



**HAL**  
open science

## Characterization of a major QTL for sodium accumulation in tomato shoot

Alexandre Héreil, M. Guillaume, R. Duboscq, Y. Carretero, E. Pelpoir, F. Bitton, C. Giraud, R. Karlova, C. Testerink, Rebecca Stevens, et al.

► **To cite this version:**

Alexandre Héreil, M. Guillaume, R. Duboscq, Y. Carretero, E. Pelpoir, et al.. Characterization of a major QTL for sodium accumulation in tomato shoot. 2024. hal-04617354

**HAL Id: hal-04617354**

**<https://hal.inrae.fr/hal-04617354>**

Preprint submitted on 19 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Characterization of a major QTL for sodium accumulation in tomato shoot

Héreil A.<sup>a</sup>, Guillaume M.<sup>b</sup>, Duboscq R.<sup>a</sup>, Carretero Y.<sup>a</sup>, Pelpoir E.<sup>a</sup>, Bitton F.<sup>a</sup>, Giraud C.<sup>c</sup>, Karlova R.<sup>d</sup>, Testerink C.<sup>d</sup>, Stevens R.<sup>a</sup> and Causse M.<sup>a</sup>

<sup>a</sup>INRAE, UR1052 GAFL, 84143 Montfavet, France | <sup>b</sup>GAUTIER Semences, Route d'Avignon, 13630 Eyragues, France | <sup>c</sup>INRAE, UE A2M, 84143 Montfavet, France | <sup>d</sup>Laboratory of Plant Physiology, Wageningen University, 6700 AA Wageningen, The Netherlands

**Correspondence:** Mathilde Causse [mathilde.causse@inrae.fr](mailto:mathilde.causse@inrae.fr) (ORCID : 0000-0002-0407-4985)

**Summary statement :** The genetic diversity of sodium accumulation in tomato shoot under salt stress is associated with an expression QTL of *SIHKT1.2*. This gene, already identified as involved in tomato salinity tolerance, has been differentially selected during domestication and is of interest for the development of salinity-tolerant varieties.

## Abstract

Soil salinity is a serious concern for tomato culture, affecting both yield and quality parameters. Although some genes involved in tomato salt tolerance have been identified, their genetic diversity has been rarely studied. In the present study, we assessed salt tolerance-related traits at juvenile and adult stages in a large core collection and identified salt tolerance QTLs by genome-wide association study (GWAS). The results suggested that a major QTL is involved in leaf sodium accumulation at both physiological stages. We were able to identify the underlying candidate gene, a well-known sodium transporter, called *SIHKT1.2*. We showed that an eQTL for the expression of this gene colocalized with the sodium content QTL. A polymorphism putatively responsible for its variation was identified in the gene promoter. Finally, to extend the applicability of these results, we carried out the same analysis on a test-cross panel composed of the core collection crossed with a distant line. The results indicated that the identified QTL retained its functional impact even in a hybrid genetic context: this paves the way for its use in breeding programs aimed at improving salinity tolerance in tomato cultivars.

## Key words

Salt stress, GWAS, *Solanum lycopersicum*, HKT1, breeding

## 1. Introduction

Soil salinity is a major agricultural challenge: excess soil salinity occurs under all climatic conditions and results from both natural events and human-induced actions. Saline soils occur in arid and semi-arid regions where rainfall is insufficient to meet water requirements of crops and leach mineral salts away from the root zone. Precise world statistics on the global extent of soil salinisation does not exist, but many sources claim that between 25% and 30% of irrigated lands are affected by high salt and are commercially unproductive. Furthermore, salinity is expected to increase in many areas of the world due to global warming [1]. From an economic point of view, the cost of salt-induced soil degradation is difficult to estimate, as it depends on many factors, such as the nature of crops, the extent and severity of soil degradation, the effectiveness of soil and water management practices, but it is estimated at an average cost of \$441 ha<sup>-1</sup> [2].

From a biological perspective, the adverse effects of salinity on plant growth arise primarily because of the increase in osmotic pressure, triggered by excessive Na<sup>+</sup> accumulation, causing a subsequent decline in the soil's water potential, inhibiting root water absorption and consequently diminishing water availability. Beyond osmotic stress, an altered cytoplasmic Na<sup>+</sup>/K<sup>+</sup> balance impairs the functionality of a myriad of K<sup>+</sup>-dependent enzymes [3]. Finally, the cumulative effects of salt-induced osmotic stress and ionic disequilibrium compromise photosynthesis, perturbing the plant's energy metabolism. These conditions may catalyse the generation of reactive oxygen species (ROS), instigating oxidative degradation within cellular structures [4]. Salinity tolerance in plants is not a ubiquitous trait but varies between and within species. This variability provides fertile ground for the identification of specific genetic loci and natural variations which can play an essential role in improving salt stress tolerance. The first studies, carried out mainly on the model plant *Arabidopsis thaliana*, identified a plethora of genes involved in salt stress response [3]. However, this species exhibits relatively low salt tolerance and is not easily phenotyped for relevant agronomic traits such as yield. In contrast, tomato (*Solanum lycopersicum*) stands out as an agriculturally valuable species, and its role as a model organism for various biological processes, including abiotic stress tolerance, has been gaining momentum. Native to western South America, numerous wild relatives of tomato, such as *S. pennellii*, *S. chilense*, and *S. cheesmaniae*, are adapted to saline environments. Notably, these species have demonstrated higher salinity tolerance than *S. lycopersicum* [5–7], indicating a promising avenue for identifying salt resistance genes and quantitative trait loci (QTLs).

Studies of the genetic architecture underlying salt tolerance in tomato have mainly relied on classical methodologies, notably QTL mapping on biparental populations [8,9,10]. Such approaches inherently limit the discovery of variants present in only one or a limited set of wild donor genotypes. However, by embracing the potential of Genome-Wide Association Studies (GWAS) using a diversity panel encompassing a large number of accessions, a more comprehensive understanding of the genetic landscape can be obtained. Such an approach was successfully implemented to analyse the root Na<sup>+</sup>/K<sup>+</sup>

ratio in a diverse panel of tomato genotypes [11]. In this study, the authors identified and validated a HAK/KUP/KT transporter, named *S/HAK20*, which plays a pivotal role in the  $\text{Na}^+/\text{K}^+$  ratio in roots. Furthermore, natural variation in the components of the SOS-pathway, specifically *S/SOS1* and *S/SOS2*, has been associated with the root  $\text{Na}^+/\text{K}^+$  ratio in tomato within the same panel [12,13].

In the present study, we introduced a novel GWAS for tomato salt tolerance, with an emphasis on mineral content traits, primarily  $\text{Na}^+$  and  $\text{K}^+$ . These elements were analysed in different organs at different physiological stages. We used two distinct but related panels : an inbred core collection and an F1 test-cross panel derived from the former [14]. This second panel aimed to evaluate the applicability and relevance of the QTLs identified at the hybrid level. The first part of this work is dedicated to trait analysis, by elucidating their significance in salt tolerance. Then, a GWAS was carried out, revealing several potential candidate genes. A major QTL hotspot associated with  $\text{Na}^+$  content was identified on chromosome 7, with two underlying candidate genes previously studied, *S/HKT1.1* and *S/HKT1.2* [15]. Here we showed that this association also co-localised with a root expression QTL for *S/HKT1.2*. Finally, The analysis of the evolution of the haplotypes of this gene suggested that domestication may have resulted in the fixation of a more highly expressed variant of *S/HKT1.2*.

## 2. Materials and methods

### 2.1. Plant materials and growth conditions

Two panels of tomato genotypes were used to analyse the genetic response to salt stress. The first panel was composed of 166 homozygous lines (mainly cherry tomato accessions and 15 *S. pimpinellifolium* accessions) from a core collection previously characterised in other studies [16,17]. We will refer to this first panel as CC (Core Collection). The second panel was composed of 144 F1 hybrids obtained by crossing 144 lines of the CC with a large-fruited line Ferum-TMV (FTMV) used as a tester. We will refer to this panel as TC (Test Cross). The list of accessions is provided in [Supplementary Table 1](#).

### 2.2. Growth conditions and measurement of biomass/agronomic traits

The CC panel has been grown in two experiments, one in hydroponic condition in a phytotron where plants were grown for 20 days, another in greenhouse under agronomic conditions. The TC panel has been studied only in greenhouse condition.

#### Phytotron trial

Temperatures were set at 24°C (day) /18°C (night), with a photoperiod of 16/8h and 65% relative humidity. Light was supplied by cool white fluorescent lamps at a photosynthetic photon flux density of 165  $\mu\text{mol m}^{-2} \text{s}^{-1}$  at the leaf level. The hydroponic setup consisted of opaque plastic containers (50 L) covered by a drilled opaque lid with 20 holes ( $d=1.43$  cm) at regular intervals. A 2-cm-long rockwool plug holder was placed in each hole. Containers were filled with a standard nutrient solution (Liquoplant

Rose, Plantin, Courthézon, France) diluted to 4% (2.0 mS) and adjusted to pH 5.7 using hydrochloric acid. The solution was renewed every 5 days. The seeds were sown directly in rockwool plug holders and germinated on a heated tray (25°C). At 7 days after germination (d.a.g), seedlings were transferred to the hydroponic system. The salt-stress condition was initiated at 10 d.a.g by adding 60 mM NaCl to the nutrient solution. The salinity was then increased to 120 mM at 12 d.a.g. Another group of plants was used as controls (without any addition of NaCl). The design of the experiment was structured as follows: It comprised 24 containers, each holding 32 plants. Among these containers, 12 were used for control condition, while the other 12 were used for stress condition. The plants were arranged in a random but symmetrical layout, which facilitated simultaneous sampling of the two sets of plants and enabled plasticity traits to be calculated by minimising spatio-temporal effects. Also, two replicates *per* genotype were placed side by side in each container. The experiment was carried out twice to obtain between 4 and 6 biological replicates per genotype for each condition. Plant harvesting started at 20 d.a.g and was carried out over a 3 h time window. For each replicated genotype grown in the same container, whole roots from two plants were weighted, washed with deionised water, pooled, snap-frozen in liquid nitrogen, and stored at -80°C for later analysis. The shoot parts were also weighted, pooled, washed with deionised water and dried in a forced-air oven (70°C) for 48h, leading to fresh weight (FW) and dry matter content. (DMC).

### **CC greenhouse trial**

The CC collection was grown from March to July 2021 in a glasshouse at INRAE GAFL, Avignon, under soilless conditions in rockwool. Plants were irrigated with a standard fertigation solution (Liquoplant Rose, Plantin, Courthézon, France) where NaCl was added, resulting in a final EC of approximately 8 dS.m<sup>-1</sup>. In order to gradually adapt plants to salt stress, the introduction of NaCl into the fertigation solution began one week after the plantation. Then, 50 mmol L<sup>-1</sup> of NaCl was added for another week before increasing to the final level (80 mM). For each accession of the CC collection, the experiment was conducted with three biological replicates. Plant height (in cm), stem diameter (in mm) and leaf length (in cm) under the 4th truss were measured. Red-ripe fruits were harvested twice a week to assess mean fruit weight.

### **TC greenhouse trial**

The TC panel was grown in a greenhouse at Gautier Semences, Eyragues, France, twice in 2021, first in spring under control condition, then in autumn under salt stress, using the same protocol as for the CC, but in addition fruit yield was assessed.

## **2.3. Mineral content measurement**

Two fully developed leaves were harvested on each plant above the fifth flower cluster prior to the ripening of the first cluster. Petioles were stripped of their leaf blades and promptly stored at -20°C until analysis. Petiole sap was extracted using an extraction bag and a semi-automated homogenizer

(HOMEX 6; Bioreba, Reinach, Switzerland). The filtered petiole sap obtained was utilized to determine the sodium (Na<sup>+</sup>) and potassium (K<sup>+</sup>) content using an ion-selective electrode meter specific for each ion following the manufacturer's instructions (LAQUAtwin Na-11 and K-11; Horiba Instruments Inc., Irvine, CA). For the evaluation of Na<sup>+</sup> and K<sup>+</sup> levels in fruits, a minimum of five red ripe fruits were collected from the third to the sixth truss and pooled for each replication. Subsequently, Na<sup>+</sup> and K<sup>+</sup> content was assessed on crushed fruit pericarps using the same device. For ionome measurement in young plants, dried leaf samples were ground, weighted, and digested with nitric acid and hydrogen peroxide using a block digester. Analysis of Na, K, Mg, Mn, Ca, Fe, and Cu levels was conducted via MP-AES (Microwave Plasma Atomic Emission Spectrometry) (Agilent 4200MP-AES, Santa Clara, CA) at the BPMP laboratory (Montpellier, France).

## 2.4. Estimation of genotype means and heritabilities

Prior to association studies, phenotypic data were aggregated: for phenotypes whose repetitions were carried out at different times (i.e., all fruit derived phenotypes: Na<sup>+</sup> and K<sup>+</sup> fruit content), best linear unbiased predictors (BLUPs) were estimated using the R lme4 package to correct for the harvest-day effect:

$$y_{ik} = \mu + D_k + G_i + \varepsilon_{ik}$$

Where  $\mu$  is the overall mean,  $G_i$  represents the  $i^{\text{th}}$  genotype effect,  $D_k$  represents the  $k^{\text{th}}$  harvest-day effect and  $\varepsilon_{ik}$  the  $ik^{\text{th}}$  residual variation with  $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ .

Biomass traits from the greenhouse trial (measured on a single day) were aggregated by a simple mean. Biomass and leaf mineral contents measured in the phytotron were also corrected using a mixed effect model:

$$y_{ij} = \mu + G_i + R_j + S_b + RS_{bj} + \varepsilon_{ij}$$

Where  $R_j$  represents the  $j^{\text{th}}$  replicate effect,  $S_b$  is the  $b^{\text{th}}$  block effect,  $RS_{bj}$  is the  $bj^{\text{th}}$  block  $\times$  replicate interaction effect with  $RS_{bj} \sim N(0, \sigma_{RS}^2)$ .

For traits measured on seedlings under control and salt stress conditions, a random-effect analysis of variance was conducted on the whole population to test for genotype (G), treatment (T), and their interaction (G $\times$ T) effects with the following model using R/lme4:

$$y_{ij} = \mu + G_i + T_j + GT_{ij} + \varepsilon_{ij}$$

Where  $T_j$  is the effect of the  $j^{\text{th}}$  treatment, and  $GT_{ij}$  is the  $ij^{\text{th}}$  genotype  $\times$  treatment interaction effect with  $GT_{bij} \sim N(0, \sigma_{GT}^2)$ . Broad-sense heritability was calculated from the above model according to [18]:

$$H^2 = 1 - \frac{\overline{\vartheta}_\Delta^{BLUP}}{2\sigma_G^2}$$

Where  $\sigma_G^2$  is the genotype variance,  $\overline{\vartheta}_\Delta^{BLUP}$  the average standard error of the genotypic BLUPs.

Phenotypic plasticity traits were computed as the ratio of stress to control values. This index was then used as a phenotype *per se* for QTL analysis. We will thus refer to (C), (S) and (D) for traits measured in control, stress and plasticity conditions. When all the genotypes were taken into account, the average effect of the stress was reported as the mean relative difference between control and stress conditions and converted into the percentage increase or decrease due to the stress. The significance of the treatment effect was then calculated by using a likelihood ratio test from R/lmtest, by comparing the goodness-of-fit between two models, one considering the treatment effect and a second one not taking it into account. Prior to GWAS analyses, we applied a Box-Cox transformation of the means calculated from replicates to correct for heteroscedasticity and non-normality of error terms [19]. Trait means of all accessions and treatments are presented in [Supplementary Tables 2 and 3](#).

## 2.5. Genotypic data and GWAS analysis

To detect associations between genotypic values and genomic markers, GWAS was performed using a set of 6,251,927 SNPs from genome resequencing. All procedures used for alignment, SNP calling and post processing are described in [20]. SNP positions refer to the Heinz\_1706 v.4.0.0 reference genome. A univariate GWAS was performed by implementing the following linear mixed model in R/GENESIS [21]:

$$y = X\beta + Zu + \varepsilon$$

where  $y$  is the vector of phenotypic means,  $X$  is the molecular marker score matrix,  $\beta$  is the vector of marker effects,  $Z$  is an incidence matrix,  $u$  is the vector of random background polygenic effects with variance  $\sigma_u^2 = K \sigma_G^2$  (where  $K$  is the kinship matrix, and  $\sigma_G^2$  is the genetic variance), and  $\varepsilon$  is the vector of residuals. The null model was fitted using the `NullModel()` function using only the fixed-effect covariates. We included the first three eigenvectors estimated from the PCA-based overall genotypic matrix using PLINK 2.0 [22]. Single-variant association tests were performed with the `assocTestSingle()` function using the Average Information REML (AIREML) procedure to estimate the variance components and score statistics. We then used an arbitrary genome-wide significance thresholds of  $\text{LOD} = 6$  for SNP significance. To estimate the proportion of variance explained (PVE) by lead SNP, we used the following formula proposed by [23] using outputs from R/GENESIS:

$$\text{PVE} = \frac{2 \hat{\beta}^2 \text{MAF} (1 - \text{MAF})}{2 \hat{\beta}^2 \text{MAF} (1 - \text{MAF}) + (\text{se}(\hat{\beta}^2))^2 2n \text{MAF}(1 - \text{MAF})}$$

Where  $\hat{\beta}$  is the effect size,  $\text{se}(\hat{\beta}^2)$  is the standard error for the effect size,  $\text{MAF}$  the minor allele frequency and  $n$  the sample size.



## 2.6. Haplotype and diversity analysis

Due to the low number of *S. pimpinellifolium* accessions in the CC panel, we merged this panel with the Varitome core collection [24] to have a broader view of tomato genetic diversity. This led to a panel of 331 resequenced accessions. Using this combined panel, we performed different haplotype analyses. A haplotype heatmap was generated using the R/pheatmap, using Ward distance for clustering the genotypes (rows). Cluster number for *k*-means was calculated according to the elbow method using a screen plot with the identification of the optimal number of clusters when the total intra-cluster variation was minimized. The multiple mean comparison to test significant differences among haplotype clusters was conducted using a linear model. We used R/emmeans to calculate the p-value of pairwise comparisons among clusters and R/multcompView to display the Tukey test, fixing the significance threshold at 0.05. To investigate if *SHKT1.2* was under selection we computed allele-specific extended haplotype homozygosity (EHH). This analysis was conducted both upstream and downstream of the most significant SNP in *SHKT1.2* promoter sequence using R/rehh. For visualisation, EHH values were integrated across their respective genomic positions to yield an 'integrated EHH' (iHH) value for each allele [25].

To identify the conserved regions of *SHKT1.2* promoter within the tomato clade, we performed a sequence alignment involving ten *de novo* assembled genomes of wild tomato relatives - *S. galapagense*, *S. pimpinellifolium*, *S. chmielewskii*, *S. neorickii*, *S. corneliomulleri*, *S. peruvianum*, *S. chilense*, *S. pennelli*, *S. habrochaites*, and *S. lycopersicoides*. These genome sequences were obtained through Hi-C sequencing [26]. For the alignment, the promoter sequence of the *SHKT1.2* gene was extracted from the SL4.0 chromosome 7 sequence between positions 5,126,736 and 5,143,883. Then, the extracted sequence was subjected to BLAST analysis against the ten wild tomato genomes using NOBLAST [27]. Finally, the sequences corresponding to these homologous regions were aligned with the reference SL4.0 sequence using the Needleman-Wunsch pairwise alignment tool, EMBOSS Needle [28].

## 2.7. RNA extraction and RT-qPCR

A total of 92 accessions out of 166 were selected for expression analysis ([Supplementary Table 4](#)). For each genotype grown in hydroponic condition, whole roots from two seedlings (0.8 to 16 g) grown in the same container were weighted, washed with deionised water, pooled, snap-frozen in liquid nitrogen, and stored at  $-80^{\circ}\text{C}$  for later analysis. RNA extraction was performed using the “RNeasy Plant Mini Kit” (Qiagen Sciences, Germantown, MD, USA) following the manufacturer’s protocol, with “On-column DNase I Digestion Set (Sigma-Aldrich, Saint-Quentin Fallavier, France) to remove any remaining genomic DNA. The amount of RNA was quantified using a Qubit Fluorometer 3.0 and the “Qubit RNA Broad Range Assay Kit” (Invitrogen, Carlsbad, CA, USA). RNA purity (260/230 nm and 260/230 nm ratios) was assessed using the NanoDrop 1000 Spectrophotometer (Thermo Fisher,



Waltham, MA, USA); all 260/280 nm ratios were comprised between 1.8 and 2.2 and all 260/230 nm ratios were beyond 2.0. Sample integrity was verified on a Bioanalyzer 2100 (Agilent Technologies, Les Ulis, France) with the “RNA 6000 Nano Kit”; all RNA integrity numbers (RIN) were beyond 7. RNA was then reverse transcribed with the “GoScript Reverse Transcription Mix, Oligo(dT)” (Promega, Madison, WI, USA) and qRT-PCR experiments were performed with the “Brilliant III Ultra-Fast SYBR Green QPCR Master Mix” (Agilent Technologies) according to the manufacturer's specifications. The primers used for the *S/HKT1.1* (Solyc07g014690) were 5'- TCTAGCCCAAGAACTCAAAT -3' and 5'- CTAATGTTACAACCTCAAGGAATT -3', and those for *S/HKT1.2* (Solyc07g014680) were 5'- TGAGCTAGGGAATGTAATAAACG -3' and 5'- AGAGAGAACTAACGATGAACC -3', according to [15]. The genes Solyc06g005360 (5'-AAGGGCTTTCGCTTCATAGT-3' and 5'-AACTTTCAGCTGGCTACCAA-3') and Solyc10g006100 (5'-CCAATAAAGCGCTGCCACAAC-3' and 5'- TGGTCCTTGTGTGCTTACTGGC-3') were used as reference genes. Relative quantification data analysis was determined with the  $\Delta\Delta C_t$  method.

### 3. Results

#### 3.1. Impact of salt stress at juvenile and adult stages

Mean, coefficient of variation as well as percent reduction in mean due to salt stress and variance decomposition of the studied traits are shown in [Table 1](#). The impact of salt stress was only characterised for traits measured at juvenile stage (20 dag) as, in greenhouse, the CC panel was grown only in stress condition. We observed a large variation in salt response across the accessions grown in the phytotron: The reduction in total biomass due to salt stress ranged from 13% to 86%. Notably, the disparity in biomass was more pronounced in the shoot, with a decrease of 61.1%, compared to the root, which was reduced by 24.3%, on average. This difference resulted in an elevated shoot-to-root ratio between the two treatments. The variance decomposition of biomass traits remained relatively consistent, with the exception of root biomass, which exhibited an increased residual variance of 40% ([Table 1](#)). Mineral compositions, however, displayed diverse patterns of variation under salinity stress. Specifically, macro-element concentrations, including calcium (-36%), potassium (-48%), magnesium (-28%), and manganese (-26%), tended to decrease. In contrast, the concentrations of microelements, notably copper and iron, increased by 23% and 74%, respectively. Zinc was the only mineral with no significant variation in concentration under saline conditions. As expected, the sodium content difference between the two treatments was substantial, recording an increase of 2004%. The observed variance in sodium was thus predominantly governed by the treatment effect, accounting for 90.5% of the total variance. Heritability estimates spanned from moderate to high values: for the control conditions, they ranged between 0.55 to 0.73, while under stress conditions, they fluctuated between 0.49 and 0.87. The potassium and sodium contents under saline conditions showed high heritabilities of 0.79 and 0.87, respectively.

The traits evaluated in greenhouse were only characterised under salt stress condition. Broad-sense heritability estimates systematically exceeded those observed during the juvenile stage. This increased heritability could be attributed to two main factors: firstly, the pronounced phenotypic diversity, highlighted by the coefficient of variation reaching 71% (fruit sodium content), and secondly, the experimental design (described in the 'Materials and Methods' section) which effectively alleviated potential spatial effects between genotypes.

### **3.2. Correlations of mineral contents with salt tolerance at different stages and in different organs**

Based on the correlation and cluster analysis shown in [Figure 1](#), distinct groups of correlated traits were identified. All traits relative to Na<sup>+</sup> content were clustered within the same group (cluster 3), suggesting a common genetic basis for sodium homeostasis in different organs and at different growth stages. Some traits, such as petiole Na<sup>+</sup> content, were highly correlated with fruit Na<sup>+</sup> content ( $r = 0.78$ ) under stress. It is noteworthy that even with trace concentrations of Na<sup>+</sup> content in the control condition, there was a substantial correlation between the Na<sup>+</sup> concentrations of the two treatments ( $r = 0.63$ ). The traits of cluster 3 were all negatively correlated with the traits in cluster 1, which grouped cation contents (Ca<sup>2+</sup>, Mg<sup>2+</sup>, K<sup>+</sup>) under stress condition and plasticity. Interestingly, the levels of these minerals measured under control conditions (cluster 4) were not correlated with their counterparts under salt stress. Finally, biomass-related traits measured under stress conditions (cluster 1) were correlated only when measured at the same physiological stage. At the adult stage, only the stem diameter exhibited a significant negative correlation with salt content, leaf and stem length being not correlated with Na<sup>+</sup> content. Stem diameter was also highly correlated with petiole potassium content.

### **3.3. GWAS analysis reveals a QTL hotspot on chromosome 7**

In order to further understand the genetic basis underlying the observed difference in salt tolerance and sodium accumulation, we performed a single-trait GWAS on 166 small-fruited accessions for 22 salt tolerance-related traits (selected based on heritability values and variance decomposition; [Figure 2.a](#)). We found a total of 1032 SNP that were above the LOD-threshold of 6. Neighbouring SNPs were mostly in linkage disequilibrium. Thus, for subsequent analysis we clustered all markers in linkage disequilibrium, reducing the number of independent QTLs to 14 clusters ([Supplementary Table 5](#)). A strong association signal for Na<sup>+</sup> content in leaf (at young stage), fruit and petiole (at adult stage) was identified on chromosome 7, with the lowest p-value of all the traits (LOD >10). In particular, the marker found to be significant for most of the traits was SSL4.0CH07\_4555305 ([Figure 2.b](#)). In this region, more than 7000 SNP markers were significantly associated with at least 11 traits, in the interval ranging from 4,415,544 to 5,127,764 base pairs (bp). This association was found exclusively for sodium levels

in each organ type (seedling leaf, mature petiole sap, fruit pericarp), under control and salt stress conditions. This QTL was also found for potassium levels, but exclusively under salt stress conditions. For the traits leaf Na<sup>+</sup> (S) and petiole Na<sup>+</sup> (S), this region covered 50 genes, including two High-affinity Potassium Transporters (HKT) genes (*S/HKT1.2* (Solyc07g014680) and *S/HKT1.1* (Solyc07g014690)), which were both previously identified as playing a pivotal role in sodium homeostasis in several species, including tomato [15,29]. This QTL also displayed long-range linkage disequilibrium across the chromosome, suggesting that population admixture, epistatic selection, or other evolutionary forces were at work [30]. Beyond this major QTL, few other associations were detected on other chromosomes (**Supplementary Table 5**). Leaf sodium content in stress condition revealed the largest number of associations, with 11 regions identified on 6 chromosomes. We detail below only the associations detected for at least two traits. A significant association was found for leaf sodium content (S), leaf sodium content (D) and leaf Na/K (D) on chromosome 4 (SSL4.0CH04\_63269805). The twelve significant SNP for this association were found in one gene, Solyc04g015850, of unknown function but harbouring transmembrane protein characteristics. In this region, at 93 kb, Solyc04g015990 was an interesting candidate gene expressed in roots (**Supplementary Figure 1**). This gene encodes a Cation/H(+) antiporter found to be down regulated by salt stress (fold change in roots: -13.43) [31]. This gene has also been suggested as a candidate for a QTL hotspot for multiple cation xylem content in a RIL population derived from a cross between *S. lycopersicum* and *S. pimpinellifolium* [32]. Another genetic marker (SL4.0CH03\_53341434), identified on chromosome 3, exhibited a significant association with four traits related to Na<sup>+</sup> and K<sup>+</sup> concentrations at the seedling stage. The most plausible candidate gene responsible for these associations was Solyc03g096670, which encodes a Type 2C protein phosphatase (PP2C). The functional relevance of PP2C in salt tolerance is well documented: especially, its overexpression improved salt tolerance by regulating shoot Na<sup>+</sup> extrusion [33] or by regulating oxidative stress-related genes [34]. Also, RNA-seq studies revealed that Solyc03g096670 was significantly upregulated under salt stress conditions, both at the root level (7-fold change) [35] and in the leaves (4.8-fold change) [36]. These expression patterns strongly suggest that Solyc03g096670 is involved in the plant adaptive response to salt stress. From the perspective of genetic diversity, substantial variations in gene expression at the root level have been reported. The dataset from Alonge et al., [37] revealed up to a 22-fold change in expression levels between extreme genotypes. This study focused on the impact of structural variants on gene expression and identified a significant association between a deletion located 5 kb downstream of Solyc03g096670 and its expression levels. Such an association suggests the presence of a cis-expression QTL. This hypothesis is further supported by the detection of a Solyc03g096670 eQTL in a panel of introgression lines (ILs) derived from a cross between *S. lycopersicum* 'M82' and *S. pennellii*, reported by Ranjan et al., [38]. All these elements suggest that Solyc03g096670 could also be an expression QTL that deserves further analysis. We did not detect any candidate gene for plant weight plasticity at the given threshold (LOD > 6). However, the most

significant peak on chromosome 8 (LOD = 5.58) was located at the level of a Late embryogenesis abundant 3 family protein (Solyc08g074720).

### 3.4. Candidate gene mining: a focus on HKT genes

The leading SNPs within the QTL on chromosome 7 were close to two HKT1 genes (*S/HKT1.1* and *S/HKT1.2*). Both have been previously identified as playing a pivotal role in sodium homeostasis in several species, including tomato [15,29]. A first polymorphism analysis revealed that neither of the two HKT1 genes exhibited polymorphisms in the coding sequence that could account for the QTL on chromosome 7. This led us to propose that the underlying polymorphism might be a differential expression of one of these genes. To validate this hypothesis, we measured the expression levels of both HKT1 genes in root samples exposed to salt stress in phytotron, given that both genes are predominantly expressed in roots (**Supplementary Figure 1**). Our analysis revealed a substantial difference in *S/HKT1.2* expression between accessions harbouring the reference and alternative alleles. A major association was identified for *S/HKT1.2*, colocalizing with the previously mentioned QTL on chromosome 7 (**Figure 3.a**). This QTL accounted for a substantial portion of the variance in *S/HKT1.2* gene expression, explaining approximately 50% of the variance of this gene expression (**Figure 3.b**; t-test p-value =  $3.6 \cdot 10^{-11}$ ). Conversely, no significant differences were detected in *S/HKT1.1* expression. Furthermore, *S/HKT1.2* expression was negatively correlated with leaf sodium content ( $r = -0.63$ , **Figure 3.d**) and positively correlated with potassium content (**Figure 3.e**;  $r = 0.5$ ). No significant correlation was observed for *S/HKT1.1*.

As we found a colocalisation between the sodium content QTL and the eQTL for *S/HKT1.2* expression, we searched for polymorphisms that could explain this difference in gene expression. We thus first carried out a local reanalysis of *S/HKT1.2* expression association, using in addition to SNP, indels and short structural variants detected by the ANNOVAR algorithm (**Figure 4**). In the promoter region, we detected few significant polymorphisms, including three SNPs and one short indel (T/TC) at 1,564 bp from the initiation start. We also compared the Heinz1706 SL4.0 reference genome to the one of BGV007931, a *de novo* genome assembly of a Mexican *S.l. cerasiforme* tomato with a level of *S/HKT1.2* expression close to zero, compared to the reference genome accession Heinz1706 [37]. The comparison of both sequences revealed additional interesting polymorphisms, including two short tandem repeats at 2,491 and 2,680 bp from the initiation site and of lengths 10 and 43 bp, respectively (**Figure 4**). As these two polymorphisms could not be captured with our short-read sequencing, we did not include them in the GWAS analysis. Finally, we sought to elucidate the potential impact of the mutations that target the HKT1.2 promoter sequence, focusing on the alteration of promoter motifs. We only considered polymorphisms that were significant in GWAS. We identified four candidate polymorphisms located in the promoter region. Among them, two motifs were predicted to be structurally affected by mutations: a nucleotide substitution at position -972 of the transcription start site

(TSS) and an insertion at position -1,564 bp (**Table 2**). Both markers were in complete linkage disequilibrium ( $r^2 = 1$ ). The insertion at position -1,564 bp is particularly interesting, as the alternative allele disrupts an ABRE (ABscisic acid-Responsive Element) motif. ABRE motifs have already been identified in many HKT1 gene promoters in several species, including *Arabidopsis*, wheat and rice [39]. Furthermore, in *Arabidopsis*, a single ABA-responsive element in the *AtHKT1;1* gene promoter has been shown to be essential for gene expression, particularly in response to sodium chloride (NaCl) and abscisic acid (ABA) [40].

To further contextualise these findings, we investigated whether the candidate polymorphisms were located in evolutionarily conserved regions of the promoter. For this purpose, we aligned the promoter sequences of *S/HKT1.2* homologs of ten wild tomato cross-compatible species – *S. galapagense*, *S. pimpinellifolium*, *S. chmielewskii*, *S. neorickii*, *S. corneliomulleri*, *S. peruvianum*, *S. chilense*, *S. pennellii*, *S. habrochaites* and *S. lycopersicoides*. We then assessed the level of conservation by measuring the allelic frequency of the alternative alleles (**Supplementary Figure 2**). Our analysis revealed that the *S/HKT1.2* promoter sequence was highly conserved up to position -1,800 relative to the TSS in these species. Furthermore, among the four candidate polymorphisms, two (-1,593 and -1,564) were conserved across all species. In particular, the ABRE motif, whose disruptive variant was suspected of being responsible for the differential expression of the *S/HKT1.2* loci, was perfectly conserved in all the species examined.

### 3.5. Diversity and evolutionary history of HKT1.2

To gain insights into the evolutionary history of HKT1.2 alleles identified in our study, we performed a large-scale genomic analysis. In order to construct a global perspective of the distribution and demographic histories of the HKT1.2 alleles, we integrated into our core collection data an additional set of 163 genome sequences sourced from the Varitome project [24]. This expanded dataset includes genomes from *S. pimpinellifolium* (SP), *S. lycopersicum* var. *cerasiforme* (SLC), and *S. lycopersicum* var. *lycopersicum* (SLL), thereby capturing the genetic diversity at the epicentre of tomato origin of domestication. Interestingly, the allele associated with a high HKT1.2 expression is absent from the *S. pimpinellifolium* background. Its increasing prevalence in domesticated forms of tomato (piecharts are arranged in order according to the tomato domestication model proposed by [41]) suggested that the highly expressed HKT1.2 allele underwent selection during the domestication process (**Supplementary figure 3.a**). To corroborate these findings, we also examined the Extended Haplotype Homozygosity (EHH) decay patterns surrounding the HKT1.2 gene. Our analysis suggested that the 'cerasiforme' haplotype underwent artificial selection during the course of tomato domestication. Specifically, the derived haplotype (more frequent in domesticated forms) exhibited a slow decay in EHH values, corresponding to a recent selection event affecting the polymorphisms in proximity to the HKT1.2 derived allele (**Supplementary Figure 3.b**).

### 3.6. Detection of Na<sup>+</sup> content QTL under hybrid background

To assess the potential of the QTLs detected on the CC panel containing only inbred lines for breeding applications, the greenhouse trials were replicated using an hybrid population (Test-cross: TC) obtained by crossing each line of the core collection with a large-fruited tomato [14]. As expected from a test-cross panel, all the traits measured showed a significantly reduced coefficient of variation (**Supplementary Table 3**). Analysis of correlations between mineral traits measured on the CC and TC panels showed that all Na<sup>+</sup> and K<sup>+</sup> content traits were correlated with each other, ranging from  $r=0.46$  for fruit K<sup>+</sup> to  $r=0.61$  for Na<sup>+</sup> petiole content (**Supplementary Figure 4**). For Na<sup>+</sup> leaf petiole content, fruit Na<sup>+</sup> content and K<sup>+</sup> petiole content, the major QTL on chromosome 7 was also detected (**Figure 5**). The lead SNP was not identical to that identified in the core collection, probably because the composition of the two panels was slightly different (**Supplementary Table 3**), causing an imbalance in allele frequencies. Also, the percentage of variance explained was significantly lower than for QTL identified using the core collection (PVE = 0.11 for petiole Na<sup>+</sup> content in the TC panel). These results suggested that HKT1.2 QTL also had an effect at the heterozygous state, making it interesting for F1 tomato breeding. We also detected QTLs specific to the test-cross panel which had putative candidate genes. For instance, for petiole K<sup>+</sup> on chromosome 1, the lead marker in position 77,939,937 was near a K<sup>+</sup> efflux antiporter (Solyc01g094290) mostly expressed in leaf tissues.

## 4. Discussion

### 4.1. Is sodium content a reliable marker of tomato salt tolerance?

Some wild relatives of tomato have higher salt tolerance than cultivated tomato and can therefore provide alleles for genetic improvement of salinity tolerance [6,11]. This requires the identification of interesting germplasm and traits implicated in stress tolerance before the identification of QTLs. Salt tolerance is usually assessed as the percent biomass production in saline versus control conditions. However, evaluating salinity tolerance strictly on the basis of biomass differential offers limited understanding of the underlying physiological mechanisms. A more fruitful approach lies in investigating the genetic control of specific downstream traits that contribute collectively to the overall salt tolerance. Among them, the one that is undoubtedly most described is sodium exclusion. In fact, in several species, a negative correlation between leaf sodium content and salinity tolerance have been described [42,43]. In tomato, whether or not this relationship exists is unclear : several studies didn't find any correlation between foliar sodium content and salinity tolerance [44,45]. In our study, we didn't find any significant correlation between biomass-based salt tolerance and shoot sodium content at the juvenile stage. However, at the adult stage, we observed a negative correlation between sodium content and biomass proxies (stem diameter, leaf length and plant height). Thus, the relationship between leaf



Na<sup>+</sup> content and salt tolerance might be detected only at mature stage. This effect has been documented in monocotyledons: some authors showed that maize and wheat cultivars with different salt tolerance can grow at the same rate for two to four weeks, before revealing a difference in growth kinetics [4]. The kinetics of ion build-up in shoots to toxic concentrations is, however, not well documented in solanaceous species (and is also probably genotype-dependent). However, it is highly likely that the situation is similar in the case of tomato. If this is the case, screening plants for salt tolerance at too early a stage is akin to phenotyping for osmotic stress tolerance. However, as we have seen, genotypic differences in sodium accumulation are detectable at the early stage, even under control conditions (without exogenous addition of sodium to the fertigation solution).

Still, the relevance of leaf sodium content as a tolerance trait *per se* is open to debate. For instance, salt-tolerant wild relatives of tomato, such as *S. chilense* and *S. pennellii*, exhibit significantly higher leaf Na<sup>+</sup> content compared to their more sensitive counterpart, *S. lycopersicum* - up to ten-fold higher in the case of *S. chilense* [5,7]. These species seem to have a different strategy from *S. lycopersicum*, maintaining high levels of sodium leaf content, but investing more in sodium sequestration as suggested by an increased expression of tonoplast-localised NHX-type Na<sup>+</sup>/H<sup>+</sup> exchangers compared to their homologues in cultivated tomato [5,7]. In such species, the high accumulation of Na<sup>+</sup> in the leaves acts as a cheap osmoticum to minimise the high energy cost of osmolyte synthesis. This raises the question of whether leaf Na<sup>+</sup> content should be considered as a target trait for breeding salt-tolerant tomato varieties? Our data suggest a positive answer, at least in the genetic background of *S. lycopersicum*; this is further supported by studies showing that knock-down of genes responsible for sodium ion translocation from roots to leaves, in particular *SIHKT1.2* and *SISOS1*, results in increased sensitivity to salt stress [12,46]. However, the introgression of exotic genetic material from 'includer' species (*S. chilense*, *S. pennellii*) could compromise this assertion, a point that remains to be verified.

#### **4.2. Genetic diversity of above-ground sodium content is controlled by a few QTLs**

During GWAS analysis, we identified a limited number of QTLs for Na<sup>+</sup> and K<sup>+</sup> content levels. Notably, excluding the hotspot QTL located on chromosome 7, only a 13 QTL were detected. This finding is consistent with existing literature on the subject. For example, a recent GWAS study conducted on a panel of 250 *S. pimpinellifolium* accessions for Na<sup>+</sup> content in leaves also reported a little number of QTLs. Specifically, only three QTLs were detected and, one of these QTLs corresponded closely to the QTL identified in our own study on chromosome 4 [47]. Similarly, in the Wang study [11], no significant association between markers and sodium or potassium levels at shoot level was found. The QTL found on chromosome 7 is not a novel finding. It was reported by Villalta et al., [10] who analysed leaf and stem sodium content in two different RIL populations derived from the cross of a cultivated tomato accession (SLL) with two wild relatives (*S. pimpinellifolium* and *S. galapagense*). The identification of the underlying candidate genes (*HKT1.1* and *HKT1.2*) was published a few years later by the same team



[15]. Analysis of the knock-down of each HKT1 gene was conducted using two near-isogenic lines (NILs) that were homozygous for either the *S. lycopersicum* allele or the *S. galapagense* allele and the authors showed that the knock-down of HKT1.2 gene from *S. lycopersicum* and *S. galapagense* increased the leaf  $\text{Na}^+/\text{K}^+$  ratio [48] and had a significant negative impact on yield [49]. The effect of HKT1.1, however, is less clear. Indeed Jaime-Pérez, N. et al. (2017) [48] showed that silencing of *S/HKT1.1/ScHKT1.1* resulted in a non-significant effect while Romero-Aranda et al., [49] observed that the knockdown of *ScHKT1.1* resulted in the reduction of the  $\text{Na}^+/\text{K}^+$  ratio (but still no significant effect for *S/HKT1.1*). However, these studies have not been able to determine which gene is at the source of the chromosome 7 QTL.

#### **4.3. The sodium content QTL on chromosome 7 is characterised by a variation in *S/HKT1.2* gene expression**

The search for polymorphisms in the promoter region of *S/HKT1.2* did not reveal any obvious candidate polymorphisms (e.g. large deletions in the promoter) which could explain the variation in gene expression. However, several variants significantly associated with a difference in HKT1.2 expression have been detected in the promoter sequence: for example, an insertion of one base (T/TC) at 900 bp from the initiation start) was significantly associated with the change in gene expression and could be a good candidate modification for functional validation through prime editing. In addition, we found a large deletion in the *de novo* assembly sequence of *S/HKT1.2* corresponding to a tandem repeat (3000 bp). As tandem repeats are known to affect the transcription of neighbouring genes by regulating chromatin structure and composition, this structural variant could be an other interesting candidate modification. Also, in *Arabidopsis*, it was established that a deletion in a tandem repeat sequence approximately 5 kb upstream of *AtHKT1* is responsible for the reduced root expression of *AtHKT1* [50]. Unfortunately, we were not able to verify if this deletion was present in accessions with only low expression as read-alignment failed for most of the accessions due to short read sequencing. From an evolutionary perspective, the absence of this allele in the *S. pimpinellifolium* background and its increasing prevalence in domesticated forms of tomato suggest that the highly-expressed HKT1.2 allele was indeed selected during domestication. However, the evolutionary mechanism behind this process is still unclear, especially as other salinity tolerance QTLs such as *S/HAK12* and *S/SOS1* were lost during domestication [12]. Sodium, at low levels, is beneficial to tomato growth and yield [51], especially under suboptimal potassium supply; we may thus hypothesise that the fixation of an highly expressed *S/HKT1.2* allele could have occurred during the domestication of tomato and was then indirectly selected for growth under optimal conditions.

## **5. Conclusion**

The results of the genome-wide association study offer a comprehensive list of promising QTL and candidate alleles that deserve further exploration to understand the genetic and molecular mechanisms underlying variability in mineral content under salt stress conditions. In particular, we have identified a major QTL located on chromosome 7 colocalising with an eQTL for HKT1.2 gene expression in root, making this gene a prime candidate for further research. A natural variation identified in the *S/HKT1.2* promoter, which has undergone selection during tomato domestication and improvement, holds great promise for enhancing tomato salt tolerance and represents an important target for the development of salt-tolerant tomato germplasm.

### **Data availability**

Raw genome reads corresponding to the accession sequences from the core-collection can be found in NCBI (<https://www.ncbi.nlm.nih.gov/sra>) database under accession number PRJNA1014227. All other data supporting the findings of this study are available in the supplementary data published online.

### **Supplementary Material**

**Supplementary Table 1.** List of the accessions of the diversity panel.

**Supplementary Table 2:** Mean phenotypic values measured under control (C), stress (S) or plasticity (D) condition for the inbred core-collection at the young stage and in stress condition at the adult stage

**Supplementary Table 3:** Mean phenotypic values measured under stress (S) condition for the test-cross core-collection

**Supplementary Table 4:** Mean root expression of *SIHKT1.1* and *SIHKT1.2* (expressed in fold change)

**Supplementary Table 5:** Summary of QTLs detected for salt tolerance-related traits

**Supplementary Figure 1.** Expression profile of candidate genes identified through GWAS in various tissues

**Supplementary Figure 2.** Conservation profile of the HKT1.2 promoter sequence in 9 wild tomato relative species

**Supplementary Figure 3.** QTL frequency in different tomato genetic groups and Haplotype Homozygosity Decay

**Supplementary Figure 4.** Hierarchical clustering heatmap of Pearson correlation coefficients between Na/K content, fruit weight and yield in the test-cross panel

### **Acknowledgement**

We are thankful to Mélanie Andrin, Anaïs Roger, Matilde Van-Ommen and Camille Rabeau for taking care of the plants and their help for plant phenotyping, to the vegetable resources centre (CRBLeg) of GAFL for keeping the seeds of the CC panel available, to the greenhouse staff of UE A2M “Arboriculture et Maraichage Méditerranéens” and to Sandrine Chay (INRAE, UMR BPMP / SAME

platform) for ionic analysis. This work was supported by the ERA-NET SusCrop ROOT (Grant No. 771134) and PRIMA- VEG-ADAPT (Grant No. 01DH19019) projects.

### **Author contribution**

MC and AH : conceptualization, data analyses and writing the manuscript; YC, EP, RD, MG, CG : data acquisition; FB bioinformatic data analysis; MC, RK, CT and RS : funding acquisition.

### **Conflict of Interest**

The authors declare that they have no conflict of interest.

## References

- [1] Hassani, A. et al. (2021). Global predictions of primary soil salinization under changing climate in the 21st century. *Nature Communications*. DOI: 10.1038/s41467-021-26907-3.
- [2] Qadir, M. et al. (2014). Economics of salt-induced land degradation and restoration. *Natural Resources Forum*. DOI: 10.1111/1477-8947.12054.
- [3] Wu, H. et al. (2018). It is not all about sodium: revealing tissue specificity and signalling roles of potassium in plant responses to salt stress. *Plant and Soil*. DOI: 10.1007/s1104-018-3770-y.
- [4] Munns, R. (2002). Comparative physiology of salt and water stress. *Plant, Cell & Environment*. DOI: 10.1046/j.0016-8025.2001.00808.x.
- [5] Albaladejo, I. et al. (2017). Unravelling the strategies used by the wild tomato species *Solanum pennellii* to confront salt stress: From leaf anatomical adaptations to molecular responses. *Environmental and Experimental Botany*. DOI: 10.1016/j.envexpbot.2016.12.003.
- [6] Pailles, Y. et al. (2020). Diverse Traits Contribute to Salinity Tolerance of Wild Tomato Seedlings from the Galapagos Islands1 [OPEN]. *Plant Physiology*. DOI: 10.1104/pp.19.00700.
- [7] Bigot, S. et al. (2022). The Halophyte Species *Solanum chilense* Dun. Maintains Its Reproduction despite Sodium Accumulation in Its Floral Organs. *Plants*. DOI: 10.3390/plants11050672.
- [8] Foolad, M.R. et Jones, R.A. (1993). Mapping salt-tolerance genes in tomato (*Lycopersicon esculentum*) using trait-based marker analysis. *Theoretical and Applied Genetics*. DOI: 10.1007/BF00223763.
- [9] Monforte, A.J. et al. (1997). Salt tolerance in *Lycopersicon* species VI. Genotype-by-salinity interaction in quantitative trait loci detection: constitutive and response QTLs. *Theoretical and Applied Genetics*. DOI: 10.1007/s001220050616.
- [10] Villalta, I. et al. (2008). Genetic analysis of Na<sup>+</sup> and K<sup>+</sup> concentrations in leaf and stem as physiological components of salt tolerance in Tomato. *Theoretical and Applied Genetics*. DOI: 10.1007/s00122-008-0720-8.
- [11] Wang, Z. et al. (2020). Loss of salt tolerance during tomato domestication conferred by variation in a Na<sup>+</sup>/K<sup>+</sup> transporter. *The EMBO Journal*. DOI: 10.15252/embj.2019103256.
- [12] Wang, Z. et al. (2021). Natural variations in *SISOS1* contribute to the loss of salt tolerance during tomato domestication. *Plant Biotechnology Journal*. DOI: 10.1111/pbi.13443.
- [13] Hong, Y. et al. (2023). Natural variation in *SISOS2* promoter hinders salt resistance during tomato domestication. *Horticulture Research*. DOI: 10.1093/hr/uhac244.
- [14] Bineau, E. et al. (2021). Breeding Tomato Hybrids for Flavour: Comparison of GWAS Results Obtained on Lines and F1 Hybrids. *Genes*. DOI: 10.3390/genes12091443.
- [15] Asins, M.J. et al. (2013). Two closely linked tomato HKT coding genes are positional candidates for the major tomato QTL involved in Na<sup>+</sup>/K<sup>+</sup> homeostasis. *Plant, Cell & Environment*. DOI: 10.1111/pce.12051.
- [16] Albert, E. et al. (2016). Association mapping reveals the genetic architecture of tomato response to water deficit: focus on major fruit quality traits. *Journal of Experimental Botany*. DOI: 10.1093/jxb/erw411.
- [17] Bineau, E. et al. (2021). Genetic diversity of tomato response to heat stress at the QTL and transcriptome levels. *The Plant Journal*. DOI: 10.1111/tpj.15379.
- [18] Cullis, B.R. et al. (2006). On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics*. DOI: 10.1198/108571106X154443.
- [19] Box, G.E.P. et Cox, D.R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*. DOI: 10.1111/j.2517-6161.1964.tb00553.x.
- [20] Desaint, H. et al. (2023). Integration of QTL and transcriptome approaches for the identification of genes involved in tomato response to nitrogen deficiency. *BioRxiv*. DOI: <https://doi.org/10.1101/2023.10.26.564109>
- [21] Gogarten, S.M. et al. (2019). Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics (Oxford, England)*. DOI: 10.1093/bioinformatics/btz567.
- [22] Purcell, S. et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*. DOI: 10.1086/519795.

- [23] Shim, H. et al. (2015). A Multivariate Genome-Wide Association Analysis of 10 LDL Subfractions, and Their Response to Statin Treatment, in 1868 Caucasians. *PLoS ONE*. DOI: 10.1371/journal.pone.0120758.
- [24] Mata-Nicolás, E. et al. (2020). Exploiting the diversity of tomato: the development of a phenotypically and genetically detailed germplasm collection. *Horticulture Research*. DOI: 10.1038/s41438-020-0291-7.
- [25] Voight, B.F. et al. (2006). A Map of Recent Positive Selection in the Human Genome. *PLOS Biology*. DOI: 10.1371/journal.pbio.0040072.
- [26] Li, N. et al. (2023). Super-pangenome analyses highlight genomic diversity and structural variation across wild and cultivated tomato species. *Nature Genetics*. DOI: 10.1038/s41588-023-01340-y.
- [27] Lagnel, J. et al. (2009). NOBLAST and JAMBLAST: New Options for BLAST and a Java Application Manager for BLAST results. *Bioinformatics (Oxford, England)*. DOI: 10.1093/bioinformatics/btp067.
- [28] Rice, P. et al. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends in genetics: TIG*. DOI: 10.1016/s0168-9525(00)02024-2.
- [29] Riedelsberger, J. et al. (2021). Plant HKT Channels: An Updated View on Structure, Function and Gene Regulation. *International Journal of Molecular Sciences*. DOI: 10.3390/ijms22041892.
- [30] Koch, E. et al. (2013). Long Range Linkage Disequilibrium across the Human Genome. *PLoS ONE*. DOI: 10.1371/journal.pone.0080754.
- [31] Ikuyinminu, E. et al. (2023). Transcriptome, Biochemical and Phenotypic Analysis of the Effects of a Precision Engineered Biostimulant for Inducing Salinity Stress Tolerance in Tomato. *International Journal of Molecular Sciences*. DOI: 10.3390/ijms24086988.
- [32] Asins, M.J. et al. (2020). Genetic Analysis of Root-to-Shoot Signaling and Rootstock-Mediated Tolerance to Water Deficit in Tomato. *Genes*. DOI: 10.3390/genes12010010.
- [33] Chu, M. et al. (2021). The Arabidopsis phosphatase PP2C49 negatively regulates salt tolerance through inhibition of AtHKT1;1. *Journal of Integrative Plant Biology*. DOI: 10.1111/jipb.13008.
- [34] Xing, B. et al. (2021). Functional Study of BpPP2C1 Revealed Its Role in Salt Stress in *Betula platyphylla*. *Frontiers in Plant Science*.
- [35] Li, N. et al. (2022). Identification and Characterization of Long Non-coding RNA in Tomato Roots Under Salt Stress. *Frontiers in Plant Science*.
- [36] Amoroso, C.G. et al. (2023). Comparison of Tomato Transcriptomic Profiles Reveals Overlapping Patterns in Abiotic and Biotic Stress Responses. *International Journal of Molecular Sciences*. DOI: 10.3390/ijms24044061.
- [37] Alonge, M. et al. (2020). Major Impacts of Widespread Structural Variation on Gene Expression and Crop Improvement in Tomato. *Cell*. DOI: 10.1016/j.cell.2020.05.021.
- [38] Ranjan, A. et al. (2016). eQTL Regulating Transcript Levels Associated with Diverse Biological Processes in Tomato. *Plant Physiology*. DOI: 10.1104/pp.16.00289.
- [39] Zamani Babgohari, M. et al. (2014). In silico analysis of high affinity potassium transporter (HKT) isoforms in different plants. *Aquatic Biosystems*. DOI: 10.1186/2046-9063-10-9.
- [40] Shkolnik, D. et al. (2019). CALMODULIN-BINDING TRANSCRIPTION ACTIVATOR 6: A Key Regulator of Na<sup>+</sup> Homeostasis during Germination1[OPEN]. *Plant Physiology*. DOI: 10.1104/pp.19.00119.
- [41] Razifard, H. et al. (2020). Genomic Evidence for Complex Domestication History of the Cultivated Tomato in Latin America. *Molecular Biology and Evolution*. DOI: 10.1093/molbev/msz297.
- [42] Munns, R. et James, R.A. (2003). Screening methods for salinity tolerance: a case study with tetraploid wheat. *Plant and Soil*. DOI: 10.1023/A:1024553303144.
- [43] Ran, X. et al. (2022). Study on the Relationship of Ions (Na, K, Ca) Absorption and Distribution to Photosynthetic Response of *Salix matsudana* Koidz Under Salt Stress. *Frontiers in Plant Science*.
- [44] Rao, E.S. et al. (2013). Relationship between survival and yield related traits in *Solanum pimpinellifolium* under salt stress. *Euphytica*. DOI: 10.1007/s10681-012-0801-2.

- [45] Almeida, P. et al. (2014). Role of Na<sup>+</sup>, K<sup>+</sup>, Cl<sup>-</sup>, proline and sucrose concentrations in determining salinity tolerance and their correlation with the expression of multiple genes in tomato. *AoB PLANTS*. DOI: 10.1093/aobpla/plu039.
- [46] Jaime-Pérez, N. et al. (2017). The sodium transporter encoded by the HKT1;2 gene modulates sodium/potassium homeostasis in tomato shoots under salinity. *Plant, Cell & Environment*. DOI: 10.1111/pce.12883.
- [47] <https://www.biorxiv.org/content/10.1101/2023.08.15.553433v1.full>.
- [48] Jaime-Pérez, N. et al. (2017). The sodium transporter encoded by the HKT1;2 gene modulates sodium/potassium homeostasis in tomato shoots under salinity. *Plant, Cell & Environment*. DOI: 10.1111/pce.12883.
- [49] Romero-Aranda, M.R. et al. (2020). Na<sup>+</sup> transporter HKT1;2 reduces flower Na<sup>+</sup> content and considerably mitigates the decline in tomato fruit yields under saline conditions. *Plant Physiology and Biochemistry*. DOI: 10.1016/j.plaphy.2020.05.012.
- [50] Rus, A. et al. (2006). Natural Variants of AtHKT1 Enhance Na<sup>+</sup> Accumulation in Two Wild Populations of Arabidopsis. *PLOS Genetics*. DOI: 10.1371/journal.pgen.0020210.
- [51] Kemi Idowu, M. et Adote Aduayi, E. (2007). Sodium-potassium interaction on growth, yield and quality of tomato in ultisol. *Journal of Plant Interactions*. DOI: 10.1080/17429140701713803.

**Table 1. Descriptive statistics and variance decomposition of various physiological and mineral content traits**

	Treatment effect (%)	Control treatment			Stress treatment			Variance decomposition (%)
		Mean	CV	$h^2$	Mean	CV	$h^2$	
<b>Early stage traits (20 das)</b>								
Leaf Cu (µg/g)	23.73 ***	21	0.22	0.55	25	0.26	0.49	
Leaf Fe (µg/g)	73.91 ***	80	0.29	0.61	139	0.41	0.63	
Leaf Zn (µg/g)	6.79 ns	121	0.30	0.63	125	0.26	0.71	
Leaf Ca (µg/g)	-36.02 ***	22,728	0.24	0.65	14,112	0.20	0.71	
Leaf K (µg/g)	-47.98 ***	78,952	0.18	0.56	40,684	0.31	0.79	
Leaf Mg (µg/g)	-28.37 ***	6,434	0.19	0.53	4,602	0.18	0.60	
Leaf Mn (µg/g)	-26.05 ***	479	0.29	0.57	347	0.36	0.70	
Leaf Na (µg/g)	2,004.28 ***	1,614	0.34	0.54	33,297	0.31	0.87	
Shoot FW (µg/g)	-61.14 ***	8.87	0.56	0.73	3.07	0.65	0.79	
Root FW (µg/g)	-24.27 ***	1.86	0.68	0.70	1.12	0.75	0.72	
Total FW (µg/g)	-55.64 ***	10.87	0.56	0.72	4.21	0.66	0.78	
Leaf DMC (%)	34.20 ***	7	0.25	0.68	9	0.29	0.84	
<b>Mature stage traits (&gt; 90 das)</b>								
Petiole Na (ppm)	-	-	-	-	2,786	0.49	0.97	-
Petiole K (ppm)	-	-	-	-	5,154	0.27	0.97	-
Fruit Na (ppm)	-	-	-	-	229	0.71	0.98	-
Fruit K (ppm)	-	-	-	-	2,565	0.11	0.91	-
Leaf length (cm)	-	-	-	-	34	0.18	0.84	-
Stem diameter (cm)	-	-	-	-	1.52	0.22	0.92	-
Stem length (cm)	-	-	-	-	76	0.22	0.95	-

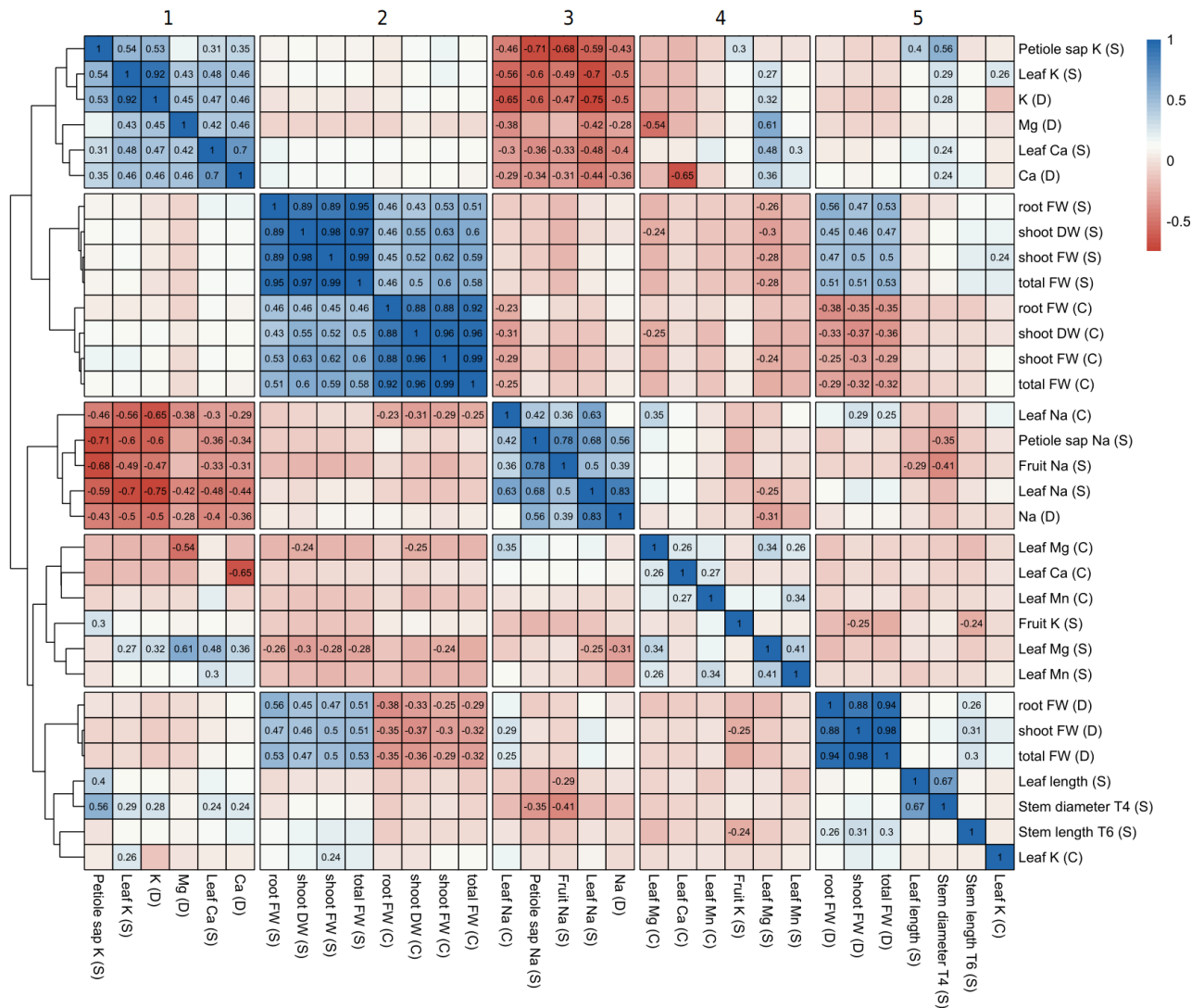
Treatment effect :  $(Y_{\text{STRESS}} - Y_{\text{CONTROL}}) * 100 / Y_{\text{CONTROL}}$  ; CV : Coefficient of variation;  $h^2$  : Cullis heritability. *P*-values for treatment effect and their level of significance are indicated as following: \*\*\*  $p < 0.001$ , ns not significant *p*. The barplot displays the proportion of each fixed effect in the total sum of squares: Genotype effect (■), Treatment effect (■), Genotype × Treatment effect (■), Random residuals (■).



**Table 2. Promoter sequence polymorphisms of *SHKT1.2***

Genome position	Position from TSS	Allele (ref/alt)	GWAS <i>p</i> -value	Cis-regulatory element	Biological function
5,127,764	-972	C/T	<b>5,10E-11</b>	Ref: (+) <b>ASF1MOTIFCAMV</b> -975 TGAC <b>G</b> -971 Alt allele = No motif	- “ <b>ASF-1 binding site</b> ” in CaMV 35S promoter found in many promoters - Are involved in transcriptional activation of several genes by auxin and/or salicylic acid
				Ref: (+) WRKY71OS -975 TGAC <b>C</b> -970 Alt allele = No motif	- “ <b>A core of TGAC-containing W-box</b> ” of, e.g., Amy32b promoter - Binding site of rice WRKY71, a transcriptional repressor of the gibberellin signalling pathway
5,127,172	-1,564	T/TC	<b>5,10E-11</b>	REF: (-) DPBFCOREDCDC3 -1565 T(-)CGTGTG -1558 Alt allele = No motif	- ABA- responsive element - ABRE/CAMTA motif - Present in the upstream region of <i>OsSOS1</i>
5,127,143	-1,593	A/C	<b>1,37E-08</b>	No differentially affected motif	
5,126,994	-1,742	G/T	<b>6,02E-10</b>	No differentially affected motif	

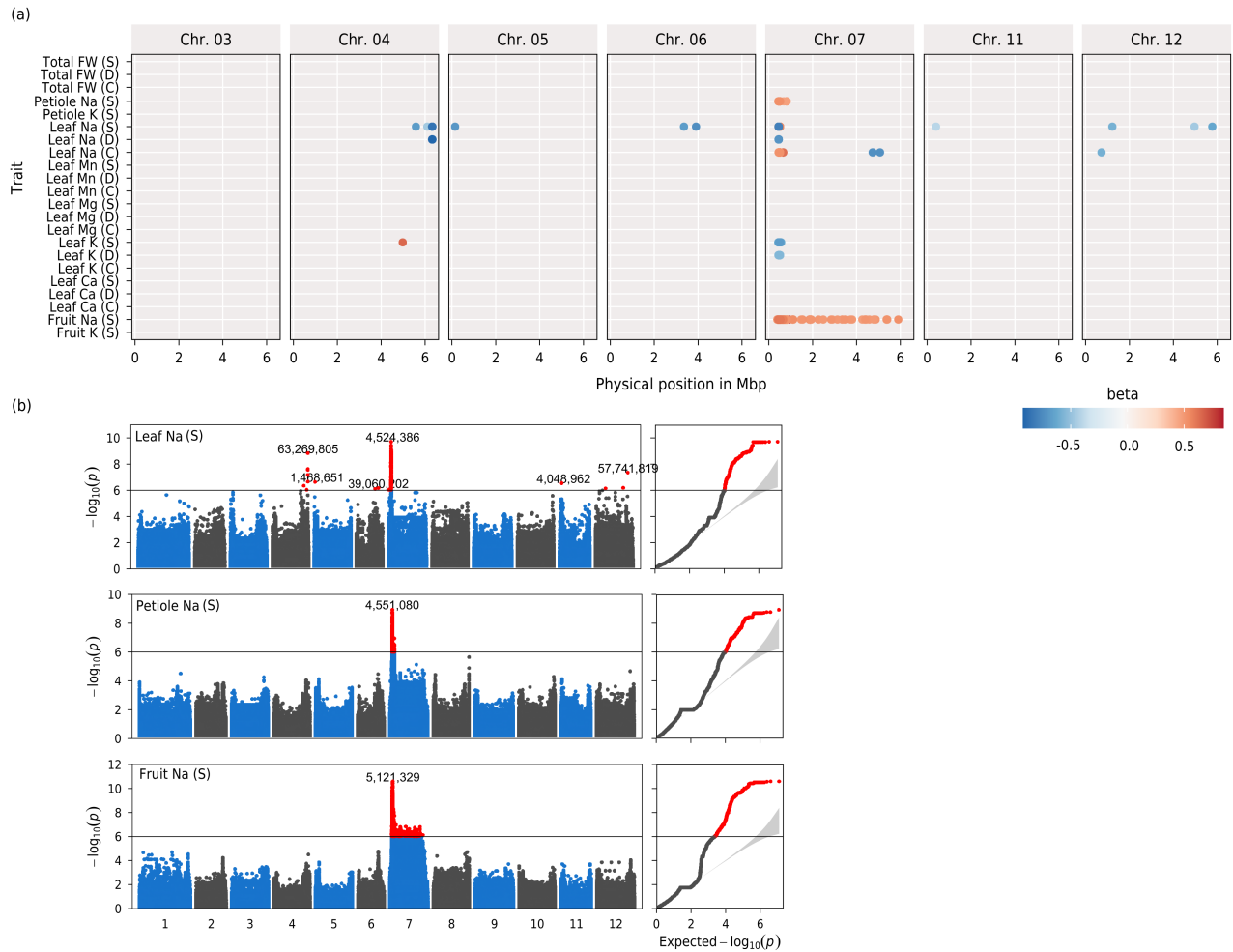
This table displays all significant polymorphisms in the *SHKT1.2* promoter sequence (-2000: 0 bp of TSS). Cis-regulatory elements were identified using the PLACE database (<https://www.dna.affrc.go.jp/PLACE/?action=newplace>). ‘Genome position’ corresponds to the true physical position according to SL4.0. ‘Position from TSS’ corresponds to the position of the mutation according to zero, the position of the Transcription Start Site.



**Figure 1. Correlation between minerals and biomass-related traits.**

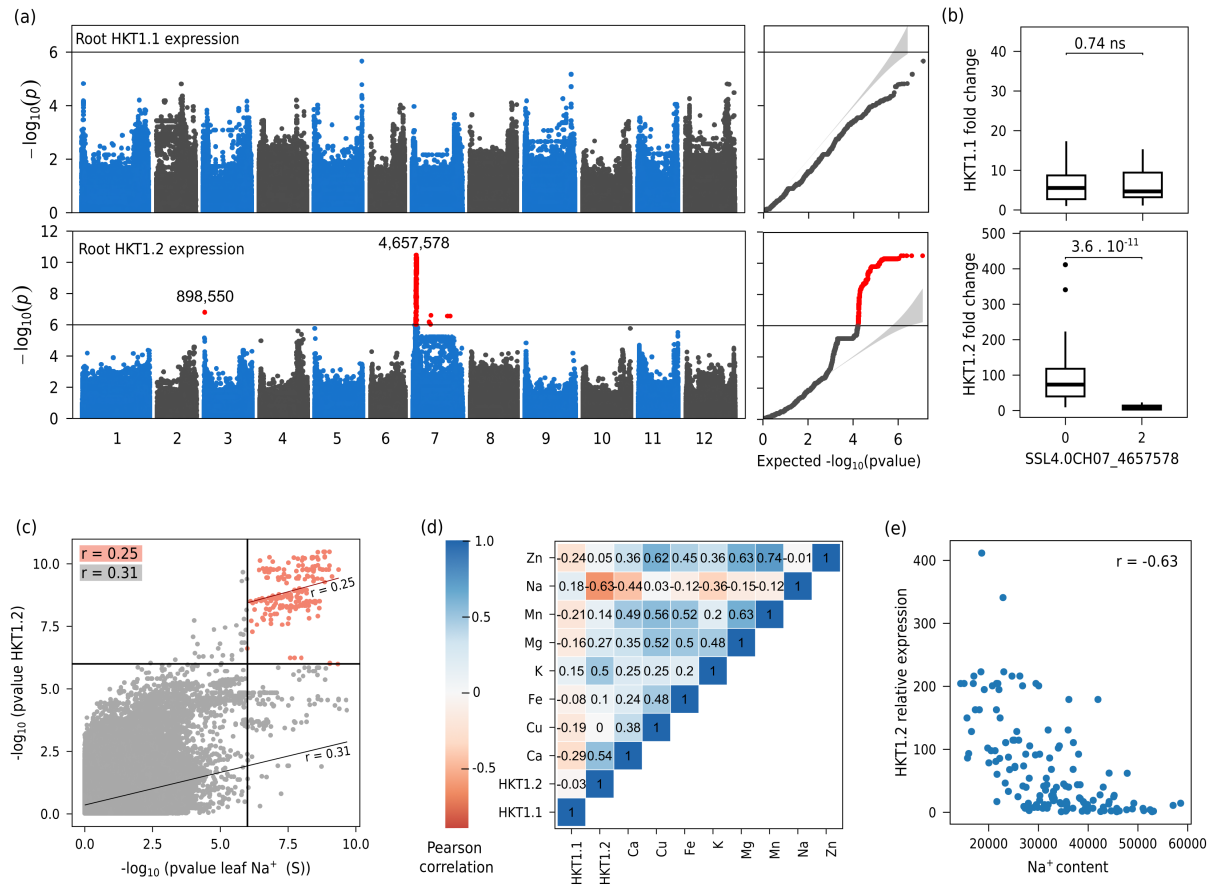
Hierarchical clustering heatmap of Pearson correlation coefficients between minerals and biomass-related traits.

The matrix represents the recorded traits under three conditions: Control (C), Salinity Treatment (S), and the ratio of the two treatments (D). The colour gradient provides insight into the correlation type; a value of 1 means a complete positive correlation (represented by dark red), while a value of -1 indicates a full negative correlation (represented by dark blue) between two traits. Only those correlations that are statistically significant ( $p$ -value  $< 0.01$ ) are shown. Numbers on top of the matrix indicated the clusters of traits.



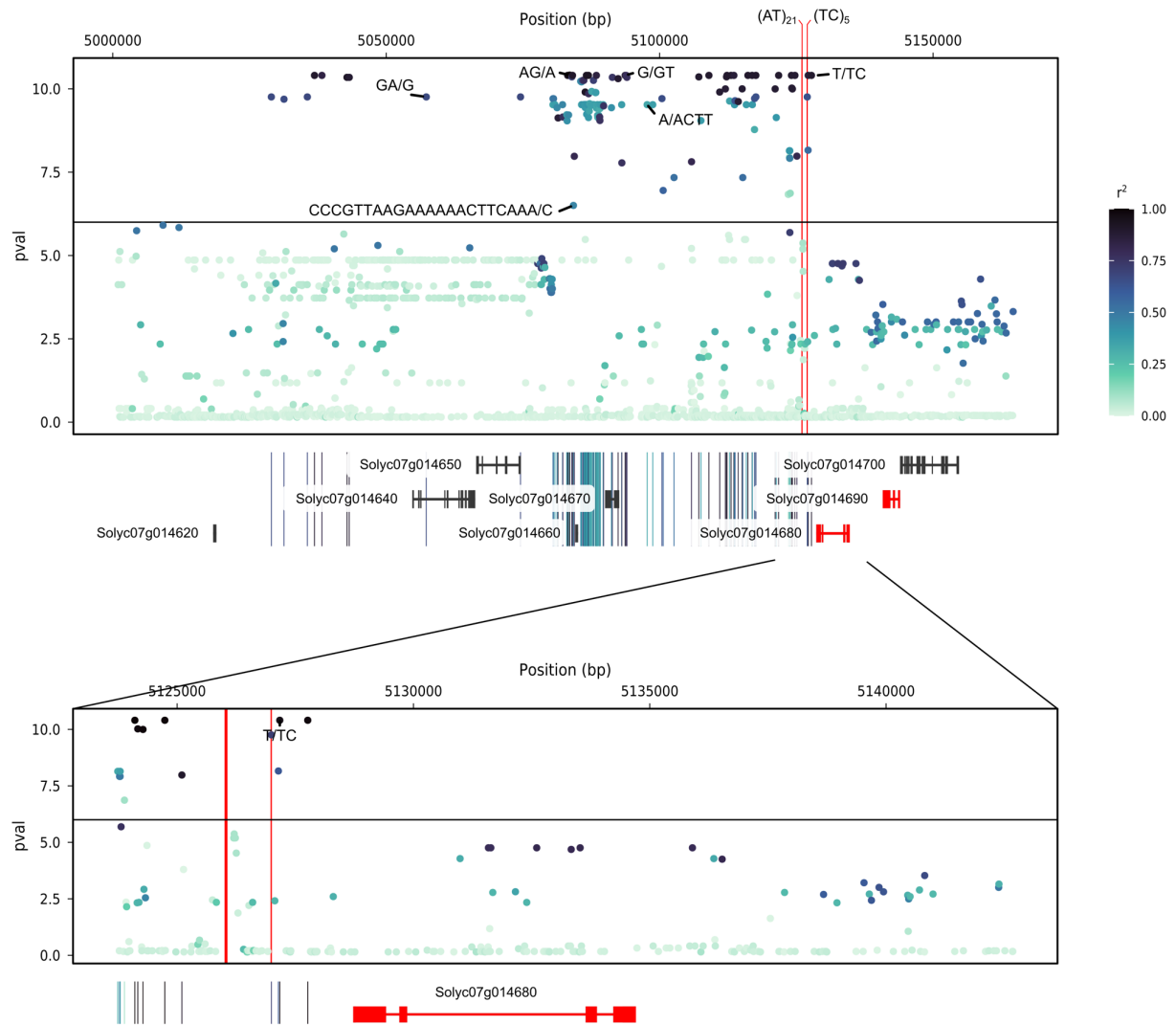
**Figure 2. Chromosome distribution of QTLs associated with biomass and mineral content traits.**

(a) QTL summary showing the physical positions of GWA significant SNP (LOD > 6). Each row represents the QTL mapping of a single trait. The dot colour denotes regression beta coefficients (blue: negative effect; red: positive effect). The x-axis indicates the physical positions on the SL4.0 tomato genome. In trait nomenclature, C/S/D stands for control, salt and delta conditions (i.e. trait plasticity). (b) Manhattan plot and qq-plots for Leaf Na<sup>+</sup> (S), Petiole sap Na<sup>+</sup> (S) and Fruit Na<sup>+</sup> (S). Red dots indicate SNPs that have reached GWAS significance. For ease of interpretation, the position of only the SNP with the highest significant p-value is indicated, along with their respective physical positions.



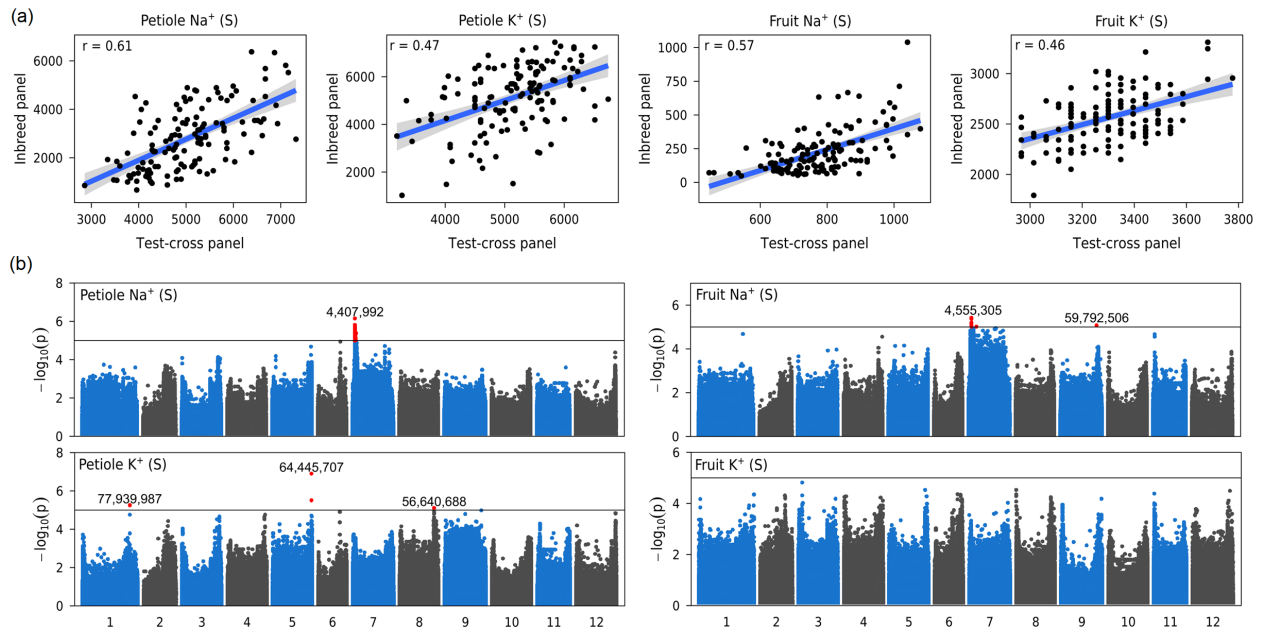
**Figure 3. Colocalisation between sodium content QTL and expression QTL (eQTL) for *S/HKT1.2* root expression**

**(a)** Manhattan plot and qqplot of GWAS analysis of *S/HKT1.1* and *S/HKT1.2* expression in root tissues under salt stress. **(b)** Boxplot displaying the difference in *S/HKT1.1* and *S/HKT1.2* expression between the three genotypes of SSL4.0CH07\_4657578 (0 reference alleles, 1 heterozygous, 2 alternative alleles). **(c)** Scatterplot displaying the correlation between GWAS p-values for the Leaf Na<sup>+</sup> (S) and GWAS p-values for *S/HKT1.2*. Significant markers for both traits are colored in red. **(d)** Correlation matrix displaying the correlations between the expression of the two HKT1 genes in root tissues with leaf mineral content. **(e)** Relationship between leaf sodium content and *S/HKT1.2* relative expression.



**Figure 4. Local visualisation of GWAS analysis for *S/HKT1.2* root expression**

The upper panel shows the local Manhattan plot of the GWAS results. The colour legends correspond to the linkage disequilibrium ( $r^2$ ) between the most significant markers and all other positions. Short indels are labelled only if they are significant for GWAS. The red genes highlight *S/HKT1.2* (Solyc07g014680) and *S/HKT1.1* (Solyc07g014690). The position of two short tandem repeats (not analysed by GWAS) is represented by the red vertical lines. The lower panel shows the annotated genes (SL4.0 genome version) in the region below the peak.



**Figure 5. GWAS for Na<sup>+</sup> and K<sup>+</sup> content in an hybrid genetic background**

(a) Relationships between traits measured in the core collection (Inbred panel) and the test-cross panel and Pearson correlation coefficient (b) Manhattan plot showing marker-trait associations for Petiole Na<sup>+</sup> and K<sup>+</sup> in stress condition (S) and Fruit Na<sup>+</sup> and K<sup>+</sup> in stress condition (S) in the Test cross panel.