THE EMBO JOURNAL

Loss of salt tolerance during tomato domestication conferred by variation in a Na+/K+ transporter

Zhen Wang, Yechun Hong, Guangtao Zhu, Yumei Li, Qingfeng Niu, Juanjuan Yao, Kai Hua, Jinjuan Bai, Yingfang Zhu, Huazhong Shi, Sanwen Huang, and Jian-Kang Zhu

Review timeline:

Submission date:	19th Aug 2019
Editorial Decision:	26th Sep 2019
Revision received:	10th Dec 2019
Editorial Decision:	10th Jan 2020
Revision received:	2nd Feb 2020
Editorial Decision:	11th Feb 2020
Revision received:	12th Feb 2020
Accepted:	13th Feb 2020

Editor: Ieva Gailite

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 26th Sep 2019

Thank you for submitting your manuscript for consideration by The EMBO Journal. We have now received three referee reports on your manuscript, which are included below for your information.

As you will see from the comments, all reviewers appreciate the work and the topic. However, they also raise a number of substantial concerns regarding core aspects of the study that would need to be conclusively addressed before they can support publication here. Based on the overall interest expressed in the reports, I would like to invite you to submit a revised version of your manuscript, in which you address the comments of all three referees.

I should add that it is The EMBO Journal policy to allow only a single major round of revision and that acceptance or rejection of the manuscript will depend on the completeness of your responses in this revised version. Since a large number of key experiments potentially impacting on the main message of the study would be required to fulfill all the referee requests, please contact me if you would like to discuss the feasibility of any of the revision aspects.

REFEREE REPORTS:

Referee #1:

The paper by Wang et al. consists on the study of a Na/K transporter of tomato from the HAK/KUP/KT family, SlHAK20, that may be a determinant of salt tolerance in tomato and probably other plant species such as rice. The GWAS-based approach used is very interesting and clearly points to SlHAK20 as a putative determinant of the Na/K ratio in plants exposed to salinity, and probably as a determinant of salt tolerance. It is very interesting that a transporter of the HAK/KUP/KT family, SlHAK20, is characterized as a K as well as a Na transporter, which adds

support to the previously published putative role of some group-IV HAK/KUP/KT transporters in Na transport. Although the results clearly show that SlHAK20 may be a determinant of salt tolerance in tomato plants, the mechanisms involved in SlHAK20-meditated salt tolerance are not sufficiently characterized. In addition, the authors propose that different SlHAK20 haplotypes show different Na transport capacities which would lead to different levels of salt tolerance. However, the data provided do no support this important conclusion.

Specific comments are as follows:

Comments to Figure 3:
Although the authors claim that differences in Na transport are found among the different SlHAK20 haplotypes, I think that data on Figures 3 g and h do not support differences in Na transport among the Hap1, Hap2 and Hap3. Figure 3 h shows SlHAK20-mediated Na uptake in yeast from very low Na concentrations, probably high-affinity Na uptake, although a kinetic characterization for this Na transport is not shown. In Figure 3h no differences among the three HAK20 haplotypes Hap1, Hap2 and Hap3 are observed. As deduced from the data presented, these three types of transporters probably mediate Na uptake with the same kinetic properties. But if the authors want to claim that these properties are different, Km and Vmax values of Na uptake should be calculated for them. In addition, data on Figure 3g do not support that differences among the three haplotypes exist. Figure 3g shows the optical density of yeast suspensions after a period of time of cells growing in media with different Na concentrations in the millimolar range up to 200 mM. First of all, if SlHAK20 is a high-affinity Na transporter, as it probably is, no differences in Na uptake nor yeast growth will be observed in the millimolar range of concentrations, because in this range SlHAK20-medaited Na uptake is saturated and most of the Na enters the yeast cells through endogenous transporters. Secondly, the square-fitting of data used makes difficult to identify differences in yeast growth. Therefore, to demonstrate differences in yeast growth, proper yeast growth rate studies at differenct external Na concentrations should be performed, determining yeast optical density at different time points, plotting the optical density versus time and fitting those data to an exponential curve to determining the growth rate.
In conclusion, according to yeast experiments shown in Figure 3 it can be concluded that SlHAK20 is a K and a Na transporter, probably with high-affinity for both cations, but these experiments do not demonstrate different ion transport capacities among the different HAK20 versions.

Regarding the generation of CRISPR edited lines, it is common that, in addition to the desired mutation in the chosen target sequence, other off-target edition in the plant genome occurs. The authors should provide information on probable off-target DNA regions of the gRNA used and that the occurrence of physiologically relevant editions in these off-targets has not occurred.

The idea presented in lines 220-224 is very interesting: "These results suggest that SlHAK20 may be involved in Na+ loading/unloading from xylem in roots and shoots under salt stress conditions. It is possible that SlHAK20 is primarily a K+ uptake transporter under normal growth conditions, but it promotes Na+ loading into the xylem in roots but restricts Na+ unloading from the xylem in shoots under salt stress conditions." However, this is highly speculative and it is difficult to understand how a transporter functions differently in roots and shoots. In fact, the authors state in the Discussion that the role of SlHAK20 in shoot remains elusive (line 314). That a HAK/KUP/KT transporter mediates Na transport is not a new idea and that it may play a role in salt tolerance either. Therefore, the authors need to provide more support to the mechanism involved in SlHAK20-mediated salt tolerance and Na distribution within the plant. K and Na uptake and translocation rates within the plant should be provided. Ideally this should be performed with radiotracers, but because some radiotracers may be difficult to obtain, Rb may be used as a tracer for K.

The results with the transgenic TS-670 line constitutively expressing the SlHAK20Hap1 (Fig. 6a, b) are very interesting but to demonstrate the relation between SaHAK20Hap1 overexpression and salt tolerance, K and Na concentrations and distributions within the plant should be also provided.

Some references are not properly used as for example in the sentence (lines 53-55): "HKT1 (High-affinity Potassium Transporter 1) was first identified as a Na+/K+ symporter in wheat that contributes to salt tolerance by unloading Na+ from the transpiration stream (Rubio et al, 1995)." Or in the sentence (lines 68-70): "In addition to maintaining K+ and Na+ homeostasis in plant tissues, these transporters are even involved in other cellular processes such as auxin movement and

adenylate cyclase activation (Osakabe et al, 2013; Santa-Maria et al, 1997)."

Some references are missing as for example in the sentence (lines 66-68): "The HAK/KUP/KT (High-Affinity K+/K+ Uptake/K+ Transporter) family transporters primarily mediate K+ fluxes, but some members of this family also play important roles in Na+ and Cs+ transport"

Some discrepancies may be observed in GUS staining in roots in Fig 3 and Fig S6. In Figure 3 shoots as well as roots are stained in blue indicating SlHAK20 expression in both organs whereas in Fig S6 only shoots are stained. The authors should provide and explanation for this.

Referee #2:

The authors perform GWAS for salt tolerance (measured as Na+/K+ content in roots) in a population of 369 tomato accessions, including cultivated tomato and its closest wild relative S. pimpinellifolium. The mayor GWAS peak is close to SlHAK20, a locus coding for a high affinity potassium transporter. The gene contains several coding and regulatory polymorphisms differentiating accessions with high and low salt tolerance. CRISPR editing or overexpression of SlHAK20 decreases and increases salt tolerance respectively, proving that this locus is involved in stress tolerance in tomato.
The findings in this work are very exciting and of high interest for fundamental and applied plant science. The language used in the manuscript would greatly benefit from English corrections.

My only major concern in this manuscript is in the identification of natural polymorphisms at the SlHAK20 locus as the cause for natural variation in salt tolerance.
The authors follow on a 200kb region surrounding a significant SNP at position 2156747 in Chr4, and focus on a potassium transporter (SlHAK20) based on its annotation. However, the authors do not mention the presence of another potassium transporter right next to SlHAK20, which is even closer to the leading SNP in the region (pos 2072311, Supplementary Dataset 2). The authors need to comment on why this is not a good candidate gene for salt tolerance: Does it contain nonsynonymous mutations? What is the % identity with SlHAK20? Tandem gene duplication is a common source of phenotypic variation: Are both potassium transporter genes present in all accessions considered?
I do not understand the choice of polymorphisms to build haplotypes. The candidate gene shows 5 nonsynonymous SNPs, one in-frame indel in the first exon (indel48), and two non-coding indels in the promoter. According to figure 2c, none of these, except for indel48, is significantly associated with salt tolerance. Why are the authors using SNP3093 in addition to indel48 to build haplotypes? From Fig. 2d, it seems that SNP3093 has very limited effect in salt tolerance. The manuscript would gain in simplicity if the authors would only use indel48. However, if the authors prefer to use three haplotypes, they will need to combine the panels in Fig. 2e and Fig. 2f to show the three haplotypes instead of the individual mutations (in other words, how are these two mutations associated during tomato evolution?).
The authors use Figure 2b to claim signals for domestication and improvement sweeps in the chromosomal region of SlHAK20. The figure shows nucleotide diversity (pi) and in the case of positive selection we would see a local reduction in pi from cultivated tomato (or cerasiforme or both), but not in S. pimpinellifolium. This is not observed in Fig 2b. In any case, a better method to observe these signatures would be to plot the ratio (pi wild / pi cerasiforme) or (pi cerasiforme / pi cultivated) as in Lin et al. 2014. With the data shown, no signatures of selection can be claimed.
Finally, I am not convinced by the differences between the alleles (Hap1, Hap2 and Hap3) in Fig3. There is a complete lack of statistics to prove their claims that "SlHAK20Hap1 and SlHAK20Hap2 are more effective in Na+ transport than SlHAK20Hap3". For example, can they analyze the differences between lines in Fig. 3f, 3g and 3h? Also for figures 3f and 3g: Why are the lines in these panels only horizontal or vertical? The points do not seem to fit very well this pattern. Fitting a curve, like in Fig 3h, seems more appropriated.
In summary, the closing sentence of this section: "These results clearly indicate that the two natural variations in SlHAK20 are the causal sites for the difference in Na+/K+ ratios among the accessions" does not hold for two reasons: First, SNP3093 does not seem to have an effect in salt tolerance. Second, the authors fail to show significant differences between the alleles in the yeast

essays. I think the rest of the manuscript shows without doubt that SlHAK20 affects salt tolerance, but the identification of the cause of this variation is vague. I would definitely look at (and discuss in the manuscript) the potassium transporter next to the SlHAK20 to look for a stronger case of natural variants affecting salt tolerance between accessions.

Minor comments:
Figure 1c. Numbers for 'n' in the legend are wrong. Also, the authors should give more detail about which accessions are used in this graph and how were they selected. For example, it is interesting that the three points with the highest Na/K ratios in roots in Figure 1b (values around 3) are not present in Figure 1c (max values around 2). Why were they removed?
Line 155 "based on the updated heinz1706 tomato genome assembly", "updated" should be replaced with the genome version number.
Supplementary Dataset 2: We cannot see the gene ids (Solyc...).
How were the markers for GWAS obtained? Can the authors offer tables with the data in order to replicate the experiment?
How was pi calculated in Figure 3b? Can the authors offer tables with the data in order to replicate the plot?
How many amino acids are affected by the 6bp (indel48) found at SlHAK20? This is interesting information for the readers to judge the putative effect of the mutation on the protein. In fact, it would be very kind to make a supplemental dataset with the sequence of the alleles that are frequently used in the paper (TS-21, TS-422, TS-577 and TS-670).
Line 163 mentions " Four nonsynonymous variations in the coding region show no effect on the phenotypic variation (Fig. 2c and Supplementary Table 3)". However it is impossible to find these variants in Suppl table 3 because the positions are different from Fig 3c and the variants do not contain information on their effect in the protein. Could this information be added to Suppl. table 3? Ideally the able could contain columns with the sequence affected at each allele and its effect in the protein (such as intergenic/synonymous/nonsynonymous).
The authors write: "Based on the identified significant variations, 328 accessions were classified into three haplotype groups". Why didn't they use all 369 accessions from the GWAS panel?
Supplementary Figure 3. The legend cannot be Na/K in root.
Figure 6B: legends are wrong.

Referee #3:

Soil salinity is a major threat to the global food security. Salinity tolerance was present in wild relatives of modern crops but lost during their domestication. In this work, Wang and co-authors have undertaken a genetic approach (GWAS and a loss-of-function mutation analyses) to compare mechanisms of salinity tolerance in cultivated and wild tomato accessions. The authors show that the natural variations in SlHAK20, a member of HAK/KUP/KT family clade IV, significantly contribute to Na+ homeostasis and is a critical determinant of salinity tolerance in tomato species. To the best of my knowledge, this is the first report of this sort that narrows down the loss of salinity tolerance trait to operation of merely one gene. Also, until now the role of HAK transporters was attributed largely to high-affinity K+ uptake. In this context, the proposed role of HAK20 in xylem Na+ loading is intriguing and highly interesting.

Overall, I am very positive about this work and would eventually like to see it published. There are a few issues, however, that prevent me from recommending acceptance in the current form. Specifically:

1. I am really puzzled with reported xylem ion concentrations data (Fig 4). Taking K+ as an example, the reported xylem sap [K+] ranges between 30 and 40 micrograms per microliter. This makes it 40 grams per liter, or 40000 ppm, or over 1M. There is absolutely no way this may be true. The same question is applicable to xylem [Na+] that exceeds 850 mM. Obviously, something went very wrong here. The numbers (units?) need to be checked and corrected.

2. Apart from the obvious mistake in reported data, I have some serious issue with using Na/K ratio as a physiological marker in this study. I completely agree that such ratio is one of the most critical

determinants of salinity tolerance; however, it this the CYTOSOLIC Na/K ratio that matters. Using the whole plant Na and K data comes with a caveat that accessions with better vacuolar Na sequestration ability would have relatively high Na/K ration at the tissue level yet remain salt tolerant. This is not obvious from the text, so the reader will be misled. Thus, to justify the use of the whole tissue Na and K content and conclusions made in this work, the authors need to demonstrate that the vacuolar Na sequestration ability was not different between contrasting accessions. By no means I am asking to do all 369 accessions. However, as a very least the authors should compare four accessions selected for details studies (depicted in Fig 1D-F) and provide the evidence of the intracellular Na distribution (using fluorescent Na dyes).

3. While the idea of HAK20 operating in xylem loading is highly interesting, all the reporting evidence are merely circumstantial. Instead of making multiple assumptions, the authors should compare the rate of Na and K loading into the xylem between overexpressing lines and those lacking functional HAK20 in direct experiments. Such measurements can be easily conducted on isolated stellar tissue using ion-selective microelectrode technique(s).

4. While HAK20 may indeed contribute to xylem Na loading, it is unlikely that this transporter will be the only one involved in this process. Hence, the authors should demonstrate that the role of other possible candidate transporters (e.g. SOS1 and CCC) are relatively minor and not differ between contrasting accessions.

Minor issues:
- Ln 47-50. Potassium comes out of blue here, without explanation if its essentiality and a role in salinity tolerance. A brief summary of the literature is needed here, to justify the experimental approach. This should also include newly discovered role of K+ as a second messenger and a determinant of the cell fate under saline conditions.
- Ln 53. The reference to 1994 is too old. Please refer to more recent reviews published in the last 1-2 years.
- Ln 91. Why roots? Some justification is needed.

1st Revision - authors' response 10th Dec 2019

Point-by-point response to referees' comments

Referee #1:

The paper by Wang et al. consists on the study of a Na/K transporter of tomato from the HAK/KUP/KT family, SlHAK20, that may be a determinant of salt tolerance in tomato and probably other plant species such as rice. The GWAS-based approach used is very interesting and clearly points to SlHAK20 as a putative determinant of the Na/K ratio in plants exposed to salinity, and probably as a determinant of salt tolerance. It is very interesting that a transporter of the HAK/KUP/KT family, SlHAK20, is characterized as a K as well as a Na transporter, which adds support to the previously published putative role of some group-IV HAK/KUP/KT transporters in Na transport. Although the results clearly show that SlHAK20 may be a determinant of salt tolerance in tomato plants, the mechanisms involved in SlHAK20-meditated salt tolerance are not sufficiently characterized. In addition, the authors propose that different SlHAK20 haplotypes show different Na transport capacities which would lead to different levels of salt tolerance. However, the data provided do no support this important conclusion.

Specific comments are as follows:

1) Comments to Figure 3:

Although the authors claim that differences in Na transport are found among the different SlHAK20 haplotypes, I think that data on Figures 3 g and h do not support differences in Na transport among the Hap1, Hap2 and Hap3. Figure 3 h shows SlHAK20-mediated Na uptake in yeast from very low Na concentrations, probably high-affinity Na uptake,

although a kinetic characterization for this Na transport is not shown. In Figure 3h no differences among the three HAK20 haplotypes Hap1, Hap2 and Hap3 are observed. As deduced from the data presented, these three types of transporters probably mediate Na uptake with the same kinetic properties. But if the authors want to claim that these properties are different, Km and Vmax values of Na uptake should be calculated for them. In addition, data on Figure 3g do not support that differences among the three haplotypes exist. Figure 3g shows the optical density of yeast suspensions after a period of time of cells growing in media with different Na concentrations in the millimolar range up to 200 mM. First of all, if SlHAK20 is a high-affinity Na transporter, as it probably is, no differences in Na uptake nor yeast growth will be observed in the millimolar range of concentrations, because in this range SlHAK20-mediated Na uptake is saturated and most of the Na enters the yeast cells through endogenous transporters. Secondly, the square-fitting of data used makes difficult to identify differences in yeast growth. Therefore, to demonstrate differences in yeast growth, proper yeast growth rate studies at different external Na concentrations should be performed, determining yeast optical density at different time points, plotting the optical density versus time and fitting those data to an exponential curve to determining the growth rate.

In conclusion, according to yeast experiments shown in Figure 3 it can be concluded that SlHAK20 is a K and a Na transporter, probably with high-affinity for both cations, but these experiments do not demonstrate different ion transport capacities among the different HAK20 versions.

Response: We appreciate greatly this excellent point raised by the reviewer! According to the Referee #2's suggestion and the results shown in Figure 2C, we rebuilt haplotypes of *SlHAK20* as Hap1 (tolerance allele) and Hap2 (sensitive allele) based on Indel48, which is significantly associated with root Na^+/K^+ ratios and salt tolerance (Fig 2D, Appendix Fig S8 in the revision). We therefore used two haplotypes, Hap1 and Hap2, in the yeast experiments for simplicity and better understanding. We tested the growth inhibition of yeast mutant cells by externally supplemented Na^+, and the results showed clear differences among $SlHAK20_{Hap1}$, $SlHAK20_{Hap2}$ and the empty vector control when 10 or 20 mM NaCl was present in the medium (Fig R1). Based on this result, we used AP medium with 10 mM Na^+ to determine the growth inhibition by SlHAK20 variants in the yeast mutant. In the revision, Figure 3G shows that the growth of ANT3 yeast mutant with $SlHAK20_{Hap1}$ or $SlHAK20_{Hap2}$ was more inhibited than that with empty vector in the presence of 10 mM NaCl, which indicates that SlHAK20 mediates Na^+ influx transport in yeast. The Na^+ uptake kinetics analysis was carried out on the yeast cells growing in AP medium with 60 μM Na^+, and the result in Fig 3H indicated that $SlHAK20_{Hap1}$ (K_m = 3.9 μM, V_{max} = 23.2 μM/min) is more effective in Na^+ transport than $SlHAK20_{Hap2}$ (K_m = 12.3 μM, V_{max} = 22.7 μM/min), supporting that Hap1 is a more active allele of *SlHAK20*. Our results also suggest that SlHAK20 is a high-affinity Na^+ transporter (Figure 3H) with K^+ permeability (Figure 3F).

Figure for reviewers removed.

B, C The growth of the ANT3 yeast strain transformed with the *SlHAK20$_{Hap1}$*, *SlHAK20$_{Hap2}$*, or empty vector (p416-GPD) at 30 mM (B) and 50 mM (C) Na^+. The data shown are mean values of three replicates.

2) Regarding the generation of CRISPR edited lines, it is common that, in addition to the desired mutation in the chosen target sequence, other off-target edition in the plant genome occurs. The authors should provide information on probable off-target DNA regions of the gRNA used and that the occurrence of physiologically relevant editions in these off-targets has not occurred.
Response: Thanks for your suggestion! To address the specificity of gene editing by CRISPR-Cas9, we performed Sanger sequencing of the four potential off-target loci predicted by the web tool Cas-OFFinder (www.rgenome.net/cas-offinder/). We did not find editing of any of these potential off-target loci in our gene editing lines. We showed these results in Fig EV5 and added some description in the results section in the revised manuscript (Line 257-262).

3) The idea presented in lines 220-224 is very interesting: "These results suggest that SlHAK20 may be involved in Na+ loading/unloading from xylem in roots and shoots under salt stress conditions. It is possible that SlHAK20 is primarily a K+ uptake transporter under normal growth conditions, but it promotes Na+ loading into the xylem in roots but restricts Na+ unloading from the xylem in shoots under salt stress conditions." However, this is highly speculative and it is difficult to understand how a transporter functions differently in roots and shoots. In fact, the authors state in the Discussion that the role of SlHAK20 in shoot remains elusive (line 314). That a HAK/KUP/KT transporter mediates Na transport is not a new idea and that it may play a role in salt tolerance either. Therefore, the authors need to provide more support to the mechanism involved in SlHAK20-mediated salt tolerance and Na distribution within the plant. K and Na uptake and translocation rates within the plant should be provided. Ideally this should be performed with radiotracers, but because some radiotracers may be difficult to obtain, Rb may be used as a tracer for K.
Response: Thanks for your constructive suggestions! We agree that using radiotracers is an ideal way to measure SlHAK20-mediated K^+/Na^+ uptake and translocation rates, but unfortunately our institute does not have a permit to use isotopes with γ radioactivity such as Na-22, and our institute does not have access to Rb-86 since it has been discontinued from PerkinElmer. We instead used non-invasive, scanning ion-selective electrode technique to determine the net Na^+ and K^+ flux in the root of *slhak20-1*, *slhak20-2*, and TS-21 after salt stress. This method has been used in several recent studies to measure Na+ and K+ fluxes in plants (Shabala et al., 2006, doi: 10.1111/j.1574-6976.2006.00019.x). The measurement showed that Na^+ efflux was decreased in the root maturation zone of *slhak20* mutants compared with the wild type TS-21, while K^+ flux did not show substantial difference between the mutants and wild type after salt stress (Appendix Fig S5). The net ion fluxes in the roots of *SlHAK20$_{Hap1}$* overexpression line (Hap1OE-1) and the wild type TS-670 were also determined, and higher Na^+ efflux but comparable K^+ flux was observed in roots of Hap1OE-1 when compared with wild type TS-670 after salt stress (Appendix Fig S9). These results suggest that SlHAK20 functions as a Na^+ efflux transporter in the root under salt stress.

4) The results with the transgenic TS-670 line constitutively expressing the SlHAK20Hap1 (Fig. 6a, b) are very interesting but to demonstrate the relation between SlHAK20Hap1 overexpression and salt tolerance, K and Na concentrations and distributions within the plant should be also provided.
Response: Thanks for your suggestion! We determined the Na^+ and K^+ contents in the shoots and roots of three Hap1OE lines and TS-670 wild type. The results showed that less Na^+ and more K^+ accumulated in the shoot of Hap1OE lines than in TS-670 after salt treatment for 7 days. The shoot Na^+/K^+ ratios were lower in Hap1OE lines than in TS-670,

indicating that overexpression of *SlHAK20$_{Hap1}$* reduces Na^+ accumulation but enhances K^+ accumulation in the shoot under salt stress. Both Na^+ and K^+ contents were decreased in the root of Hap1OE lines compared with TS-670 under salt stress, while the root Na^+/K^+ ratios did not show significant difference between Hap1OE lines and TS-670 plants, indicating that overexpression *SlHAK20$_{Hap1}$* also reduces Na^+ accumulation in roots. Our results suggest that *SlHAK20$_{Hap1}$* overexpression improves salt tolerance in tomato by decreasing Na^+ accumulation. These results are presented in the Figure 6E - G, and also described in the manuscript (line 289-299).

5) Some references are not properly used as for example in the sentence (lines 53-55): "HKT1 (High-affinity Potassium Transporter 1) was first identified as a Na+/K+ symporter in wheat that contributes to salt tolerance by unloading Na+ from the transpiration stream (Rubio et al, 1995)." Or in the sentence (lines 68-70): "In addition to maintaining K+ and Na+ homeostasis in plant tissues, these transporters are even involved in other cellular processes such as auxin movement and adenylate cyclase activation (Osakabe et al, 2013; Santa-Maria et al, 1997)."
Response: Thanks! We have corrected these citations, and the references listed below are now cited in the manuscript.

Schachtman DP, Schroeder JI (1994) Structure and transport mechanism of a high-affinity potassium uptake transporter from higher plants. *Nature* 370: 655-8

Vicente-Agullo F, Rigas S, Desbrosses G, Dolan L, Hatzopoulos P, Grabov A (2004) Potassium carrier TRH1 is required for auxin transport in Arabidopsis roots. *Plant J* 40: 523-35

Al-Younis I, Wong A, Gehring C (2015) The Arabidopsis thaliana K(+)-uptake permease 7 (AtKUP7) contains a functional cytosolic adenylate cyclase catalytic centre. *Febs Lett* 589: 3848-52

6) Some references are missing as for example in the sentence (lines 66-68): "The HAK/KUP/KT (High-Affinity K+/K+ Uptake/K+ Transporter) family transporters primarily mediate K+ fluxes, but some members of this family also play important roles in Na+ and Cs+ transport"
Response: Thanks! The following references are now added in the revision:

Benito B, Garciadeblas B, Rodriguez-Navarro A (2012) HAK Transporters from Physcomitrella patens and Yarrowia lipolytica Mediate Sodium Uptake. *Plant Cell Physiol* 53: 1117-1123

Nieves-Cordones M, Mohamed S, Tanoi K, Kobayashi NI, Takagi K, Vernet A, Guiderdoni E, Perin C, Sentenac H, Very AA (2017) Production of low-Cs(+) rice plants by inactivation of the K(+) transporter OsHAK1 with the CRISPR-Cas system. *Plant J* 92: 43-56

7) Some discrepancies may be observed in GUS staining in roots in Fig 3 and Fig S6. In Figure 3 shoots as well as roots are stained in blue indicating SlHAK20 expression in both organs whereas in Fig S6 only shoots are stained. The authors should provide and explanation for this.
Response: Thanks for your comment! We examined the expression levels of *SlHAK20* in roots of TS-21 and TS-670 using qRT-PCR, and the results showed that the expression levels of *SlHAK20* in the roots of 7- and 14-day-old seedlings were much lower than those of 18- and 21-day-old tomato plants (Fig. R2). This is the reason that the GUS activity was not detected in the roots of 7- and 14-day-old transgenic tomato plants which were used in Fig. S6.

Figure for reviewers removed.

Referee #2:

The authors perform GWAS for salt tolerance (measured as Na+/K+ content in roots) in a population of 369 tomato accessions, including cultivated tomato and its closest wild relative S. pimpinellifolium. The mayor GWAS peak is close to SlHAK20, a locus coding for a high affinity potassium transporter. The gene contains several coding and regulatory polymorphisms differentiating accessions with high and low salt tolerance. CRISPR editing or overexpression of SlHAK20 decreases and increases salt tolerance respectively, proving that this locus is involved in stress tolerance in tomato.

1) The findings in this work are very exciting and of high interest for fundamental and applied plant science. The language used in the manuscript would greatly benefit from English corrections.
Response: Thank you very much for your positive comments! The manuscript has been carefully revised by a native English-speaking scientist.

2) My only major concern in this manuscript is in the identification of natural polymorphisms at the SlHAK20 locus as the cause for natural variation in salt tolerance. The authors follow on a 200kb region surrounding a significant SNP at position 2156747 in Chr4, and focus on a potassium transporter (SlHAK20) based on its annotation. However, the authors do not mention the presence of another potassium transporter right next to SlHAK20, which is even closer to the leading SNP in the region (pos 2072311, Supplementary Dataset 2). The authors need to comment on why this is not a good candidate gene for salt tolerance: Does it contain nonsynonymous mutations? What is the % identity with SlHAK20? Tandem gene duplication is a common source of phenotypic variation: Are both potassium transporter genes present in all accessions considered?
Response: We are sorry for this confusion. In fact, the sequences of these two genes, Solyc04g008450 and Solyc04g008455, belong to a single gene but were incorrectly annotated in the tomato annotation version ITAG3.2 according to genome assemble SL3.0. Actually, the sequence of *SlHAK20* was correctly annotated as a single gene Solyc04g008450 in the older version of tomato annotation ITAG2.40 based on genome assemble SL2.50. We used Sanger sequencing to determine the cDNA sequence of *SlHAK20* and confirmed that *SlHAK20* is a single gene. In the revision, the corrected annotation information is used in Dataset EV2.

3) I do not understand the choice of polymorphisms to build haplotypes. The candidate gene shows 5 nonsynonymous SNPs, one in-frame indel in the first exon (indel48), and two non-coding indels in the promoter. According to figure 2c, none of these, except for indel48, is significantly associated with salt tolerance. Why are the authors using SNP3093

in addition to indel48 to build haplotypes? From Fig. 2d, it seems that SNP3093 has very limited effect in salt tolerance. The manuscript would gain in simplicity if the authors would only use indel48. However, if the authors prefer to use three haplotypes, they will need to combine the panels in Fig. 2e and Fig. 2f to show the three haplotypes instead of the individual mutations (in other words, how are these two mutations associated during tomato evolution?).
Response: Thanks for your excellent suggestion! We agree that the Indel48 is the only significant polymorphism associated with salt tolerance, and have removed SNP3093 from the haplotypes and now use only two haplotypes in the revised manuscript. We revised the related figures accordingly, but our conclusions are still consistent with those in the previous version. We also combined the panels of Figure 2e and 2f in the previous version into Fig 2E in the revised manuscript, and the data still support that these two haplotypes are involved in domestication and improvement sweeps during selection for larger fruits. Please refer to the result in Figure 2E.

4) The authors use Figure 2b to claim signals for domestication and improvement sweeps in the chromosomal region of SlHAK20. The figure shows nucleotide diversity (pi) and in the case of positive selection we would see a local reduction in pi from cultivated tomato (or cerasiforme or both), but not in S. pimpinellifolium. This is not observed in Fig 2b. In any case, a better method to observe these signatures would be to plot the ratio (pi wild / pi cerasiforme) or (pi cerasiforme / pi cultivated) as in Lin et al. 2014. With the data shown, no signatures of selection can be claimed.
Response: Thank you for your suggestion! In the revision, the layout of the diversity (π) has been substituted by the ratio of π_{PIM}/π_{CER} or π_{CER}/π_{BIG} in Figure 2B. The results show that *SlHAK20* is involved in the domestication sweep rather than the improvement sweep. This result is consistent with the analysis of *SlHAK20* allele frequency. The frequency of *SlHAK20* wild allele in PIM, CER, and BIG is 75.0%, 10.2%, 1.9% (Fig 2E), respectively, indicating that this gene underwent stringent selection during domestication and improvement for large fruits.

5) Finally, I am not convinced by the differences between the alleles (Hap1, Hap2 and Hap3) in Fig3. There is a complete lack of statistics to prove their claims that "$SlHAK20_{Hap1}$ and $SlHAK20_{Hap2}$ are more effective in Na+ transport than $SlHAK20_{Hap3}$". For example, can they analyze the differences between lines in Fig. 3f, 3g and 3h? Also for figures 3f and 3g: Why are the lines in these panels only horizontal or vertical? The points do not seem to fit very well this pattern. Fitting a curve, like in Fig 3h, seems more appropriated.
Response: Thanks for your comments and suggestions! Following your suggestion, two haplotypes (Hap1 and Hap2) were built based on Indel48, and the figures, including Fig 3f, 3g and 3h, were re-organized using fit curve in the revision. The results showed that Hap1 is a stronger transporter for Na^+ and K^+ than Hap2 in yeast (Fig 3F - H), indicating that Hap1 is a more active allele of *SlHAK20* than Hap2.

6) In summary, the closing sentence of this section: "These results clearly indicate that the two natural variations in SlHAK20 are the causal sites for the difference in Na+/K+ ratios among the accessions" does not hold for two reasons: First, SNP3093 does not seem to have an effect in salt tolerance. Second, the authors fail to show significant differences between the alleles in the yeast essays. I think the rest of the manuscript shows without doubt that SlHAK20 affects salt tolerance, but the identification of the cause of this variation is vague. I would definitely look at (and discuss in the manuscript) the potassium transporter next to the SlHAK20 to look for a stronger case of natural variants affecting salt tolerance between accessions.
Response: We greatly appreciate your excellent points! As mentioned earlier, *SlHAK20* is the only transporter gene at this genetic locus and Solyc04g008455 was incorrectly annotated as a gene. We have corrected the annotation information in Dataset EV2.

Minor comments:

1) Figure 1c. Numbers for 'n' in the legend are wrong. Also, the authors should give more detail about which accessions are used in this graph and how were they selected. For example, it is interesting that the three points with the highest Na/K ratios in roots in Figure 1b (values around 3) are not present in Figure 1c (max values around 2). Why were they removed?

Response: Thanks for your comments and questions! The legend was corrected and the fruit weight of 195 accessions was added into Dataset EV1 in the revision. The fruit weights of accessions grown in the same filed were chosen for the determination of the relationship with root Na^+/K^+ ratios in Fig 1C, but unfortunately the three concerned accessions were not included in this experiment. We have now determined the fruit weights of these three accessions and the results showed that the means of fruit weight are 188.2, 181.6 and 264.2 g for TS-188, TS-272 and TS-577, respectively, which is correlated with high root Na^+/K^+ ratios.

2) Line 155 "based on the updated heinz1706 tomato genome assembly", "updated" should be replaced with the genome version number.

Response: Thanks for your suggestion! We have revised the corresponding statement in the manuscript as “based on the Heinz 1706 tomato genome assembly (version SL2.50) as a reference genome”.

3) Supplementary Dataset 2: We cannot see the gene ids (Solyc...).

Response: Thanks! Correction has been made.

4) How were the markers for GWAS obtained? Can the authors offer tables with the data in order to replicate the experiment?

Response: Thanks! We have added the SNP calling into the methods section. We used SOAP2 to map all the sequencing reads from each accession to the tomato reference genome (Version SL2.50) with the following parameters: -m 100, -x 888, -s 35, -l 32, -v 3. Mapped reads were filtered to remove PCR duplicates. Both paired-end and single-end mapped reads were then used for SNP calling throughout the entire collection of tomato accessions using SOAPsnp with the following parameters: -L 100 -u -F 1. We generated the genotype likelihood across the population for each SNP with quality ≥ 40 and base quality ≥ 40. False positive SNPs were filtered in the population following the method previously described by Lin et al., 2014. A total of 2,824,130 SNPs (minor allele frequency ≥ 0.05 and missing ratio < 10 %) for these 369 accessions have been used to perform the association studies. The raw re-sequencing data can be downloaded from NCBI under Number SRP045767, PRJNA353161 and PRJEB5226–PRJEB5228, and PRJEB5253.

5) How was pi calculated in Figure 2b? Can the authors offer tables with the data in order to replicate the plot?

Response: The level of genetic diversity (π) of the whole chromosome 4 was calculated using a 200-kb (as the same size of linkage disequilibrium) window with a step size of 10 kb in PIM, CER and BIG. The regions affected by domestication should have substantially lower diversity in CER than in PIM. Improvement sweeps should show a much stronger reduction of diversity in BIG in comparison to CER. By scanning the ratios of genetic diversity between PIM and CER (π_{PIM}/π_{CER}) as well as between CER and BIG (π_{CER}/π_{BIG}), The regions with the top 10% of ratios (1.82 and 4.27 for domestication and improvement, respectively) was defined as the candidate sweeps of domestication and improvement, respectively. The interval of *SlHAK20* was located in the top 10% region (Dataset EV3).

6) How many amino acids are affected by the 6 bp (indel48) found at SlHAK20? This is interesting information for the readers to judge the putative effect of the mutation on the

protein. In fact, it would be very kind to make a supplemental dataset with the sequence of the alleles that are frequently used in the paper (TS-21, TS-422, TS-577 and TS-670).
Response: Thanks for your suggestion! Two alanine residues are encoded by the in-frame indel48 (6 bp: GCGGCG) in the SlHAK20 protein. We listed the amino acid sequence of SlHAK20 of TS-21, TS-422, TS-577 and TS-670 in Dataset EV4.

7) Line 163 mentions " Four nonsynonymous variations in the coding region show no effect on the phenotypic variation (Fig. 2c and Supplementary Table 3)". However it is impossible to find these variants in Suppl table 3 because the positions are different from Fig 2c and the variants do not contain information on their effect in the protein. Could this information be added to Suppl. table 3? Ideally the able could contain columns with the sequence affected at each allele and its effect in the protein (such as intergenic/synonymous/nonsynonymous).
Response: Thanks for your comments! We added the details of the position and sequence information of each variation to Table EV3 (Suppl. Table 3).

8) The authors write: "Based on the identified significant variations, 328 accessions were classified into three haplotype groups". Why didn't they use all 369 accessions from the GWAS panel?
Response: Thanks for the question! Some accessions were classified into other haplotype groups of *SlHAK20* showing no statistical significance because the number of accessions within these haplotype groups is fewer than six. Most of the variations in *SlHAK20* were not confirmed by Sanger sequencing in some of the accessions within these haplotype groups. In the revision, we confirmed Indel48 of *SlHAK20* using Sanger sequencing in the population, and the 369 accessions were classified into two haplotype groups based on this variation (Fig 2D).

9) Supplementary Figure 3. The legend cannot be Na/K in root.
Response: Thanks! Corrected.

10) Figure 6B: legends are wrong.
Response: Thanks! Corrected.

Referee #3:
Soil salinity is a major threat to the global food security. Salinity tolerance was present in wild relatives of modern crops but lost during their domestication. In this work, Wang and co-authors have undertaken a genetic approach (GWAS and a loss-of-function mutation analyses) to compare mechanisms of salinity tolerance in cultivated and wild tomato accessions. The authors show that the natural variations in SlHAK20, a member of HAK/KUP/KT family clade IV, significantly contribute to Na+ homeostasis and is a critical determinant of salinity tolerance in tomato species. To the best of my knowledge, this is the first report of this sort that narrows down the loss of salinity tolerance trait to operation of merely one gene. Also, until now the role of HAK transporters was attributed largely to high-affinity K+ uptake. In this context, the proposed role of HAK20 in xylem Na+ loading is intriguing and highly interesting.

Overall, I am very positive about this work and would eventually like to see it published. There are a few issues, however, that prevent me from recommending acceptance in the current form. Specifically:

1) I am really puzzled with reported xylem ion concentrations data (Fig 4). Taking K+ as an example, the reported xylem sap [K+] ranges between 30 and 40 micrograms per microliter. This makes it 40 grams per liter, or 40000 ppm, or over 1M. There is absolutely no way this may be true. The same question is applicable to xylem [Na+] that exceeds 850

mM. Obviously, something went very wrong here. The numbers (units?) need to be checked and corrected.
Response: We are sorry for this mistake on the units. The correct unit is μg/mL for both Na^+ and K^+ concentrations in the xylem sap. We changed the units for Figure 4E and 4F in the revision.

2) Apart from the obvious mistake in reported data, I have some serious issue with using Na/K ratio as a physiological marker in this study. I completely agree that such ratio is one of the most critical determinants of salinity tolerance; however, it this the CYTOSOLIC Na/K ratio that matters. Using the whole plant Na and K data comes with a caveat that accessions with better vacuolar Na sequestration ability would have relatively high Na/K ration at the tissue level yet remain salt tolerant. This is not obvious from the text, so the reader will be misled. Thus, to justify the use of the whole tissue Na and K content and conclusions made in this work, the authors need to demonstrate that the vacuolar Na sequestration ability was not different between contrasting accessions. By no means I am asking to do all 369 accessions. However, as a very least the authors should compare four accessions selected for details studies (depicted in Fig 1D-F) and provide the evidence of the intracellular Na distribution (using fluorescent Na dyes).
Response: Thanks for your suggestion! We have determined the relative intracellular Na^+ contents using CoroNa Green dye in the roots of TS-21, TS-422, TS-577 and TS-670 according to previously reported method by Wu *et al*. 2019 (doi: 10.1111/tpj.14424). The results showed that the cytosolic Na^+ levels of root cells of TS-21 and TS-422 were notably lower than those of TS-577 and TS-670, whereas there were no significant difference in the vacuolar Na^+ levels among these four accessions. These results suggest that the vacuolar Na^+ sequestration ability was not affected in these four contrasting accessions (Appendix Fig S1).

3) While the idea of HAK20 operating in xylem loading is highly interesting, all the reporting evidence are merely circumstantial. Instead of making multiple assumptions, the authors should compare the rate of Na and K loading into the xylem between overexpressing lines and those lacking functional HAK20 in direct experiments. Such measurements can be easily conducted on isolated stellar tissue using ion-selective microelectrode technique(s).
Response: Thanks for your suggestion! We failed to isolate stele tissues from tomato roots, although we agree that the suggested measurements could provide strong evidence to support our conclusions. Instead, we detected the net Na^+ and K^+ flux in the roots of *slhak20-1*, *slhak20-2*, TS-21, TS-670 and Hap1OE-1 plants treated with 50 mM NaCl for 3 days. The *slhak20* mutant showed lower Na^+ efflux but comparable K^+ flux in the root when compared with the wild type after salt stress (Appendix Fig S5). This result indicates that Na^+ exclusion in the root cells of the mutant plants is restricted under salt stress. Moreover, higher Na^+ efflux but comparable K^+ flux in the root of HapOE-1 plants were observed when compared with wild type TS-670 after salt stress (Appendix Fig 9), further supporting that SlHAK20 plays an important role in Na^+ efflux in the root.

4) While HAK20 may indeed contribute to xylem Na loading, it is unlikely that this transporter will be the only one involved in this process. Hence, the authors should demonstrate that the role of other possible candidate transporters (e.g. SOS1 and CCC) are relatively minor and not differ between contrasting accessions.
Response: Thanks for your suggestion! We extracted the SNP information of *SlSOS1*(Gene ID: Solyc01g005020), which is located in Chr 1, in the 369 population, and the data showed that the *P* values of all SNPs are higher than that of the significance threshold (Dataset EV5), suggesting that *SlSOS1* variations are not involved in controlling root Na^+/K^+ ratio. We identified two homologous genes of *AtCCC*, Solyc02g021620 and Solyc02g070290, in tomato, but they were not in the 400 kb window of the signal (02_17572854) on the Chromosome 2 in this study (Dataset EV2). Moreover, the *P* values

of SNPs in these two *SlCCCs* are also higher than that of the threshold according to the root Na^+/K^+ ratio (Dataset EV5). SlHKT1;2 was reported to regulate shoot Na^+/K^+ homeostasis during salt stress in tomato (doi: 10.1111/pce.12883). However, we did not detect any SNPs in the *SlHKT1;2* gene (Gene ID: Solyc07g014680) that is associated with the root Na^+/K^+ ratios in the 369 population. Meanwhile, we checked the coding sequence of these four genes in TS-21, TS-422, TS-577 and TS-670 using Sanger sequencing, and the result indicates that the amino acid sequences of these four genes are not different between the contrasting accessions (Dataset EV5). For the above-mentioned reasons, we think that *SlCCCs*, *SlSOS1* and *SlHKT1;2* are not casual genes for root Na^+/K^+ ratio in our study. This is now discussed in the manuscript (Line 350-354).

Minor issues:
1) Ln 47-50. Potassium comes out of blue here, without explanation if its essentiality and a role in salinity tolerance. A brief summary of the literature is needed here, to justify the experimental approach. This should also include newly discovered role of K+ as a second messenger and a determinant of the cell fate under saline conditions.
Response: Thanks for your suggestion! In the revision, we added a brief summary to illustrate the role and essentiality of potassium, highlighting its role as a messenger of salinity stress response (Line 51-60).

2) Ln 53. The reference to 1994 is too old. Please refer to more recent reviews published in the last 1-2 years.
Response: Thanks for your suggestion! Three recent reviews are now cited for this viewpoint in the revision, which are listed below:

Isayenkov SV, Maathuis FJM (2019) Plant Salinity Stress: Many Unanswered Questions Remain. *Front Plant Sci* 10: 80

Rubio F, Nieves-Cordones M, Horie T, Shabala S (2019) Doing 'business as usual' comes with a cost: evaluating energy cost of maintaining plant intracellular K(+) homeostasis under saline conditions. *The New phytologist*

Wu HH, Zhang XC, Giraldo JP, Shabala S (2018) It is not all about sodium: revealing tissue specificity and signalling roles of potassium in plant responses to salt stress. *Plant Soil* 431: 1-17

3) Ln 91. Why roots? Some justification is needed.
Response: Thanks for your comment! We also analyzed the shoot Na^+/K^+ ratio by GWAS with a Bonferroni-adjusted correction, and the results showed that no significant signal is correlated with this trait (Appendix Fig S2), whereas nine major signals associated with root Na^+/K^+ ratio were detected using the same method (Fig 2A).

2nd Editorial Decision 10th Jan 2020

Thank you for submitting a revised version of your manuscript. The study has now been re-evaluated by all original referees, and I have included their reports below for your information.

As you can see, while all reviewers appreciate the revised manuscript and find that is significantly improved, reviewers #1 and #2 find that additional experiments and textual explanations are needed before they can recommend publication of the manuscript. Therefore, I would like to invite you to submit a revised manuscript, addressing the remaining referee comments, especially regarding providing appropriate data to demonstrate the differential Na uptake properties of SlHAK20 Hap1 and Hap2 as indicated by reviewer #1. From my side, I find that inclusion of this information would

significantly strengthen the manuscript, and the experiments appear relatively straightforward to perform. Please contact me if you encounter any difficulties in performing these assays.

Please also address the following editorial issues.

REFEREE REPORTS:

Referee #1:

I consider that the authors have revised satisfactorily most of the questions that I raised in my first round of revision. The paper has increased in clarity by reducing the studies in yeast to only two haplotypes, the tolerant (Hap1) and the sensitive (Hap2). However, my main concern is that, contrarily to the authors claim, the results presented with yeast expression do not demonstrate a difference in Na uptake between the two haplotypes. The required kinetic characterization of uptake is still missing and the presented data contain important technical and conceptual flaws.
The preliminary experiments to determine the sensitivity of the yeast cells expressing the two haplotypes to Na shown in Figure R1A are OK (although I do not understand the results on Figure R1B and C. In fact, these results on Figure R1 B and C are not mentioned in the response letter from the authors). So, I understand that the authors chose 10 mM Na to evaluate sensitivity of yeast cells expressing the tomato transporters to Na (Figure 3G). However, after the revision of the manuscript, the results presented on Figure 3G do not provide yet a solid demonstration that SlHAK20 is mediating Na uptake in yeast and, more importantly, that differences between the two haplotypes exist. The results show a longer growth lag phase of yeast cells expressing the two haplotypes of SlHAK20, but similar growth rates (exponential phase of growth) of cells expressing the empty vector and the two haplotypes of SlHAK20.
As I also mentioned in my first round of revision, if SlHAK20 is a high affinity Na transporter, analyzing the yeast growth in the millimolar range of Na concentrations to disclose differences between the two haplotypes is not the best approach.
MyM: Line 480: The authors state "For kinetic analysis of K uptake in yeast,..." However, measuring OD at different time points is a study of yeast growth, not a study of K uptake or a kinetic analysis of K uptake.
Lines 488-495: The authors describe Na uptake experiments in yeast from only one external Na concentration (60 µM Na). By using this approach, they can determine the initial rate of Na uptake from 60 µM external Na, and how the Na flux is changing versus time. These data cannot be used for a kinetic characterization to determine Km and Vmax values of Na uptake. To determine Km and Vmax, initial rates of Na uptake at different external Na concentrations should be determined. From the plot of the Initial rates of Na uptake versus the external Na concentration, the Km and Vmax values could be calculated. In summary, what the authors really show on Figure 3H is in fact external Na versus time. It is clear that expression of the SlHAK20 transporter in yeast leads to a much higher rate of Na uptake in comparison with the cells expressing the empty vectors (Figure 3H). These results show that SlHAK20 behaves as a Na transporter. However, concluding that the Na uptake characteristics of Hap1 and Hap2 are different form the results presented is not possible. Again, as I suggested in my first round of revision, a proper kinetic characterization should be performed, which includes a statistical analysis to disclose significant differences between Hap1 and Hap2 Na uptake properties.
From the plot shown on Figure 3H, Km and Vmax cannot be calculated. In addition, I do not understand what is the label of the Y axis (Na uptake (µM). Is it a rate of uptake or the external Na concentration? Na uptake is not a concentration of Na but the amount of Na accumulated within the cell per unit of time and unit of yeast weight. I do not understand either how the Km and Vmax values provided on Figure 3H have been obtained.

Point 2 is revised satisfactorily

Point 3. The authors have described Na and K fluxes by invasive, scanning ion-selective electrode technique. This is OK to characterize ion fluxes at the root. The results show that the slhak20 mutant lines show decreased Na efflux and the OX line increased Na efflux.
Point 4, 5,6 and 7 are revised satisfactorily

Referee #2:

The authors perform GWAS for salt tolerance (measured as Na+/K+ content in roots) in a population of 369 tomato accessions, including cultivated tomato and its closest wild relative S. pimpinellifolium. The mayor GWAS peak is close to SlHAK20, a locus coding for a high affinity potassium transporter. The gene contains several coding and regulatory polymorphisms differentiating accessions with high and low salt tolerance. CRISPR editing or overexpression of SlHAK20 decreases and increases salt tolerance respectively, proving that this locus is involved in stress tolerance in tomato. The findings in this work are very exciting and of high interest for fundamental and applied plant science.
After revision, I have no mayor comments on the manuscript, although I still have some minor comments:
I still have problems with the closing sentence of the "A variation in SlHAK20 is associated with root Na+/K+ ratio in tomato during salt stress" section (line 195): "These results clearly indicate that this natural variation in SlHAK20 is the causal sites for the difference in Na+/K+ ratios among the accessions". The authors do not have experiments that show that "this" variation at the SlHAK20 locus is causal for the phenotype. It could be any of the mutations in the gene or in the promoter (or maybe even in the introns?). In order to prove that a specific variation at SlHAK20 is causal, they will need to transform a mutant line (the CRISPR line for example) with an allele that carries single and combinations of mutations until they find the specific mutation (or combination) that affects the phenotype. For this reason I think the best is to change the sentence to: "These results indicate that natural variation at the SlHAK20 locus is strongly associated with the difference in Na+/K+ ratios among the accessions".
In this same section and in Figure 2: I think it is not proper to call "haplotype" to a group of alleles that share a similar polymorphism. It gives the impression that all alleles in Hap1 have the same sequence although this is clearly not the case, and they only share indel48. Could you call them group1 / group2, or indel48+ and indel48-, or something along these lines?
Related to the above, in all that refers to Figure 3 is necessary that the authors state in the text what are the alleles behind Hap1 and Hap2, and what are the differences between them: Is it only indel48, or are there more differences?

Referee #3:

The authors did a good job addressing all reviewers' concerns and I have no more critical questions about this MS. The authors should be congratulated with producing a very nice and interesting piece of work.

2nd Revision - authors' response 2nd Feb 2020

Point by point response to referees' comments

Referee #1:

I consider that the authors have revised satisfactorily most of the questions that I raised in my first round of revision. The paper has increased in clarity by reducing the studies in yeast to only two haplotypes, the tolerant (Hap1) and the sensitive (Hap2). However, my main concern is that, contrarily to the authors claim, the results presented with yeast expression do not demonstrate a difference in Na uptake between the two haplotypes. The required kinetic characterization of uptake is still missing and the presented data contain important technical and conceptual flaws.

Response: Thanks for carefully evaluating our manuscript and for providing critical comments.

1) The preliminary experiments to determine the sensitivity of the yeast cells expressing the two haplotypes to Na shown in Figure R1A are OK (although I do not understand the results on Figure R1B and C. In fact, these results on Figure R1 B and C are not mentioned in the response letter from the authors). So, I understand that the authors chose 10 mM Na to evaluate sensitivity of yeast cells expressing the tomato transporters to Na (Figure 3G). However, after the revision of the manuscript, the results presented on Figure 3G do not provide yet a solid demonstration that SlHAK20 is mediating Na uptake in yeast and, more importantly, that differences between the two haplotypes exist. The results show a longer growth lag phase of yeast cells expressing the two haplotypes of SlHAK20, but similar growth rates (exponential phase of growth) of cells expressing the empty vector and the two haplotypes of SlHAK20.
As I also mentioned in my first round of revision, if SlHAK20 is a high affinity Na transporter, analyzing the yeast growth in the millimolar range of Na concentrations to disclose differences between the two haplotypes is not the best approach.

Response: Thanks for your comments! In the revision, we show the difference of Na^+ uptake between the yeast cells expressing *SlHAK20*$_{Hap1}$ and *SlHAK20*$_{Hap2}$ in the presence of 60 µM external Na^+ in Figure 3G. The result indicates that the rate of Na^+ uptake by yeast expressing *SlHAK20*$_{Hap1}$ is higher than that of the yeast expressing *SlHAK20*$_{Hap2}$, supporting that SlHAK20 mediates Na^+ transport and the variant SlHAK20$_{Hap1}$ has higher Na^+ transport activity than SlHAK20$_{Hap2}$.

2) MyM: Line 480: The authors state "For kinetic analysis of K uptake in yeast,..." However, measuring OD at different time points is a study of yeast growth, not a study of K uptake or a kinetic analysis of K uptake.

Response: Thanks for pointing out our mistake! This is now corrected as “For the analysis of yeast growth under different concentrations of K^+, …”.

3) Lines 488-495: The authors describe Na uptake experiments in yeast from only one external Na concentration (60 µM Na). By using this approach, they can determine the initial rate of Na uptake from 60 µM external Na, and how the Na flux is changing versus time. These data cannot be used for a kinetic characterization to determine Km and Vmax values of Na uptake. To determine Km and Vmax, initial rates of Na uptake at different external Na concentrations should be determined. From the plot of the Initial rates of Na uptake versus the external Na concentration, the Km and Vmax values could be calculated. In summary, what the authors really show on Figure 3H is in fact external Na versus time. It is clear that expression of the SlHAK20 transporter in yeast leads to a much higher rate of Na uptake in comparison with the cells expressing the empty vectors (Figure 3H). These results show that SlHAK20 behaves as a Na transporter. However, concluding that the Na uptake characteristics of Hap1 and Hap2 are different form the results presented is not possible. Again, as I suggested in my first round of revision, a proper kinetic characterization should be performed, which includes a statistical analysis to disclose significant differences between Hap1 and Hap2 Na uptake properties.

Response: Thank you for your constructive suggestion! In the revision, we have now determined the Na^+ uptake kinetics by culturing yeast cells with different external Na^+ concentrations and the result is shown in Figure 3H and described in the Results section (Line 215 - 219). The calculated Km for SlHAK20$_{Hap1}$ and SlHAK20$_{Hap2}$ are 26.8 ± 4.4 µM and 40.1 ± 4.7 µM, respectively. This indicates that SlHAK20$_{Hap1}$ has higher Na^+-binding affinity than SlHAK20$_{Hap2}$. The V_{max} values for SlHAK20$_{Hap1}$ (76.5 ± 3.9 nmol / 10^7 cells / h) and SlHAK20$_{Hap2}$ (V_{max} = 74.1 ± 3.2 nmol / 10^7 cells / h) are comparable. Taken together

the results shown in Figure 3G and 3H, we conclude that SlHAK20$_{Hap1}$ is a more active allele than SlHAK20$_{Hap2}$.

4) From the plot shown on Figure 3H, Km and Vmax cannot be calculated. In addition, I do not understand what is the label of the Y axis (Na uptake (µM). Is it a rate of uptake or the external Na concentration? Na uptake is not a concentration of Na but the amount of Na accumulated within the cell per unit of time and unit of yeast weight. I do not understand either how the Km and Vmax values provided on Figure 3H have been obtained.

Response: Following your suggestion, we have measured the Na^+ uptake kinetics by using different concentrations of NaCl, and the kinetics result is shown in Figure 3H in the revision. We have re-organized the data in the original Figure 3H (now Figure 3G in this revision) showing higher Na^+ transport activity of SlHAK20$_{Hap1}$ than SlHAK20$_{Hap2}$. The Y-axis in the Figure 3G in this revision is the Na^+ concentration in the culture medium. The initial Na^+ present in the medium is 60 µM and the concentration decreases over the culture time due to the uptake of Na^+ by the yeast cells.

5) Point 2 is revised satisfactorily

Response: Thanks!

6) Point 3. The authors have described Na and K fluxes by invasive, scanning ion-selective electrode technique. This is OK to characterize ion fluxes at the root. The results show that the slhak20 mutant lines show decreased Na efflux and the OX line increased Na efflux.

Response: Thanks!

7) Point 4, 5,6 and 7 are revised satisfactorily

Response: Thanks!

Referee #2:

The authors perform GWAS for salt tolerance (measured as Na+/K+ content in roots) in a population of 369 tomato accessions, including cultivated tomato and its closest wild relative S. pimpinellifolium. The mayor GWAS peak is close to SlHAK20, a locus coding for a high affinity potassium transporter. The gene contains several coding and regulatory polymorphisms differentiating accessions with high and low salt tolerance. CRISPR editing or overexpression of SlHAK20 decreases and increases salt tolerance respectively, proving that this locus is involved in stress tolerance in tomato. The findings in this work are very exciting and of high interest for fundamental and applied plant science.
After revision, I have no mayor comments on the manuscript, although I still have some minor comments:
1) I still have problems with the closing sentence of the "A variation in SlHAK20 is associated with root Na+/K+ ratio in tomato during salt stress" section (line 195): "These results clearly indicate that this natural variation in SlHAK20 is the causal sites for the difference in Na+/K+ ratios among the accessions". The authors do not have experiments that show that "this" variation at the SlHAK20 locus is causal for the phenotype. It could be any of the mutations in the gene or in the promoter (or maybe even in the introns?). In order to prove that a specific variation at SlHAK20 is causal, they will need to transform a mutant line (the CRISPR line for example) with an allele that carries single and combinations of mutations until they find the specific mutation (or combination) that affects the phenotype. For this reason I think the best is to change the sentence to: "These results indicate that natural variation at the SlHAK20 locus is strongly associated with the difference in Na+/K+ ratios among the accessions".

Response: We appreciate your suggestion. This sentence is now changed as suggested (Line 193 - 195).

2) In this same section and in Figure 2: I think it is not proper to call "haplotype" to a group of alleles that share a similar polymorphism. It gives the impression that all alleles in Hap1 have the same sequence although this is clearly not the case, and they only share indel48. Could you call them group1 / group2, or indel48+ and indel48-, or something along these lines?

Response: Thanks for your suggestion! We think it is more convenient and less confusing to use Hap1 and Hap2 to represent the two haplotype groups among the 369 accessions based on the indel48. This kind of designation was also used by Wang et al., 2016 (doi:10.1038/ng.3636).

3) Related to the above, in all that refers to Figure 3 is necessary that the authors state in the text what are the alleles behind Hap1 and Hap2, and what are the differences between them: Is it only indel48, or are there more differences?

Response: Thanks for the suggestion and question! The indel 48 is the only difference between Hap1 and Hap2. In Figure 3, the Hap1 and Hap2 were used to represent the coding sequences of *SlHAK20*$_{TS\text{-}21}$ and *SlHAK20*$_{TS\text{-}670}$, respectively. The amino acid sequence of these two alleles are shown in the Dataset EV4. We also added a sentence to clarify this in the revised manuscript (Line 184 - 185).

Referee #3:

The authors did a good job addressing all reviewers' concerns and I have no more critical questions about this MS. The authors should be congratulated with producing a very nice and interesting piece of work.

Response: Thank you!

3rd Editorial Decision 11th Feb 2020

Thank you for submitting a revised version of your manuscript. Your study has now been seen by reviewer #1, who is satisfied with the included data and now supports publication of the manuscript. There now remain only a couple of editorial issues that have to be addressed before I can extend formal acceptance of the manuscript.

REFEREE REPORTS:

Referee #1:

I consider that the revised manuscript successfully addressed all the questions that I raised in my review. Now, the data clearly show differences in Na uptake kinetics between the two haplotypes of SlHAK20 and the methodology has been improved.

3rd Revision - authors' response 12th Feb 2020

The authors performed the requested editorial changes.

4th Editorial Decision 13th Feb 2020

Thank you for implementing the final editorial revisions in your manuscript. I am now pleased to inform you that your manuscript has been accepted for publication in The EMBO Journal.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name:	
Journal Submitted to:	
Manuscript Number:	

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be iustified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	There was not a pre-specified effect size.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	No samples were excluded.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	No
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	No
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Yes, we confirmed that the data showed normal distribution.
Is there an estimate of variation within each group of data?	The variation has been shown by standard deviation (SD)

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repor

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tumo

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Is the variance similar between the groups that are being statistically compared?	Yes

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	GFP antibody (No.11814460001, Roche), Actin antibody (M20010L,Abmart)
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	NA
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	NA
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	No,our study does not fall under dual use research restrictions.