent dioxygenase, is a main player in irreversible auxin degradation (Hayashi et al., 2021) and a higher production of adventitious roots has been shown in the *dao1-1* loss of function mutants (Lakehal et al., 2019). Downregulation of *DAO1* could enhance root development and Fe acquisition, being beneficial in high pH habitats. AHH catalyzes the degradation of allantoin yielding CO_2 and four molecules of NH_4 (Werner et al., 2008). High levels of allantoin and downregulation of allantoin degradation has been found to enhance salt stress tolerance in *A. thaliana* (Lescano et al., 2016). Our findings regarding the bicarbonate tolerance of the *ahh* mutant (Supporting Information: Figure S3A) and the lower root expression of AHH in $A1_{(c+)}$ indicate (Figure 5c,d) a role for allantoin also in bicarbonate tolerance. In *A. thaliana*, JAZ proteins act as repressors of JA signalling and plants overexpressing *AtJAZ10* are insensitive to JA (Chung & Howe, 2009; Sun et al., 2017). In our case, *jaz10* exhibited a better response to bicarbonate stress than the WT and *JAZ10* was less expressed in the tolerant $A1_{(c+)}$ deme, suggesting that a downregulation of this gene could trigger JA signals important for tolerant responses.

Trichome Birefringence Like 19 -*tbl19* (m11) did not grow or produced more siliques than Col-0 but was not affected by the carbonate treatment, exhibiting higher tolerance (Supporting Information: Figure S3A,B). In both demes *TBL19* was downregulated in roots and upregulated in leaves (Figure 5). Under bicarbonate exposure, the knock-out mutant had high root biomass and the highest fitness as expressed in number of siliques (Supporting Information: Figure S3). TBL19, also known as AtXYBAT1 [63], has O-acetyltransferase activity and is probably involved in the acetylation of cellohexaose, a hemicellulose in *A. thaliana* cell walls. Root apoplastic iron in *A. thaliana* is mainly stored in the operationally defined HC1 fraction of hemicellulose and this bound-Fe can be mobilized by coumarin-type phenolic exudation (Lei et al., 2014). Under bicarbonate exposure, the tolerant $A1_{(c+)}$ has higher exudation of these phenolics than the sensitive $T6_{(c-)}$ (Terés et al., 2019). The decrease of acetylation may enhance Fe binding to cell wall hemicellulose and its accessibility to the coumarin-type phenolics, so enhancing the root-to-shoot Fe translocation specifically in $A1_{(c+)}$.

Alkaline/Neutral Invertase (INV-E) encodes a chloroplast-targeted protein, although it is expressed in root tissues in *A. thaliana* (TAIR-BAR). INV-E catalyzes the irreversible cleavage of sucrose into glucose and fructose and some *At*INV have been linked to the control of root cell elongation mediated by sugars (Ruan et al., 2010). *inv-e* (m17) mutant was clearly affected by the bicarbonate content both in hydroponics and in soil (Supporting Information: Figure S3A,B). The SNP changes detected in $A1_{(c+)}$ and $T6_{(c-)}$ were evaluated and we found that $T6_{(c-)}$ INV-E gene contains a large number of variants that could affect the expression of the gene (Supporting Information: Figure S4, Supporting Information: Data set SD9), supporting the hypothesis that a proper regulation of INV-E could play an important role in bicarbonate tolerance.

In conclusion, differential bicarbonate tolerance in *A. thaliana* is associated with enhanced root-to-shoot translocation of iron, proper distribution of C/N fixation, and better root development. Association analysis of genomics and transcriptomics with natural demes well- or mal-adapted to soil carbonate allowed the selection of 18 candidate genes for this differential tolerance. Evaluation of genomic variation and the corresponding knock-out mutants suggested implications of AHH, CSLD2, DAO1, GDH2, FBN6, INV-E, JAZ10 and TBL19. We propose these genes as candidates for being further explored through complementary analysis under bicarbonate stress conditions.

ACKNOWLEDGEMENTS

Special thanks to Sian Bray for her bioinformatic programming support. We are grateful to Maria José Almira for her collaboration in laboratory experiments and graphics; and to Rosa Padilla for processing the ionome data. This research was funded by the Spanish Ministry of Science & Innovation (PID2019-104000RB-I00).

ORCID

Roser Tolrà http://orcid.org/0000-0001-8513-3520
Charlotte Poschenrieder http://orcid.org/0000-0002-3818-0874

REFERENCES

Ahmad, P., Ozturk, M., Sharma, S. & Gucel, S. (2014) Effect of sodium carbonate-induced salinity-alkalinity on some key osmoprotectants, protein profile, antioxidant enzymes, and lipid peroxidation in two mulberry (*Morus alba* L.) cultivars. *Journal of Plant Interactions*, 9, 460–467. Available from: https://doi.org/10.1080/17429145.2013.855271

Al-Busaidi, A. & Cookson, P. (2003) Salinity—pH relationships in calcareous soils of Oman. *Agricultural and Marine Sciences*, 8, 41–46. Available from: https://journals.squ.edu.om/index.php/jams/article/view/615/614

Bartlett, M.E. & Whipple, C.J. (2013) Protein change in plant evolution: tracing one thread connecting molecular and phenotypic diversity. *Frontiers in Plant Science*, 4, 382. Available from: https://doi.org/10.3389/fpls.2013.00382

Benjamini, Y. & Yekutieli, D. (2001) The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 2001(29), 1165–1188. Available from: https://doi.org/10.1214/aos/1013699998

Benowicz, A., Stoehr, M., Hamann, A. & Yanchuk, A.D. (2020) Estimation of the F2 generation segregation variance and relationships among growth, frost damage, and bud break in coastal Douglas-Fir (*Pseudotsuga menziesii* (Mirb.) Franco) wide-crosses. *Annals of Forest Science*, 77, 28. Available from: https://doi.org/10.1007/s13595-020-0925-9

Bernal, A.J., Yoo, C.-M., Mutwil, M., Jensen, J., Hou, G., Blaukopf, C. et al. (2008) Functional analysis of the cellulose synthase-like genes CSLD1, CSLD2, and CSLD4 in tip-growing Arabidopsis cells. *Plant Physiology*, 148, 1238–1253. Available from: https://doi.org/10.1104/pp.108.121939

Busoms, S., Paajanen, P., Marburger, S., Bray, S., Huang, X.Y., Poschenrieder, C. et al. (2018) Fluctuating selection on migrant adaptive sodium transporter alleles in coastal *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, 115, E12443–E12452. Available from: https://doi.org/10.1073/pnas.1816964115

Busoms, S., Terés, J., Huang, X.-Y., Bomblies, K., Danku, J., Douglas, A. et al. (2015) Salinity is an agent of divergent selection driving local adaptation of Arabidopsis to coastal habitats. *Plant Physiology*, 168, 915–929. Available from: https://doi.org/10.1104/pp.15.00427

Busoms, S., Terés, J., Yant, L., Poschenrieder, C. & Salt, D.E. (2021) Adaptation to coastal soils through pleiotropic boosting of ion and stress hormone concentrations in wild *Arabidopsis thaliana*. *New Phytologist*, 232, 208–220. Available from: https://doi.org/10.1111/nph.17569

Cao, Y., Song, H. & Zhang, L. (2022) New insight into plant saline-alkali tolerance mechanisms and application to breeding. *International Journal of Molecular Sciences*, 23, 16048. Available from: https://doi.org/10.3390/ijms232416048

Castilla, A.R., Méndez-Vigo, B., Marcer, A., Martínez-Minaya, J., Conessa, D., Picó, F.X. et al. (2020) Ecological, genetic and evolutionary drivers of regional genetic differentiation in *Arabidopsis thaliana*. *BMC Evolutionary Biology*, 20, 71. Available from: https://doi.org/10.1186/s12862-020-01635-2

Chung, H.S. & Howe, G.A. (2009) A critical role for the TIFY motif in repression of jasmonate signaling by a stabilized splice variant of the JASMONATE ZIM-domain protein JAZ10 in Arabidopsis. *The Plant Cell*, 21, 131–145. Available from: https://doi.org/10.1105/tpc.108.064097

Cingolani, P. (2022) Variant Annotation and Functional Prediction: SnpEffIn: Ng, C. & Piscuoglio, S. (Eds.) *Variant calling. Methods in molecular biology*, 2493. New York, NY: Humana. https://doi.org/10.1007/978-1-0716-2293-3_19

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A. et al. (2011) The variant call format and VCFtools. *Bioinformatics*, 27, 2156–2158. Available from: https://doi.org/10.1093/bioinformatics/btr330

Derakhshani, B., Ayalew, H., Mishina, K., Tanaka, T., Kawahara, Y., Jafary, H. et al. (2020) Comparative analysis of root transcriptome reveals candidate genes and expression divergence of homoeologous genes in response to water stress in wheat. *Plants*, 9, 9596. Available from: https://doi.org/10.3390/plants9050596

Dreyer, M., Wichmann, M., Rischen, M., Görlach, B.M., Ehmke, A., Pitann, B. et al. (2020) Ammonium-driven nitrification plays a key role in increasing Mn availability in calcareous soils. *Journal of Plant Nutrition and Soil Science*, 183, 389–396. Available from: https://doi.org/10.1002/jpln.201900131

Duan, X., Yu, Y., Duanmu, H., Chen, C., Sun, X., Cao, L. et al. (2018) GsSLAH3, a *Glycine soja* slow type anion channel homolog, positively modulates plant bicarbonate stress tolerance. *Physiologia Plantarum*, 164, 145–162. Available from: https://doi.org/10.1111/ppl.12683

Edwards, M.D. & Gifford, D.K. (2012) High-resolution genetic mapping with pooled sequencing. *BMC Bioinformatics*, 13(Suppl 6), S8. Available from: https://doi.org/10.1186/1471-2105-13-S6-S8

Felle, H.H. (2001) pH: signal and messenger in plant cells. *Plant Biology*, 3, 577–591.

Filippi, P., Jones, E.J., Ginns, B.J., Whelan, B.M., Roth, G.W. & Bishop, T.F.A. (2019) Mapping the depth-to-soil pH constraint, and the relationship with cotton and grain yield at the within-field scale. *Agronomy*, 9, 251. Available from: https://doi.org/10.3390/agronomy9050251

Guo, Z., Cai, L., Chen, Z., Wang, R., Zhang, L., Guan, S. et al. (2020) Identification of candidate genes controlling chilling tolerance of rice in the cold region at the booting stage by BSA-Seq and RNA-Seq. *Royal Society Open Science*, 7, 201081. Available from: https://doi.org/10.1098/rsos.201081

Hayashi, K., Arai, K., Aoi, Y., Tanaka, Y., Hira, H., Guo, R. et al. (2021) The main oxidative inactivation pathway of the plant hormone auxin. *Nature Communications*, 12, 6752. Available from: https://doi.org/10.1038/s41467-021-27020-1

Hui, X., Luo, L., Wang, S., Cao, H., Huang, M., Shi, M. et al. (2019) Critical concentration of available soil phosphorus for grain yield and zinc nutrition of winter wheat in a zinc-deficient calcareous soil. *Plant and Soil*, 444, 315–330. Available from: https://doi.org/10.1007/s11104-019-04273-w

Joshi, N.A. & Fass, J.N. (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. Available from: https://github.com/najoshi/sickle

Kiba, T., Kudo, T., Kojima, M. & Sakakibara, H. (2011) Hormonal control of nitrogen acquisition: roles of auxin, abscisic acid, and cytokinin. *Journal of Experimental Botany*, 62, 1399–1409. Available from: https://doi.org/10.1093/jxb/erq410

Lakehal, A., Dob, A., Novák, O. & Bellini, C. (2019) A DAO1-mediated circuit controls auxin and jasmonate crosstalk robustness during adventitious root initiation in Arabidopsis. *International Journal of Molecular Sciences*, 20, 4428. Available from: https://doi.org/10.3390/ijms20184428

Lee, J.A. & Woolhouse, H.W. (1969) Root growth and dark fixation of carbon dioxide in calcicoles and calcifuges. *New Phytologist*, 68, 247–255. Available from: https://doi.org/10.1111/j.1469-8137.1969.tb06437.x

Lee, K., Lehmann, M., Paul, M.V., Wang, L., Luckner, M., Wanner, G. et al. (2020) Lack of FIBRILLIN6 in *Arabidopsis thaliana* affects light acclimation and sulfate metabolism. *New Phytologist*, 225, 1715–1731. Available from: https://doi.org/10.1111/nph.16246

Lei, G.J., Zhu, X.F., Wang, Z.W., Dong, F., Dong, N.Y. & Zheng, S.J. (2014) Abscisic acid alleviates iron deficiency by promoting root iron reutilization and transport from root to shoot in Arabidopsis. *Plant, Cell & Environment*, 37, 852–863. Available from: https://doi.org/10.1111/pce.12203

Lescano, C.I., Martini, C., González, C.A. & Desimone, M. (2016) Allantoin accumulation mediated by allantoinase downregulation and transport by Ureide Permease 5 confers salt stress tolerance to Arabidopsis plants. *Plant Molecular Biology*, 91, 581–595. Available from: https://doi.org/10.1007/s11103-016-0490-7

Li, H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27, 2987–2993. Available from: https://doi.org/10.1093/bioinformatics/btr509

Li, N., Liu, H., Sun, J., Zheng, H., Wang, J., Yang, L. et al. (2018) Transcriptome analysis of two contrasting rice cultivars during alkaline stress. *Scientific Reports*, 8, 9586.

Li, Q., Yang, A. & Zhang, W.-H. (2016) Efficient acquisition of iron confers greater tolerance to saline-alkaline stress in rice (*Oryza sativa* L.). *Journal of Experimental Botany*, 67, 6431–6444. Available from: https://doi.org/10.1093/jxb/erw407

Li, Z. & Xu, Y. (2022) Bulk segregation analysis in the NGS era: a review of its teenage years. *The Plant Journal*, 109, 1355–1374. Available from: https://doi.org/10.1111/tpj.15646

Loeppert, R.H. & Suarez, D.L. (1996) Carbonate and gypsum. In: Sparks, D.L., Page, A.L., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T. & Sumner, M.E. (Eds.) *Methods of soil analysis. Part 3 chemical methods, 5.3*. Madison, USA: Publisher: Soil Science Society of America and American Society of Agronomy, chapter 15. pp. 437–474.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A. et al. (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20, 1297–1303. Available from: https://doi.org/10.1101/gr.107524.110

Mehari, T.G., Xu, Y., Umer, M.J., Shiraku, M.L., Hou, Y., Wang, Y. et al. (2021) Multi-omics-based identification and functional characterization of Gh_A06G1257 proves its potential role in drought stress tolerance in *Gossypium hirsutum*. *Frontiers in Plant Science*, 12, 746771. Available from: https://doi.org/10.3389/fpls.2021.746771

Monroe, J.G., Srikant, T., Carbonell-Bejerano, P., Becker, C., Lensink, M., Exposito-Alonso, M. et al. (2022) Mutation bias reflects natural selection in *Arabidopsis thaliana*. *Nature*, 602, 101–105. Available from: https://doi.org/10.1038/s41586-021-04269-6

Pérez-Martín, L., Busoms, S., Almira, M.J., Azagury, N., Terés, J., Tolrà, R. et al. (2022) Evolution of salt tolerance in *Arabidopsis thaliana* on siliceous soils does not confer tolerance to saline calcareous soils. *Plant and Soil*, 476, 455–475. Available from: https://doi.org/10.1007/s11104-022-05439-9

Pérez-Martín, L., Busoms, S., Tolrà, R. & Poschenrieder, C. (2021) Transcriptomics reveals fast changes in salicylate and jasmonate signaling pathways in shoots of carbonate-tolerant *Arabidopsis thaliana* under bicarbonate exposure. *International Journal of Molecular Sciences*, 22, 1226. Available from: https://doi.org/10.3390/ijms22031226

Ruan, Y.L., Jin, Y., Yang, Y.J., Li, G.J. & Boyer, J.S. (2010) Sugar input, metabolism, and signaling mediated by invertase: roles in development, yield potential, and response to drought and heat. *Molecular Plant*, 3, 942–955. Available from: https://doi.org/10.1093/mp/ssq044

Sagervanshi, A., Naeem, A., Kaiser, H., Pitann, B. & Mühling, K.H. (2021) Early growth reduction in *Vicia faba* L. under alkali salt stress is mainly caused by excess bicarbonate and related to citrate and malate over accumulation. *Environmental and Experimental Botany*, 192, 104636. Available from: https://doi.org/10.1016/j.envexpbot.2021.104636

Salomé, P.A., Bomblies, K., Fitz, J., Laitinen, R.A.E., Warthmann, N., Yant, L. et al. (2012) The recombination landscape in *Arabidopsis thaliana* F2 populations. *Heredity*, 108, 447–455. Available from: https://doi.org/10.1038/hdy.2011.95

Sarkar, B., Hasanuzzaman, M. & Adak, M.K. (2022) Insights into the role of iron supplementation in conferring bicarbonate-mediated alkaline stress tolerance in maize. *Journal of Soil Science and Plant Nutrition*, 22, 2719–2734. Available from: https://doi.org/10.1007/s42729-022-00839-3

Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. (2012) NIH image to ImageJ: 25 years of image analysis. *Nature Methods*, 9, 671–675. Available from: https://doi.org/10.1038/nmeth.2089

Scholl, R.L., May, S.T. & Ware, D.H. (2000) Seed and molecular resources for Arabidopsis. *Plant Physiology*, 124, 1477–1480. Available from: https://doi.org/10.1104/pp.124.4.1477

Sivitz, A.B., Hermand, V., Curie, C. & Vert, G. (2012) Arabidopsis bHLH100 and bHLH101 control iron homeostasis via a FIT-independent pathway. *PLoS One*, 7, e44843. Available from: https://doi.org/10.1371/journal.pone.0044843

Sun, Q., Wang, G., Zhang, X., Zhang, X., Qiao, P., Long, L. et al. (2017) Genome-wide identification of the TIFY gene family in three cultivated Gossypium species and the expression of JAZ genes. *Scientific Reports*, 2017(7), 1–9. Available from: https://doi.org/10.1038/srep42418

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J. et al. (2019) String v11: protein-protein association networks with increased coverage, supporting functional discovery in genome –wide experimental datasets. *Nucleic Acids Research*, 47, D607–D613. Available from: https://doi.org/10.1093/nar/gky1131

Terés, J., Busoms, S., Perez Martín, L., Luís-Villarroya, A., Flis, P., Álvarez-Fernández, A. et al. (2019) Soil carbonate drives local adaptation in *Arabidopsis thaliana*. *Plant, Cell & Environment*, 42, 2384–2398. Available from: https://doi.org/10.1111/pce.13567

Thanati, F., Karatzas, E., Baltoumas, F., Stravopodis, D.J., Eliopoulos, A. & Pavlopoulos, G. (2021) FLAME: a web tool for functional and literature enrichment analysis of multiple gene lists. *Biol. (Basel)*, 10, 665. Available from: https://doi.org/10.3390/biology10070665

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z. et al. (2017) agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Research*, 45, W122–W129. Available from: https://doi.org/10.1093/nar/gkx382

Waters, B.M., Amundsen, K. & Graef, G. (2018) Gene expression profiling of iron deficiency chlorosis sensitive and tolerant soybean indicates key roles for phenylpropanoids under alkalinity stress. *Frontiers in Plant Science*, 9, 10. Available from: https://doi.org/10.3389/fpls.2018.00010

Werner, A.K., Sparkes, I.A., Romeis, T. & Witte, C.P. (2008) Identification, biochemical characterization, and subcellular localization of allantoate amidohydrolases from Arabidopsis and Soybean. *Plant Physiology*, 146, 323–324. Available from: https://doi.org/10.1104/pp.107.110809

Wilm, A., Aw, P.P.K., Bertrand, D., Yeo, G.H.T., Ong, S.H., Wong, C.H. et al. (2012) LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Research*, 40, 11189–11201. Available from: https://doi.org/10.1093/nar/gks918

Xia, A. & Wu, Y. (2022) Joint interactions of carbon and nitrogen metabolism dominated by bicarbonate and nitrogen in *Orychophragmus violaceus* and *Brassica napus* under simulated karst habitats. *BMC Plant Biology*, 22, 264. Available from: https://doi.org/10.1186/s12870-022-03646-1

Ye, S., Yan, L., Ma, X., Chen, Y., Wu, L., Ma, T. et al. (2022) Combined BSA-Seq based mapping and RNA-Seq profiling reveal candidate genes associated with plant architecture in *Brassica napus*. *International Journal of Molecular Sciences*, 23, 2472. Available from: https://doi.org/10.3390/ijms23052472

Yin, L., Verhertbruggen, Y., Oikawa, A., Manisseri, C., Knierim, B., Prak, L. et al. (2011) The cooperative activities of CSLD2, CSLD3, and CSLD5 are required for normal Arabidopsis development. *Molecular Plant*, 4, 1024–1037. Available from: https://doi.org/10.1093/mp/ssr02

Zhang, H., Liu, X.-L., Zhang, R.-X., Yuan, H.-Y., Wang, M.-M., Yang, H.-Y. et al. (2017) Root damage under alkaline stress is associated with reactive oxygen species accumulation in rice (*Oryza sativa* L.). *Frontiers of Plant Science*, 8, 1–12. Available from: https://doi.org/10.3389/fpls.2017.01580

Zou, C.L., Wang, Y.B., Wang, B., Liu, D., Liu, L. & Li, C.F. (2021) Effects of alkali stress on dry matter accumulation, root morphology, ion balance, free polyamines, and organic acids of sugar beet. *Acta Physiologiae Plantarum*, 43, 13. Available from: https://doi.org/10.1007/s11738-020-03194-x

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Busoms, S., Pérez-Martín, L., Terés, J., Huang, X.-Y., Yant, L., Tolrà, R. et al. (2023) Combined genomics to discover genes associated with tolerance to soil carbonate. *Plant, Cell & Environment*, 1–13. https://doi.org/10.1111/pce.14691