

# SCIENTIFIC REPORTS



OPEN

## Yield-related salinity tolerance traits identified in a nested association mapping (NAM) population of wild barley

Stephanie Saade<sup>1</sup>, Andreas Maurer<sup>2</sup>, Mohammed Shahid<sup>3</sup>, Helena Oakey<sup>1</sup>, Sandra M. Schmöckel<sup>1</sup>, Sónia Negrão<sup>1</sup>, Klaus Pillen<sup>2</sup> & Mark Tester<sup>1</sup>

Received: 03 May 2016  
Accepted: 03 August 2016  
Published: 02 September 2016

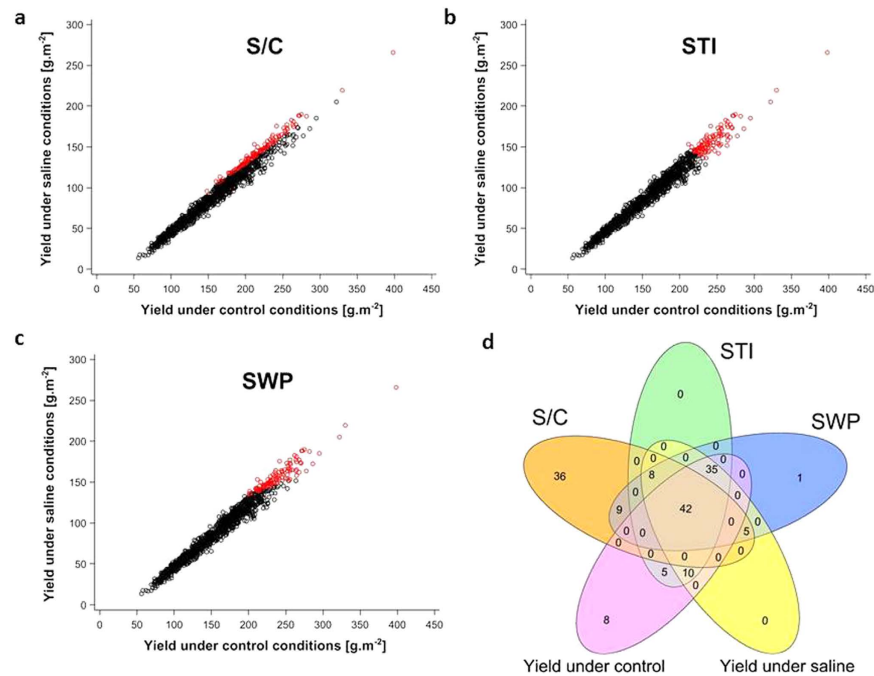
Producing sufficient food for nine billion people by 2050 will be constrained by soil salinity, especially in irrigated systems. To improve crop yield, greater understanding of the genetic control of traits contributing to salinity tolerance in the field is needed. Here, we exploit natural variation in exotic germplasm by taking a genome-wide association approach to a new nested association mapping population of barley called HEB-25. The large population (1,336 genotypes) allowed cross-validation of loci, which, along with two years of phenotypic data collected from plants irrigated with fresh and saline water, improved statistical power. We dissect the genetic architecture of flowering time under high salinity and we present genes putatively affecting this trait and salinity tolerance. In addition, we identify a locus on chromosome 2H where, under saline conditions, lines homozygous for the wild allele yielded 30% more than did lines homozygous for the Barke allele. Introgressing this wild allele into elite cultivars could markedly improve yield under saline conditions.

Over one billion hectares of the world's land are affected by soil salinity, a major constraint on agricultural production<sup>1</sup>. Salinity tolerance in plants is under complex polygenic control and several genes have been proposed to be involved with salinity-tolerance traits<sup>2</sup>. However, what matters to the farmer and brings economic benefit to the agricultural sector is a plant's tolerance to high salinity under field conditions during the reproductive stage.

In addition to its economic importance, barley is a suitable model for studying salinity tolerance in plants because it is the most salt tolerant of the cereal crops<sup>3</sup>. In this study, we used a nested association mapping (NAM) population of barley called "HEB-25". HEB-25 was developed using wild barley (*Hordeum vulgare* ssp. *spontaneum*) donors from the Fertile Crescent<sup>4</sup> that are likely adapted to high salinity and thus are likely to harbor salinity-tolerance alleles that were not eroded during domestication. At the same time, approximately 72% of HEB-25's genome comes from the German elite cultivar Barke<sup>4</sup>, rendering it capable of reaching the reproductive stage and producing high yields in the field.

Here, we present a genome-wide association study (GWAS) in the wild barley NAM population HEB-25. This population, comprising 1,336 individuals belonging to 25 families (Supplementary Table 1), was evaluated for ten traits related to agronomic performance and yield in the field for two consecutive years (Supplementary Table 2) to identify loci associated with salinity tolerance using single nucleotide polymorphism (SNP) information. Plants were stressed with salt via drip irrigation in the sandy soil and dry environment of our field site, located at the International Center for Biosaline Agriculture (ICBA) in Dubai, UAE. This controlled, yet field-based, environment has for the first time, enabled the quantification of the effects of salinity on a large number of yield-related traits such as flowering time and the harvest index in a large mapping population. We dissect the genetic architecture of flowering time under high salinity and suggest putative genes underlying the observed quantitative trait loci (QTL). In addition to the loci identified previously, we identify a strong exotic QTL, associated with a favorable effect on yield under high salinity. The favorable effect of this QTL is due to the wild allele carried by lines in

<sup>1</sup>King Abdullah University of Science and Technology (KAUST), Biological and Environmental Sciences and Engineering (BESE), Thuwal, 23955-6900, Saudi Arabia. <sup>2</sup>Institute of Agricultural and Nutritional Sciences, Martin Luther University Halle-Wittenberg, Betty-Heimann-Str. 3, 06120 Halle, Germany. <sup>3</sup>International Center for Biosaline Agriculture (ICBA), Dubai, United Arab Emirates. Correspondence and requests for materials should be addressed to M.T. (email: mark.tester@kaust.edu.sa)



**Figure 1.** Comparison of the performance of three stress-tolerance indices, S/C, STI and SWP, using yield under saline conditions as a function of yield under control conditions. Red circles indicate the best performing 100 HEB lines identified by each stress-tolerance index ((a) S/C, (b) STI, (c) SWP) when lines are sorted by descending order for the index. Black dots indicate the remaining lines. (d) Venn diagram of shared/unique numbers of lines (out of the 100 top lines for each stress-tolerance index when lines are sorted in descending order) among the three indices and yield under saline and control conditions: numbers in the outer part of each oval correspond to the unique number of lines for each index while numbers in overlapping areas correspond to the number of shared lines that are identified by each index.

HEB family 1. Altogether, we describe the power of a barley NAM population developed using wild relatives of barley to locate yield-related salinity tolerance loci. We demonstrate that HEB-25 may be a valuable resource to develop improved elite barley cultivars that are better adapted to yield well under high salinity.

## Results and Discussion

The wild barley NAM population HEB-25 (Supplementary Table 1) was evaluated for ten traits related to agronomic performance and yield at ICBA (Supplementary Table 2). Plants were irrigated with low saline water, referred to hereafter as the control condition, or with saline water, referred to hereafter as the saline condition (Supplementary Figure 1). In addition to the ten traits measured in the field, we also derived indices for salinity tolerance (Supplementary Tables 3 and 4). The definition of what is desirable in regard to salinity tolerance in the field is not trivial. While lines that are salt tolerant *per se* are desirable, lines that are simultaneously capable of high yields under non-stress conditions are of practical agricultural importance. To differentiate lines in terms of stress tolerance and performance, stress-tolerance indices have previously been proposed<sup>5,6</sup>. For example, S/C is a simple index that indicates the tolerance of the plant to the stress regardless of the plant's performance under non-stress conditions<sup>5</sup>. Another stress-tolerance index, STI, proposed by Fernandez<sup>6</sup>, enables identification of lines that are both stress tolerant and yield highly.

We propose a new index, stress-weighted performance (SWP, Supplementary Tables 3 and 4), which is capable of differentiating the top 100 high-yielding and salt-tolerant plants from other lines (Fig. 1). In SWP, lines with higher values are better performers under control conditions and are more salt tolerant than are lines with smaller values. SWP selects for lines that have above-average S/C and also have high yield under control conditions. This can be visualized in Fig. 1c, where lines with the top 100 values for this index have been colored red. SWP outperforms STI in stress-tolerance research because some lines that do not perform well but have above average S/C scores are identified by STI (Fig. 1b). Although STI captures the best lines based on yield performance (high yield), SWP captures the more salt-tolerant lines (Fig. 1c). In addition to the 42 lines common to all three indices seen in Fig. 1d, the salinity-tolerance index, S/C, has other lines in common with SWP, but not with STI. To achieve an improved identification of lines that are capable of high yield and also more tolerant to saline conditions, we thus advocate the use of SWP.

Salinity significantly reduces all traits measured in our field trials relative to control conditions (ANOVA,  $p < 2e-16$  for all traits, Table 1). On average, yield is reduced by 43% under saline compared to control conditions.

There is a strong correlation ( $|r| > 0.8$ ) between the control and saline conditions in seven of the ten traits (flowering time, maturity time, ear number per plant, grain number per ear, dry mass per  $m^2$ , yield and harvest index). Only ripening period, plant height, and thousand grain mass do not exhibit this correlation (Fig. 2

Trait <sup>a</sup>	Condition	Barke	HEB-25 population		
		Mean	Mean	SE <sup>b</sup>	CV (%) <sup>c</sup>
HEA (days)	Control	80.6	82.6	0.221	9.79
	Saline	77.2	79.3	0.220	10.1
MAT (days)	Control	110	112	0.163	5.35
	Saline	104	106	0.193	6.64
RIP (days)	Control	29.4	29.1	0.0827	10.4
	Saline	26.3	26.8	0.0690	9.40
HEI (cm)	Control	75.1	75.1	0.334	16.3
	Saline	66.4	63.1	0.202	11.7
TGW (g)	Control	40.3	37.8	0.140	13.5
	Saline	35.3	33.2	0.115	12.7
EAR	Control	4.84	3.94	0.0270	25.0
	Saline	3.77	2.50	0.0158	23.1
GPE	Control	20.0	14.5	0.0885	22.3
	Saline	15.9	10.6	0.0746	25.8
DRY_WT (g.m <sup>-2</sup> )	Control	627	511	1.59	11.4
	Saline	521	422	1.48	12.9
YLD (g.m <sup>-2</sup> )	Control	254	157	1.27	29.5
	Saline	165	89.3	0.978	40.0
HI	Control	0.404	0.305	0.00196	23.6
	Saline	0.320	0.210	0.00194	33.8

**Table 1. Agronomic performance of HEB-25 lines and Barke under control and saline conditions.**

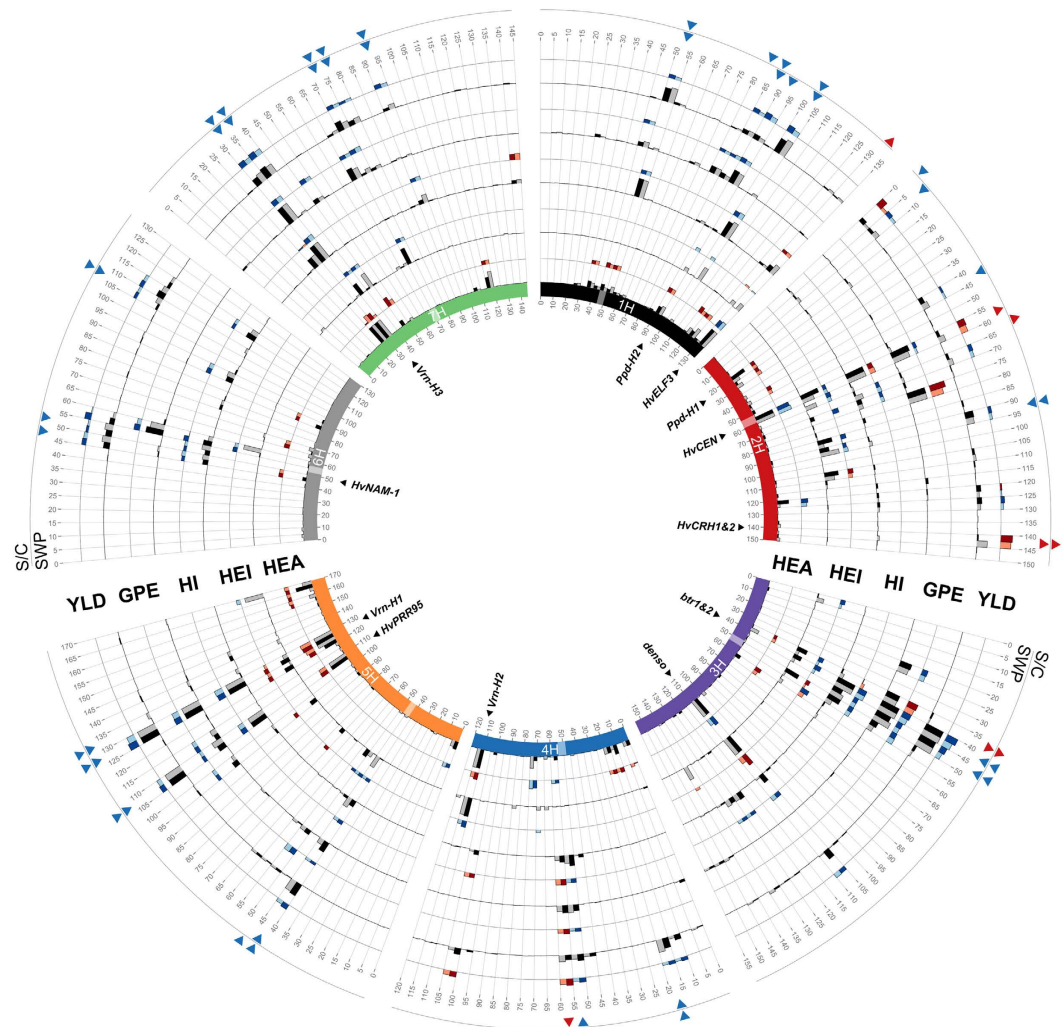
<sup>a</sup>Flowering time (HEA), maturity time (MAT), ripening period (RIP), plant height (HEI), thousand grain mass (TGW), ear number per plant (EAR), grain number per ear (GPE), dry mass per m<sup>2</sup> (DRY\_WT), yield (YLD), and harvest index (HI). <sup>b</sup>SE stands for standard error. <sup>c</sup>CV stands for coefficient of variation.

		Plants grown under control conditions									
Trait <sup>a</sup>	Trait <sup>a</sup>	HEA	MAT	RIP	HEI	TGW	EAR	GPE	DRY_WT	YLD	HI
		Plants grown under saline conditions	HEA	0.942	0.952	-0.795	-0.165	-0.321	-0.305	-0.348	-0.0851
MAT	0.953		0.946	-0.57	-0.125	-0.291	-0.314	-0.365	-0.0916	-0.387	-0.445
RIP	-0.527		-0.243	0.717	0.194	0.283	0.194	0.209	0.0467	0.25	0.294
HEI	-0.143		-0.132	0.0886	0.508	0.211	0.0763	0.146	0.184	0.0493	-0.0233
TGW	-0.298		-0.277	0.179	0.315	0.444	0.0991	0.0773	0.119	0.192	0.185
EAR	-0.4		-0.41	0.131	0.0987	0.143	0.876	0.735	0.425	0.69	0.677
GPE	-0.378		-0.386	0.127	0.141	0.117	0.783	0.991	0.501	0.79	0.763
DRY_WT	-0.0459		-0.0409	0.0321	0.216	0.144	0.446	0.489	0.987	0.656	0.348
YLD	-0.422		-0.422	0.166	0.11	0.232	0.763	0.815	0.603	0.987	0.93
HI	-0.485		-0.487	0.187	0.057	0.229	0.75	0.796	0.351	0.952	0.981

**Figure 2. Pearson correlations between the ten studied traits.** Correlations under control conditions are above the diagonal and correlations under saline conditions are below the diagonal. A heat map is used to color these correlations: blue indicates negative correlations, red indicates positive correlations, and the color intensity indicates the strength of the correlation (the darker the color the stronger the correlation). Within trait correlations between control and saline conditions are placed on the diagonal axis (white cells). All correlations are significant ( $p < 0.05$ ) except for those in the yellow cells. <sup>a</sup>Flowering time (HEA), maturity time (MAT), ripening period (RIP), plant height (HEI), thousand grain mass (TGW), ear number per plant (EAR), grain number per ear (GPE), dry mass per m<sup>2</sup> (DRY\_WT), yield (YLD), and harvest index (HI).

Supplementary Notes). This is reflected in the results presented in Fig. 3 where few quantitative trait loci (QTL) are shown to be treatment specific and most are significant under both control and saline conditions.

We observed loci controlling the flowering time trait with differing effects under control versus saline conditions, suggesting that these loci influence salinity tolerance. The significance of this observation is that flowering time negatively correlates with yield, yield components, and yield-derived traits (Fig. 2, Supplementary Notes).



**Figure 3. Genetic architecture of five of the studied traits: flowering time (HEA), plant height (HEI), harvest index (HI), grain number per ear (GPE), and yield (YLD).** The data in this Circos plot results from 100 cross-validated (20 times 5-fold) GWAS runs performed within each treatment for the studied traits. Barley chromosomes are shown on the inner circle with different colors and centromeres are indicated with transparent boxes. For each trait, the first (inner) track represents the frequency of QTL detection in a 5-cM window while the outer track represents the effect of this QTL. The maximum height of the effect bars for each trait are 9 days for HEA, 17 cm for HEI, 0.1 for HI, 5 grains for GPE and  $60 \text{ g} \cdot \text{m}^{-2}$  for YLD. Window positions (in cM, following Maurer *et al.*<sup>4</sup>) are ordered clockwise per chromosome. In the inner track, QTL appearing under control and saline conditions are represented with black and gray bars, respectively. The effect of the QTL conferred by the wild allele relative to Barke is represented on the outer track, where blue and red bars indicate decreasing and increasing wild barley QTL effects, respectively for each treatment. On the two outermost circles, significant QTL for stress tolerance indices (SWP and S/C, from inside to outside) are shown. Blue arrows pointing inwards and red arrows pointing outwards indicate decreasing and increasing effects from the wild barley alleles, respectively. Candidate genes, potentially explaining the observed QTL effects, are indicated inside the inner circle.

Under our conditions, salinity accelerates flowering by an average of 3.3 days. Here, we identify the genetic basis for this earlier flowering, which is associated with higher salinity tolerance.

Loci on chromosomes 1H (at 130–135 cM; cM positioning throughout this study follows Maurer *et al.*<sup>4</sup>) and 2H (at 55–60 cM), where *HvELF3*<sup>7</sup> and *HvCEN*<sup>8</sup> are respectively located, are significantly associated with flowering time under both control and saline conditions. The wild allele at the *HvELF3* locus causes earlier flowering and maturity under both control and saline conditions and increases the harvest index. The effect of the wild allele on promoting flowering under saline conditions is larger than it is under control conditions. Also, the extension of the ripening period caused by the wild allele is less pronounced under saline stress, allowing plants to complete their life cycles faster. We observe that the wild allele at the *HvELF3* locus decreases plant height and increases thousand grain mass under saline conditions only, which results in a significant increase in S/C, the salinity-tolerance index.

Condition	Genotype at SNP <sup>a</sup>	BOPA2_12_30822 <sup>b</sup> at 2H, 140.8 cM		SCRI_RS_116590 <sup>c</sup> at 2H, 140.8 cM	
		Yield average (g.m <sup>-2</sup> ) ± standard deviation	p value <sup>d</sup>	Yield average (g.m <sup>-2</sup> ) ± standard deviation	p value <sup>d</sup>
Control	0	180 ± 39.7	0.00323	158 ± 45.8	0.163
	1	205 ± 19.3		150 ± 42.4	
	2	216 ± 43.3		153 ± 48.8	
Saline	0	107 ± 31.6	0.0009	89.6 ± 35.1	0.145
	1	128 ± 16.4		84.2 ± 33.6	
	2	139 ± 34.8		86.3 ± 38.1	

**Table 2. Comparison of yield performance between lines by field condition (control and saline) and by genotype at SNPs BOPA2\_12\_30822 and SCRI\_RS\_116590.** <sup>a</sup>Values 0 and 2 indicate HEB lines homozygous for Barke and wild barley alleles, respectively, and value 1 indicates heterozygous HEB lines. <sup>b</sup>BOPA2\_12\_30822 only segregates in HEB family 1, with 29 vs 23 vs 3 lines carrying the genotypes homozygous Barke, homozygous wild barley and heterozygous, respectively. <sup>c</sup>SCRI\_RS\_116590 segregates in 24 out of 25 HEB families (i.e. all except HEB family 24), with 868 vs 317 vs 91 lines carrying the genotypes homozygous Barke, homozygous wild barley, and heterozygous, respectively. <sup>d</sup>p value of the t-test to check differences in yield average between lines carrying the genotype homozygous Barke and lines carrying the genotype homozygous wild barley.

Similarly, the wild allele at the *HvCEN* locus promotes earlier flowering (by approx. 6 days) and earlier maturity (by approx. 4 days) under both control and saline conditions. It also reduces plant height and dry mass per m<sup>2</sup>. This indicates that the reduced resources invested in vegetative tissue may enable the increased number of ears and increased yield (by 21 g.m<sup>-2</sup>). The reduction in grain number per ear could be related to the production of bigger grains per ear or more tillers per plant. The effect of the wild allele on reducing dry mass per m<sup>2</sup> and increasing yield results in an overall increasing effect on the harvest index. We observe that the salinity tolerance index, SWP, but not S/C, is significantly increased at the *HvCEN* locus, indicating that this locus is associated more with better yield performance than with increased salinity tolerance. We conclude that *HvCEN* may be of great interest to breeders because of its yield improvement under both control and saline conditions.

In addition to loci that improve yield by accelerating flowering, we found a strong QTL on chromosome 2H, 140–145 cM with a direct favorable effect on yield (increasing by 33 g.m<sup>-2</sup>, 21% under control conditions and 37% under saline conditions) that is conferred by the wild barley allele. This QTL appears in cross-validations 56 times under salinity and eight times under control conditions. The SNP showing the significant yield effect is BOPA2\_12\_30822 in the *alpha-glucosidase* gene (AK375658). This SNP segregates only in HEB family 1, explaining 19.8% of the phenotypic variation in yield under high salinity in this family. In HEB family 1, lines carrying the wild barley allele at this SNP have a significantly higher yield under both control and saline conditions than do lines carrying the Barke allele at this SNP (Table 2). Both S/C and SWP indices are significant, revealing a favorable effect of the wild allele at this locus. Interestingly, this QTL is only significant for yield, which implies that each yield component only slightly increases. Moreover, because both the S/C and SWP indices for yield at this QTL are also significant, we suggest that this QTL increases the salinity-tolerance component of SWP. Hence, the wild barley donor of HEB family 1, HID\_003, seems to confer an allele that increases yield under high salinity. However, the A2148G substitution that is detected by the SNP does not affect the amino acid sequence. A second SNP, SCRI\_RS\_116590, is located at the same position as BOPA2\_12\_30822 (2H, 140.8 cM) and segregates in 24 out of 25 HEB families. However, the yield effect that we observe between lines homozygous for wild and Barke alleles at BOPA2\_12\_30822 is not observed at SCRI\_RS\_116590 (Table 2). This result indicates that the favorable yield effect observed with BOPA2\_12\_30822 may be restricted to donor HID\_003, originating from northeastern Iraq.

Genes related to the barley *alpha-glucosidase* gene in this QTL have been studied in other species. It has been shown that the Arabidopsis *alpha-glucosidase* gene is involved in cell wall biosynthesis<sup>9</sup> and in the first steps of N-glycan trimming in the endoplasmic reticulum<sup>10</sup>. Furthermore, a study of knockout mutants of Arabidopsis *alpha-glucosidase* suggested that the N-glycosylation of this gene was required for salinity tolerance<sup>11</sup>. Further experimental evidence, involving manipulation of *alpha-glucosidase* expression, is required to test whether this gene is the main cause of the favorable yield effect observed in our study. Alternatively, this gene could be linked to the causative gene responsible for the phenotype observed in family 1. Interestingly, barley calreticulin 1 (*CRH1*, GenBank: L27348.1) and calreticulin 2 (*CRH2*, GenBank: L27349.1), which are calcium-binding proteins, align to position 141.78 cM and may instead be the genes underlying this QTL (2H, 140–145 cM). Xiang *et al.*<sup>12</sup> have shown that tobacco plants transformed with *Triticum aestivum* calreticulin gene (*TaCRT1*) showed higher salinity tolerance compared to their wild-type counterparts. Another candidate in the region of interest is a choline-transporter-like gene (MLOC\_63287.1). Enzymes mediating choline oxidoreduction have also been shown to be involved in salinity tolerance<sup>2</sup>. Choline is converted by choline dehydrogenase/oxidase to glycine betaine<sup>13,14</sup>, an osmoprotectant known to increase salinity tolerance<sup>2</sup>.

In several studies, analyses of transcriptomes of wild and cultivated barley grown under control and saline conditions have been conducted to identify differential gene expression in response to salinity<sup>15–20</sup>. Several members of ubiquitous gene families that were reported to be differentially expressed in those transcriptomic studies, could be located within the QTL region on chromosome 2H, 140–145 cM. These genes code for:

pentatricopeptide repeat domain-containing proteins, RNA recognition motif containing proteins, zinc finger proteins, MYB family transcription factors, heat shock hsp70 proteins and cytochrome P450. However, no further evidence is available that would support a direct involvement in the salinity tolerance of HEB family 1. Also, it needs to be remembered that genes whose expression changes in response to salinity stress may not necessarily contribute to salinity tolerance.

In addition, Wu *et al.*<sup>21</sup> compared the ionic and proteomic responses of a wild barley genotype, XZ16, and a more salt-tolerant cultivar, CM72, to salinity under hydroponic conditions. The authors found that putative class III peroxidases (GenBank: AK365672 and AK363236) were significantly upregulated in the roots of both genotypes. Based on barley genome sequence blasting, we could map these genes to chromosome 2H, 146.5cM. Given the known role of reactive oxygen species (ROS) and the ROS-scavenging system in salinity tolerance, it is conceivable that the putative class III peroxidase genes are candidates that could explain the function of the QTL on chromosome 2H, 140–145 cM. Other loci controlling yield are identified in the Supplementary notes (included in the Supplementary Information).

Using our new stress-tolerance index, SWP, on HEB-25, we identified genes putatively affecting flowering time and salinity tolerance in barley. We also located a QTL responsible for higher yield. Further studies are necessary to validate the candidate genes and to study in more depth the mechanisms leading to increased yield and salinity tolerance with the ultimate goal of introgressing salinity tolerance traits (loci) into commercial barley lines. Studying the causal genes in other crops, notably wheat, also provides significant opportunities to improve performance of major crops in saline conditions, guided by this work in the relatively salt-tolerant crop, barley.

## Materials and Methods

**Plant material.** HEB-25, a barley NAM population, comprises 25 families resulting from crosses of the German spring cultivar Barke with 25 wild donors, 24 *H. vulgare* ssp. *spontaneum* (*Hsp*) from the Fertile Crescent and one Tibetan *H. vulgare* ssp. *agriocrithon* (*Hag*) accession. The resulting F<sub>1</sub> plants were backcrossed with Barke as a female parent and then selfed three times through the application of a single seed descent procedure. Subsequently, BC<sub>1</sub>S<sub>3</sub> lines were bulk propagated to BC<sub>1</sub>S<sub>3,6</sub>. Further details on HEB-25 are provided in Maurer *et al.*<sup>4</sup>. In this study, 1,336 BC<sub>1</sub>S<sub>3,7</sub> lines of HEB-25 (Supplementary Table 1) were tested in the field under control and saline conditions.

**Field trials.** A field trial was conducted for two years at the International Center for Biosaline Agriculture (ICBA), Dubai, United Arab Emirates (N 25° 05.847; E 055° 23.464), from December to May (2013–2014 and 2014–2015).

The soil at ICBA is a fine sand (sand 98%, silt 1%, and clay 1%), calcareous (50–60% CaCO<sub>3</sub> equivalents), porous (45% porosity), and moderately alkaline (pH 8.22). The saturation percentage of the soil is 26; it has a high drainage capacity; and the electrical conductivity of its saturated extract (ECe) is 1.2 dS.m<sup>-1</sup>. ICBA soil is classified as Typic Torripsamments, carbonatic, and hyperthermic<sup>22,23</sup>. Koblenz Organic Fertilizer (manufactured by Tadweer Waste Treatments LLC (Dubai, UAE) was added to the top of the soil at a rate of 40 tonnes FW.ha<sup>-1</sup> (at 85% moisture) to increase the soil water-holding capacity and to provide some of the required nitrogen (N), potassium (K), sulfur (S), and micro-nutrients. Two weeks before sowing, fertilization with phosphorus (P) was conducted using a single supply of 45 kg P<sub>2</sub>O<sub>5</sub> ha<sup>-1</sup> (100 kg of Triple Super Phosphate from Benedict (Indonesia) with a total of 45% phosphoric anhydride). Granular urea nitrogen (N) fertilizer from Fertil (Abu Dhabi, UAE) was applied once at a rate of 30 kg N ha<sup>-1</sup> (i.e. 3 g.m<sup>-2</sup> pure N, equivalent to 7.5 g urea per plot at 46% N), three weeks after planting. An application of NPK fertilizer (20-20-20) from ADFERT (Abu Dhabi, UAE) at a rate of 30 kg.ha<sup>-1</sup> was also made six weeks after planting by fertigation.

The 1,336 NAM lines were grown, along with Barke and the 25 wild donors, irrigated with saline (referred to as the saline condition throughout the paper) and low-saline water (referred to as the control condition throughout the paper). Plots were randomized in an augmented design, and salt-tolerant check lines (116/2A, 58/1A, CM72) were sown every seven plots on average. Four rows of a local barley cultivar (58/1A) were sown around the experimental area to reduce edge effects. The plot size was 1.5 m × 0.5 m, containing five rows, each 25 cm apart. Approximately 25 seeds (at 2 cm spacing and 1 cm depth) were hand-sown per row (Supplementary Figure 1).

The plots were irrigated twice per day for ten minutes each time. Each plot received 13.3 L water per day. Control plots were irrigated with water of 1 dS.m<sup>-1</sup>; and saline plots were irrigated with 1 dS.m<sup>-1</sup> water during the first week followed by irrigation with saline water (17 dS.m<sup>-1</sup>) for the remainder of the growing period. The ionic composition of the irrigation water was monitored throughout the season using salinity sensors. The distribution of drippers was homogeneous and the distance between drippers allowed the overlapping of wetting fronts. Weather data was recorded at the field site at ICBA during the two years of field trials (Supplementary Table 5).

Ten agronomic traits were measured under control and saline conditions (Supplementary Table 2) and stress-tolerance indices were derived from those traits (Supplementary Table 3).

**Genotyping of HEB-25.** Genotyping of HEB-25, along with the parents, was performed using an Illumina Infinium iSelect HD 9k chip consisting of 7,864 SNPs as previously described<sup>8</sup>. The 5,398 informative SNPs that were polymorphic in at least one HEB family and that met certain quality criteria (<10% missing, <12.5% heterozygous, and not in complete linkage disequilibrium (LD) to another SNP in the set) were kept for further analyses. Further details on the genotypic data are available in Maurer *et al.*<sup>4</sup>.

**Statistical analysis of phenotypic data.** To correct for spatial variation in the field, a multi-environment trial (MET) analysis was conducted for two years on each trait and predicted line means adjusted for environmental variation were obtained (Supplementary Dataset 1). In the MET analysis, each year by treatment combination was considered as a separate environment or trial, with these trials correlated in a MET analysis following the factor analytic model of Smith *et al.*<sup>24</sup>. The environmental variation often present in field trials was modeled according to the work of Gilmour *et al.*<sup>25</sup> that allowed for the three possible sources of environmental variation: global, extraneous, and local. The need for adjustment for environmental variation was explored by examination of residual plots including a variogram. The environmental terms identified were then added to the MET analysis model.

Consider a data set consisting of  $\nu$  lines and  $s$  trials, the latter of which correspond to the two-treatment-by-two-year combinations with  $\mathbf{y}$  the vector of response, an appropriate linear mixed model for a MET analysis is as follows:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\tau} + \mathbf{Z}_g\mathbf{g} + \mathbf{Z}_u\mathbf{u} + \boldsymbol{\varepsilon},$$

where  $\mathbf{y}^{(n \times 1)} = (\mathbf{y}_1^T, \dots, \mathbf{y}_s^T)^T$  and  $\mathbf{y}_t^{(n_t \times 1)}$  is the vector of response for trial  $t$  and  $n = \sum_{t=1}^s n_t$ , where  $n_t$  is the number of observations (plots) in trial  $t$ ,  $\boldsymbol{\tau}$  is a vector of fixed terms consisting of an overall mean performance for each trial as well as trial specific global or extraneous trial terms such as linear row or linear column trends,  $\mathbf{X}$  is the associated design matrix,  $\mathbf{g}^{(\nu s \times 1)}$  is the vector of random line effects of the  $\nu$  lines in each of the  $s$  trials with design matrix  $\mathbf{Z}_g$  and associated variance  $\text{var}(\mathbf{g}) = \mathbf{G} \otimes \mathbf{I}_\nu$ , where  $\mathbf{I}_\nu$  is the  $(\nu \times \nu)$  identity matrix,  $\otimes$  is the Kronecker product,  $\mathbf{G}$  is the  $(s \times s)$  genetic variance matrix. Here we consider a factor analytic structure for  $\mathbf{G}^{24}$ , with two factors,  $\mathbf{G} = \mathbf{\Lambda}\mathbf{\Lambda}^T + \boldsymbol{\Psi}$ , with  $\mathbf{\Lambda}$  being a matrix of factor loadings at each of the  $s$  trials, and  $\boldsymbol{\Psi}$  a diagonal matrix with elements relating to specific variances for trial  $s$ . This formulation for  $\mathbf{G}$  allows a separate variance for each trial and a difference covariance between pairs of trials.  $\mathbf{u}$  is a vector of random effects and includes extraneous environmental variation specific to each trial such as random row or column effects or a random smoothing spline<sup>26</sup> and  $\mathbf{Z}_u$  is the associated design matrix. The residual vector,  $\boldsymbol{\varepsilon}^{(n \times 1)} = (\boldsymbol{\varepsilon}_1^T, \dots, \boldsymbol{\varepsilon}_s^T)^T$ , represents local stationary variation at the  $t$ th trial with  $\boldsymbol{\varepsilon}_s^T$  having variance  $\sigma_t^2 \boldsymbol{\Sigma}_t$  for  $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_r \otimes \boldsymbol{\Sigma}_c$ , where  $\boldsymbol{\Sigma}_t$  represents the Kronecker product between auto-regressive processes of order one (AR1) in the column and row directions for the  $t$ th trial.

The generalized heritability was calculated for each trial separately (i.e. assuming  $\mathbf{G} = \boldsymbol{\Psi}$ ) as  $1 - \frac{a}{2\theta_{gt}^2}$  where  $a$  is the average pairwise prediction error variance of line effects and  $\theta_{gt}^2$  is the specific genetic variance of trial  $t$ <sup>27</sup>. Results are shown in Supplementary Table 6. The factor analytic models were fitted in ASReml v3.0-1<sup>28</sup> for R v3.2.0<sup>29</sup>.

A linear mixed model was fitted to calculate best linear unbiased estimates (BLUEs) across years of the adjusted values per plot to take into account the year and line-by-year interactions. The adjusted HEB means across years were used in the GWAS.

Traits and indices were correlated using Pearson correlation, and analysis of variance (ANOVA) was used to study the effect of treatment on each trait.

**Association mapping analysis.** To conduct GWAS, we used the multiple linear regression Model-B of Liu *et al.*<sup>30</sup>, in which cofactors and a population effect were included in addition to the SNP under investigation. This model exhibited high predictive power in previous studies and effectively controlled for population structure when compared to other joint linkage association mapping models<sup>31</sup>. This model has been shown to perform well in the Maurer *et al.*<sup>4</sup> study of flowering time in the HEB-25 population.

The analysis was conducted with SAS 9.4 Software (SAS Institute Inc., Cary, NC, USA) using *Proc GLMSELECT*. Significant SNPs were determined by stepwise forward-backward regression. SNPs were allowed to enter or leave the model at each step if  $p < 0.001$ . To estimate the proportion of phenotypic variance explained by each significant SNP,  $R^2$  was calculated after modelling the SNP solely in a linear model. Additive effects of the wild allele relative to the Barke allele were estimated across families and were taken as the regression coefficient of the SNP from the GWAS model. Multiplication of those additive effects by a factor of two represents the absolute difference between the wild allele effect and the Barke allele, which is referred to throughout the paper.

A five-fold cross-validation was run 20 times to increase the robustness of the GWAS. The 100 subsets were taken from the total phenotypic dataset. Each subset consisted of 80% randomly chosen HEB lines used as a training set to define significant SNPs and the remaining 20% of lines were used as a test set. The phenotype of these latter lines was predicted based on the SNP effects estimated in the training set. 'R<sup>2</sup>val' was then calculated as the squared Pearson product, the moment correlation between predicted and observed phenotypes in the test set, whereas 'R<sup>2</sup>train' represents the model fit of the training set. The detection rate was calculated as the number of times, out of 100, that a SNP showed significance (results in Supplementary Table 7 and Supplementary Dataset 2). To enable comparison of shared QTL regions across cross-validation runs, detection rates were accumulated and SNP effects were averaged within 5 cM windows, as shown in the Circos plot of Fig. 3. Individual Manhattan plots for each trait and condition were created with qqman<sup>32</sup> (Supplementary Figure 2). The  $-\log_{10}(p\text{-values})$  were averaged across all 100 cross-validation runs and the mean was weighed through multiplying by the sum of occurrences (out of 100) and dividing the overall result by 100. The resulting numbers were referred to as cross-validated  $-\log_{10}(p)$  in Supplementary Figure 2. The Bonferroni-adjusted 5% significance threshold was calculated according to Holm<sup>33</sup>. This threshold has been weighed by a factor of 0.2, reflecting a detection rate of 20 out of 100 cross-validation runs.

## References

1. Wicke, B. *et al.* The global technical and economic potential of bioenergy from salt-affected soils. *Energy & Environmental Science* **4**, 2669–2681 (2011).
2. Roy, S. J., Negrão, S. & Tester, M. Salt resistant crop plants. *Curr Opin Biotechnol* **26**, 115–24 (2014).
3. Munns, R. & Tester, M. Mechanisms of salinity tolerance. *Annual Review of Plant Biology* **59**, 651–681 (2008).
4. Maurer, A. *et al.* Modelling the genetic architecture of flowering time control in barley through nested association mapping. *Bmc Genomics* **16** (2015).
5. Munns, R. *et al.* Avenues for increasing salt tolerance of crops, and the role of physiologically based selection traits. *Plant and Soil* **247**, 93–105 (2002).
6. Fernandez, G. C. J. Effective selection criteria for assessing plant stress tolerance. in *Proceedings of the international symposium on adaptation of vegetable and other food crops in temperature and water stress* 257–270 (Taiwan, 1992).
7. Faure, S. *et al.* Mutation at the circadian clock gene *EARLY MATURITY 8* adapts domesticated barley (*Hordeum vulgare*) to short growing seasons. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 8328–8333 (2012).
8. Comadran, J. *et al.* Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat Genet* **44**, 1388–92 (2012).
9. Gillmor, C. S., Poindexter, P., Lorieau, J., Palcic, M. M. & Somerville, C. Alpha-glucosidase I is required for cellulose biosynthesis and morphogenesis in *Arabidopsis*. *Journal of Cell Biology* **156**, 1003–1013 (2002).
10. Boisson, M. *et al.* Arabidopsis glucosidase I mutants reveal a critical role of N-glycan trimming in seed development. *EMBO J* **20**, 1010–9 (2001).
11. Kang, J. S. *et al.* Salt tolerance of *Arabidopsis thaliana* requires maturation of N-glycosylated proteins in the Golgi apparatus. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 5933–5938 (2008).
12. Xiang, Y. *et al.* Overexpression of a *Triticum aestivum* Calreticulin gene (*TaCRT1*) improves salinity tolerance in tobacco. *PLoS One* **10**, e0140591 (2015).
13. Mohanty, A. *et al.* Transgenics of an elite *indica* rice variety Pusa Basmati 1 harbouring the *codA* gene are highly tolerant to salt stress. *Theor Appl Genet* **106**, 51–7 (2002).
14. Holmstrom, K. O., Somersalo, S., Mandal, A., Palva, T. E. & Welin, B. Improved tolerance to salinity and low temperature in transgenic tobacco producing glycine betaine. *J Exp Bot* **51**, 177–85 (2000).
15. Ziemann, M. *et al.* Analysis of the barley leaf transcriptome under salinity stress using mRNA-Seq. *Acta Physiol Plant* **35**, 1915–1924 (2013).
16. Ueda, A., Kathiresan, A., Bennett, J. & Takabe, T. Comparative transcriptome analyses of barley and rice under salt stress. *Theor Appl Genet* **112**, 1286–94 (2006).
17. Walia, H. *et al.* Expression analysis of barley (*Hordeum vulgare* L.) during salinity stress. *Funct Integr Genomics* **6**, 143–56 (2006).
18. Ozturk, Z. N. *et al.* Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. *Plant Mol Biol* **48**, 551–573 (2002).
19. Ueda, A. *et al.* Osmotic stress in barley regulates expression of a different set of genes than salt stress does. *J Exp Bot* **55**, 2213–2218 (2004).
20. Bahieldin, A. *et al.* RNA-Seq analysis of the wild barley (*H. spontaneum*) leaf transcriptome under salt stress. *Cr Biol* **338**, 285–297 (2015).
21. Wu, D. Z. *et al.* Identification of proteins associated with ion homeostasis and salt tolerance in barley. *Proteomics* **14**, 1381–1392 (2014).
22. Soil Survey Staff. *Keys to soil taxonomy*, (United States Dept. of Agriculture, Natural Resources Conservation Service, Washington, D.C., 2010).
23. Shahid, S. A. *United Arab Emirates keys to soil taxonomy*, (Springer, Dordrecht, 2014).
24. Smith, A., Cullis, B. & Thompson, R. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics* **57**, 1138–1147 (2001).
25. Gilmour, A. R., Cullis, B. R. & Verbyla, A. Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agricultural, Biological and Environmental Statistics* **2**, 269–293 (1997).
26. Verbyla, A. P., Cullis, B. R., Kenward, M. G. & Welham, S. J. The analysis of designed experiments and longitudinal data by using smoothing splines. *Journal of the Royal Statistical Society Series C-Applied Statistics* **48**, 269–300 (1999).
27. Cullis, B. R., Smith, A. B. & Coombes, N. E. On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological and Environmental Statistics* **11**, 381–393 (2006).
28. Butler, D., Cullis, B., Gilmour, A. & Gogel, B. (eds). *ASReml R-reference manual*, (VSN International Ltd., Hemel Hempstead, UK, 2009).
29. R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. URL <http://www.R-project.org/> (2015).
30. Liu, W. *et al.* Association mapping in an elite maize breeding population. *Theor Appl Genet* **123**, 847–58 (2011).
31. Wurschum, T. *et al.* Comparison of biometrical models for joint linkage association mapping. *Heredity (Edinburgh)* **108**, 332–40 (2012).
32. Turner, S. D. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *Preprint at* <http://dx.doi.org/10.1101/005165> (2014).
33. Holm, S. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics* **6**, 65–70 (1979).

## Acknowledgements

The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST) and the German Research Foundation (DFG) via the priority program 1530: Flowering time control - from natural variation to crop improvement (Grant Pi339/7-1), via ERA-CAPS (Grant Pi339/8-1) and from funds from the Interdisciplinary Centre for Crop Plant Research (IZN), Halle, Germany. We are grateful to TraitGenetics GmbH, Gatersleben, Germany, for genotyping HEB-25 with the Infinium iSelect 9 k SNP chip. We thank Dr. Ibrahim Salah for his contributions to the design of the experiment in the field. The ICBA team is gratefully acknowledged for managing and phenotyping the field experiments in 2013–2014 and 2014–2015.

## Author Contributions

S.S. performed all data analyses and wrote the manuscript. A.M. and S.S. performed the GWAS and statistical analyses. M.S. supervised the field trials at ICBA and collected the phenotypic data. H.O. performed spatial corrections. S.M.S. helped with data analysis. K.P. developed HEB-25, a NAM population of wild barley, and provided the genotypic data. M.T., K.P. and S.N. contributed to the original concept of the project and supervised the study. M.T. conceived the project and its components and led the design of the field trials. All authors reviewed the manuscript.



### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Saade, S. *et al.* Yield-related salinity tolerance traits identified in a nested association mapping (NAM) population of wild barley. *Sci. Rep.* **6**, 32586; doi: 10.1038/srep32586 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016