












## Barley sodium content is regulated by natural variants of the Na<sup>+</sup> transporter *HvHKT1;5*

Kelly Houston<sup>1</sup>, Jiaen Qiu <sup>2,3</sup>, Stefanie Wege<sup>2,3</sup>, Maria Hrmova <sup>3,4</sup>, Helena Oakey <sup>3</sup>, Yue Qu <sup>2,3</sup>, Pauline Smith<sup>1</sup>, Apriadi Situmorang<sup>2,3</sup>, Malcolm Macaulay<sup>1</sup>, Paulina Flis <sup>5</sup>, Micha Bayer<sup>1</sup>, Stuart Roy <sup>3,6</sup>, Claire Halpin <sup>7</sup>, Joanne Russell<sup>1</sup>, Miriam Schreiber<sup>1</sup>, Caitlin Byrt <sup>2,3,8</sup>, Matt Gilliham <sup>2,3</sup>✉, David E. Salt <sup>5</sup>✉ & Robbie Waugh <sup>1,3,7</sup>✉

During plant growth, sodium (Na<sup>+</sup>) in the soil is transported via the xylem from the root to the shoot. While excess Na<sup>+</sup> is toxic to most plants, non-toxic concentrations have been shown to improve crop yields under certain conditions, such as when soil K<sup>+</sup> is low. We quantified grain Na<sup>+</sup> across a barley genome-wide association study panel grown under non-saline conditions and identified variants of a Class 1 *HIGH-AFFINITY-POTASSIUM-TRANSPORTER* (*HvHKT1;5*)-encoding gene responsible for Na<sup>+</sup> content variation under these conditions. A leucine to proline substitution at position 189 (L189P) in *HvHKT1;5* disturbs its characteristic plasma membrane localisation and disrupts Na<sup>+</sup> transport. Under low and moderate soil Na<sup>+</sup>, genotypes containing *HvHKT1;5*<sub>P189</sub> accumulate high concentrations of Na<sup>+</sup> but exhibit no evidence of toxicity. As the frequency of *HvHKT1;5*<sub>P189</sub> increases significantly in cultivated European germplasm, we cautiously speculate that this non-functional variant may enhance yield potential in non-saline environments, possibly by offsetting limitations of low available K<sup>+</sup>.

<sup>1</sup>Cell and Molecular Sciences, The James Hutton Institute, Errol Road Invergowrie, Dundee DD2 5DA Scotland, UK. <sup>2</sup>ARC Centre of Excellence in Plant Energy Biology, University of Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia. <sup>3</sup>School of Agriculture and Wine & Waite Research Institute, University of Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia. <sup>4</sup>School of Life Science, Huaiyin Normal University, 223300 Huaian, China. <sup>5</sup>Future Food Beacon of Excellence and the School of Biosciences, University of Nottingham, Nottingham NG7 2RD, UK. <sup>6</sup>ARC Industrial Transformation Research Hub for Wheat in a Hot Dry Climate, University of Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia. <sup>7</sup>School of Life Sciences, University of Dundee, Dow Street, Dundee DD1 5EH Scotland, UK. <sup>8</sup>Research School of Biology, 46 Sullivans Creek Road, The Australian National University, Canberra, ACT 2601, Australia. ✉email: [matthew.gilliham@adelaide.edu.au](mailto:matthew.gilliham@adelaide.edu.au); [David.Salt@nottingham.ac.uk](mailto:David.Salt@nottingham.ac.uk); [robbie.waugh@hutton.ac.uk](mailto:robbie.waugh@hutton.ac.uk)

In C3 plants, such as *Arabidopsis*, wheat, rice, and barley, sodium ( $\text{Na}^+$ ) is non-essential for plant growth and development<sup>1</sup>. While halophytes thrive on high  $\text{Na}^+$  containing soils<sup>2</sup>, for glycophytes, including our major cereal crops,  $\text{Na}^+$  becomes toxic when present above certain species-specific threshold concentrations. Intriguingly, many crops, including barley, have been shown to benefit from intermediate (non-toxic) concentrations of  $\text{Na}^+$ , a situation that is particularly evident when levels of  $\text{K}^+$  in the soil are low or unavailable<sup>1,3–10</sup>. Extensive historical evidence supports a requirement for non-toxic concentrations of  $\text{Na}^+$  to achieve maximal biomass growth in a wide range of plants<sup>11</sup>. As such it has been suggested that  $\text{Na}^+$  is a functional nutrient, proposed to be capable of substituting for many of the essential roles that  $\text{K}^+$  ions play in plant nutrition including osmoregulation and enzyme activation<sup>1</sup>.

Despite the reported positive attributes of  $\text{Na}^+$ , by far the majority of studies in the more recent literature focus on the negative impacts of  $\text{Na}^+$  (i.e. salinity) on plant growth<sup>12,13</sup>. These latter investigations generally seek to explore the possible mechanisms that explain how tolerance to excess  $\text{Na}^+$  can be achieved. They commonly revolve around  $\text{Na}^+$  exclusion from the transpiration stream via active removal in the root, the partitioning of excess  $\text{Na}^+$  into root and shoot vacuoles to protect the  $\text{Na}^+$  sensitive photosynthetic tissues of the shoot, or the energy balance associated with active tolerance mechanisms<sup>14,15</sup>. While understanding how to enable crops to grow more efficiently in the expanding saline environments across the globe is highly relevant, it remains important to note that the majority of temperate cereal crop production is achieved on nonsaline soils. Given the demonstrated benefits of  $\text{Na}^+$  as a functional nutrient, we have taken a combined genetic and functional approach to explore the extent and causes of natural variation in  $\text{Na}^+$  content in barley grown in nonsaline soils. Our data lead us to speculate, cautiously, that high  $\text{Na}^+$  accumulation may be a positive trait in the nonsaline conditions typical of high production agricultural environments.

## Results

**Barley grain  $\text{Na}^+$  content is under genetic control.** We used Inductively Coupled Plasma Mass Spectrometry (ICP-MS) to quantify sodium ( $\text{Na}^+$ ) and potassium ( $\text{K}^+$ ) content of whole-grain samples from five biological and five technical replicates of a small barley genome-wide association study (GWAS) panel comprised of 131 elite two-row spring genotypes. Elite barley germplasm displays significant population structure due to both winter/spring growth habit and two- versus six-row inflorescence architecture<sup>16</sup>. By focusing only on the two-row spring barley gene pool we effectively remove confounding effects of underlying population structure<sup>17</sup>. All plants were grown under optimal, nonsaline conditions. We observed an approximate six-fold variation in grain  $\text{Na}^+$  (16.07 to 98.76 ppm) and greater than two-fold variation in  $\text{K}^+$  (2459–5562 ppm) contents (Fig. 1a and Supplementary Data 1). We then used barley 50k iSelect SNP genotypic data collected from all 131 genotypes to conduct GWAS<sup>18</sup>. For grain  $\text{Na}^+$  content we observed a single highly significant association on the bottom of chromosome 4HL for the marker SCRI\_RS\_142792, with a  $-\log_{10}(p) = 11.8$ , FDR adjusted  $p$ -value =  $2.2\text{E-}08$ , and an  $R^2 = 0.45$  (Fig. 1b) consistent with prior genetic analyses of shoot  $\text{Na}^+$  content<sup>19–21</sup>. The moderately high  $R^2$  value implies other loci and mechanisms are also involved in this trait. No significant genetic association was observed for grain  $\text{K}^+$  content at this locus (Fig. 1c). The significantly associated region spanned ~6.6 Mb, from 638,211,825 nt to 644,818,273 nt on the barley 4H physical map<sup>22</sup> and contained 247 gene models (Supplementary Data 2). Common SNPs

between the 50k and 9K iSelect platforms<sup>17,18</sup> aligned this association with a region recently shown to contain *HvHKT1;5*<sup>23</sup>, and in the barley genome sequence HORVU4Hr1G087960, 0.58 kb from the top scoring SNP (SCRI\_RS\_142792), was annotated as a homolog of *OsHKT1;5*. Considering this and previous functional studies<sup>24–27</sup>, we conclude that HORVU4Hr1G087960 is the  $\text{Na}^+$  specific transporter *HvHKT1;5*.

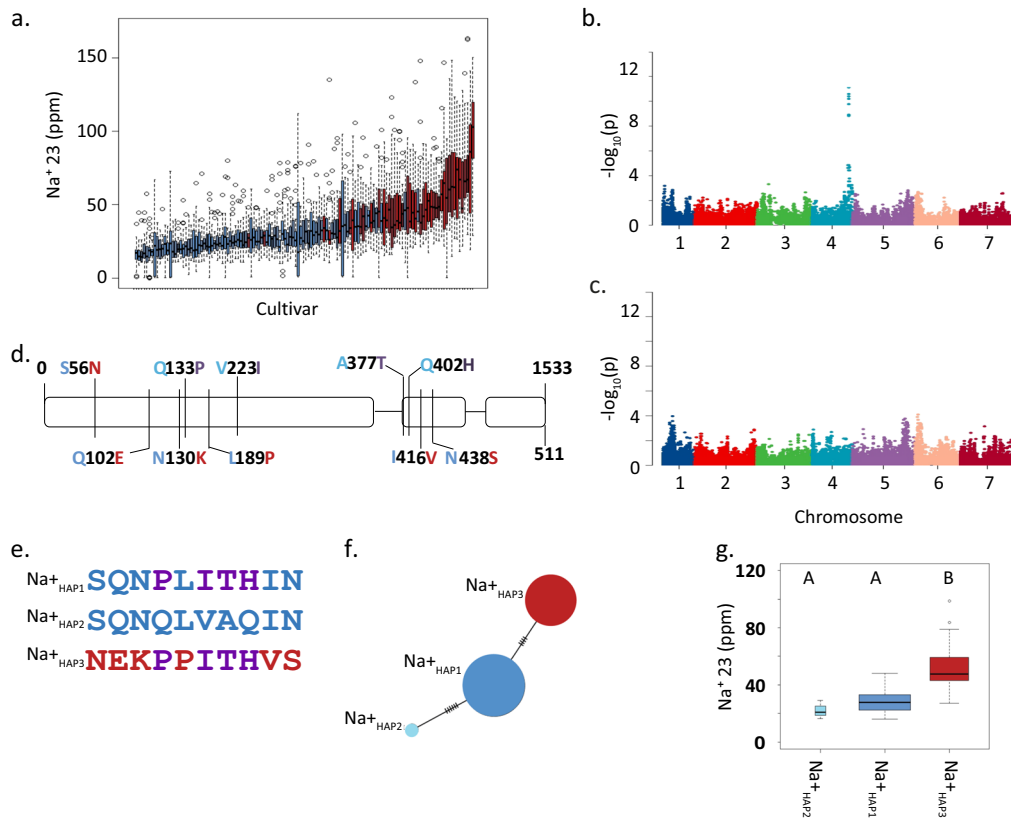
As *HvHKT1;5* is a clear candidate for causing the observed phenotypic variation we PCR-sequenced this gene from all 131 genotypes included in the GWAS. We observed 10 nonsynonymous SNPs that defined three haplotypes (defined as  $\text{Na}^+_{\text{HAP1}}$ ,  $\text{Na}^+_{\text{HAP2}}$ , and  $\text{Na}^+_{\text{HAP3}}$ ) (Fig. 1d–f).  $\text{Na}^+_{\text{HAP1}}$  and  $\text{Na}^+_{\text{HAP2}}$  correspond exactly to the HGB haplotype very recently described<sup>21</sup>. In our GWAS panel there was no association between *HvHKT1;5* haplotype and population structure. We observed a significant difference in mean grain  $\text{Na}^+$  content between genotypes containing the low grain  $\text{Na}^+$  haplotypes,  $\text{Na}^+_{\text{HAP1}}$  ( $M = 28.6$ ,  $SD \pm 72$ ) and  $\text{Na}^+_{\text{HAP2}}$  ( $M = 22.2$ ,  $SD \pm 6.48$ ), and the high grain  $\text{Na}^+_{\text{HAP3}}$  allele ( $M = 51.6$ ,  $SD \pm 15.2$ );  $t(52) = 9.07$ ,  $p = 2.66\text{E-}12$  (Fig. 1g). There was no significant difference between  $\text{Na}^+_{\text{HAP1}}$  and  $\text{Na}^+_{\text{HAP2}}$  ( $M = 28.6$ ,  $SD \pm 7.72$ ,  $M = 22.2$ ,  $SD \pm 6.48$ );  $t(2) = 1.66$ ,  $p = 0.23$ ). Lines containing  $\text{Na}^+_{\text{HAP3}}$  contain an average increase in grain  $\text{Na}^+$  content of 1.8-fold over  $\text{Na}^+_{\text{HAP1}}$  and  $\text{Na}^+_{\text{HAP2}}$ . Of the 10 nonsynonymous SNPs, six in complete linkage disequilibrium (LD) altered amino-acid residues (S56N, Q102E, N130K, L189P, I416V, and N438S) that differentiated the two low  $\text{Na}^+$  haplotypes ( $\text{Na}^+_{\text{HAP1}}$  and  $\text{Na}^+_{\text{HAP2}}$ ) from the high  $\text{Na}^+$  haplotype ( $\text{Na}^+_{\text{HAP3}}$ ). The remaining four (P133Q, I223V, T377A, and H402Q) differentiated low  $\text{Na}^+_{\text{HAP1}}$  and high  $\text{Na}^+_{\text{HAP3}}$  from  $\text{Na}^+_{\text{HAP2}}$  (Fig. 1d, e).

## Transcript abundance varies between *HvHKT1;5* haplotypes.

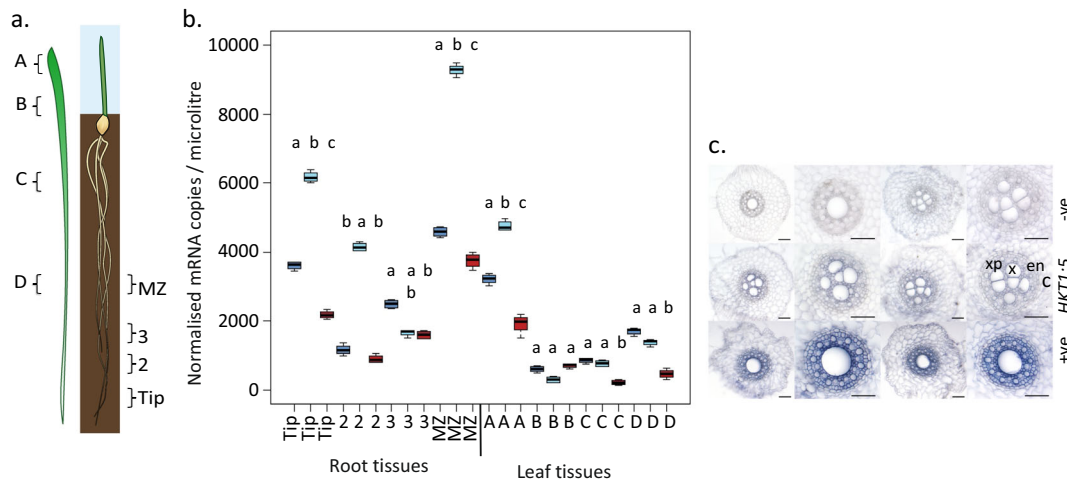
As variation in transcript abundance among *HvHKT1;5* haplotypes has previously been implicated in determining variation in shoot  $\text{Na}^+$  content<sup>21</sup>, we selected genotypes that were representative of  $\text{Na}^+_{\text{HAP1}}$  (cv. Golden Promise),  $\text{Na}^+_{\text{HAP2}}$  (cv. Viivi), and  $\text{Na}^+_{\text{HAP3}}$  (cv. Morex) and quantified *HvHKT1;5* transcript abundance by qRT-PCR in a range of tissues after growth without added  $\text{Na}^+$  (Fig. 2). Golden Promise and Morex were chosen because they had available genome assemblies and publicly available transcript data for a variety of tissues<sup>22,28</sup>. We observed significant differences ( $P < 0.05$ ) in the normalised gene expression in each haplotype across tissues. In general terms, *HvHKT1;5* was more highly expressed in roots compared to shoots, and in low  $\text{Na}^+_{\text{HAP1}}$  and  $\text{Na}^+_{\text{HAP2}}$  compared to high  $\text{Na}^+_{\text{HAP3}}$ . The highest overall expression was observed in the maturation zone of the roots in cv. Viivi ( $\text{Na}^+_{\text{HAP2}}$ ) with in situ hybridisations using sections from this region showing that *HvHKT1;5* was predominantly expressed in the xylem parenchyma and endodermal cells adjacent to the xylem vessels (Fig. 2c)<sup>24,25,29</sup>. These results agree with previous reports that *HvHKT1;5* transcript abundance in roots is inversely correlated with grain  $\text{Na}^+$  content.

## A single substitution in *HvHKT1;5*<sub>HAP3</sub> disrupts $\text{Na}^+$ transport function.

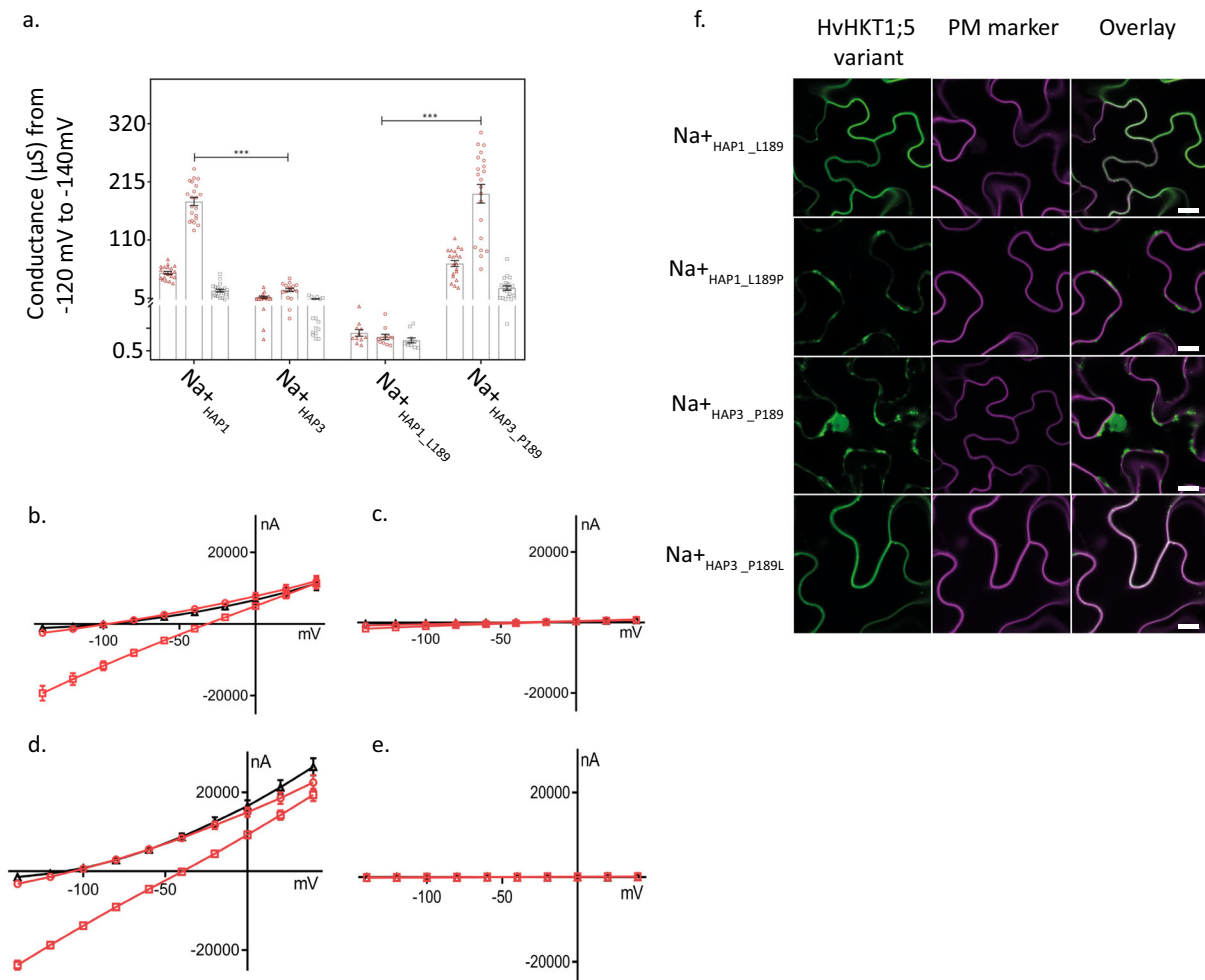
Despite the current lack of evidence for natural functional variation in *HvHKT1;5*<sup>21</sup> we were interested in testing whether the observed haplotypes influenced in vivo  $\text{Na}^+$  transport properties. We assembled and independently tested constructs expressing  $\text{Na}^+_{\text{HAP1}}$ ,  $\text{Na}^+_{\text{HAP2}}$ , or  $\text{Na}^+_{\text{HAP3}}$  in *Xenopus laevis* oocytes using two-electrode voltage-clamp (TEVC) experiments. Oocytes injected with cRNA of  $\text{Na}^+_{\text{HAP1}}$  (or  $\text{Na}^+_{\text{HAP2}}$ , Supplementary Fig. 1) showed significant inward currents in the presence of external  $\text{Na}^+$  but not  $\text{K}^+$  (Fig. 3) consistent with *HvHKT1;5* being a  $\text{Na}^+$ -specific transport protein. With the external  $\text{Na}^+$  concentration increased from 1 mM to



**Fig. 1** *HvHKT1;5* haplotypes influence grain  $\text{Na}^+$  content. **a** Box plots of grain sodium content quantified using ICP-MS,  $n = 5$  biologically independent samples. Blue bars represent lines containing L189 and red bars represent genotypes containing P189. **b** Manhattan plot of GWAS of grain  $\text{Na}^+$  23 content, FDR threshold =  $-\log_{10}(P)=6.02$ . **c** Manhattan plot of GWAS of grain  $\text{K}^+$  content using, FDR threshold =  $-\log_{10}(P)=6.02$ . **d** Gene structure and nonsynonymous polymorphisms coloured according to haplotype shown in **e**. **e** Haplotype summary based on nonsynonymous SNPs. **f** Haplotype analysis of nonsynonymous SNPs. Circles scaled to number of individuals sharing the haplotype, and short lines represent number of SNPs differentiating haplotype groups. **g** Box plots of grain  $\text{Na}^+$  contents in the three haplotypes,  $n = 3, 73,$  and  $41$ , width of plots scaled to number of individuals from each haplotype. Different letters above boxes indicate significant difference in grain sodium content  $p < 0.05$ , same letters indicate no significant difference using this threshold. For both sets of boxplots the horizontal bar of the boxplot shows the median, the box delineates the first and third quartile, and the whiskers show  $\pm 1.5 \times \text{IQR}$ .



**Fig. 2** Spatial expression patterns of *HvHKT1;5* haplotypes. **a** Sections used for quantification of transcript abundance (after Burton et al., 2004<sup>47</sup>). **b** Quantification of *HvHKT1;5* in root (Tip, 2, 3, and MZ) and shoot (A, B, C, and D) tissues from 12-day-old barley plants. Dark blue— $\text{Na}^+_{\text{HAP1}}$  (cv. Golden Promise), pale blue =  $\text{Na}^+_{\text{HAP2}}$  (cv. Vivii), red =  $\text{Na}^+_{\text{HAP3}}$  (cv. Morex) differences from after carrying out an ANOVA followed by Tukey HSD within tissue type, same letters in lower case indicate no significant difference using  $P < 0.05$ . For boxplots the horizontal bar of the boxplot shows the median, the box delineates the first and third quartile, and the whiskers show  $\pm 1.5 \times \text{IQR}$ . **c** In situ localisation of *HvHKT1;5* in 2-week-old barley root tissue from cv. Golden Promise,  $\text{Na}^+_{\text{HAP1}}$  (grown without added NaCl). Top: negative controls with no RT (reverse transcription), Middle *HvHKT1;5* with c cortex, en endodermis, x xylem, xp xylem parenchyma labelled in red. Bottom *Hv18S* rRNA (positive control). Scale bars, 100  $\mu\text{m}$ .



**Fig. 3** Heterologous expression of HvHKT1;5 variants in *Xenopus laevis* oocytes and *Nicotiana benthamiana* epidermal cells. **a** Cation ( $\text{Na}^+$  and  $\text{K}^+$ ) conductance ( $-140$  mV to  $-120$  mV) from HvHKT1;5 haplotypes cRNA-injected oocytes perfused with 1 mM (open red triangles), 30 mM  $\text{Na}^+$  (open red squares) and 30 mM  $\text{K}^+$  (open grey squares). Data are means  $\pm$  SEM of currents,  $n = 11$ –21, each triangle or square represents a single sample, (\*\*\*) $P < 0.001$ , combined from three independent experiments. **b–e** Representative I–V curves of cRNA from HvHKT1;5 haplotypes injected into *Xenopus laevis* oocytes ( $n = 4$ ) clamped at  $-140$  mV to 40 mV in  $\text{Na}^+$  or  $\text{K}^+$  solutions: Red circles (1 mM  $\text{Na}^+$ ), red squares (30 mM  $\text{Na}^+$ ), black triangles (30 mM  $\text{K}^+$ ). **b**  $\text{Na}^+_{\text{HAP1\_L189}}$ ; **c**  $\text{Na}^+_{\text{HAP3\_P189}}$ ; **d**  $\text{Na}^+_{\text{HAP3\_L189}}$ ; **e**  $\text{Na}^+_{\text{HAP3\_P189}}$ . **f** Transient co-expression of GFP-HvHKT1;5 variants with CBL1n-RFP plasma membrane marker in *Nicotiana benthamiana* leaf epidermal cells. GFP signal in the left panel (green), RFP-signal in the middle (magenta), overlay on the right (colocalisation of green and magenta signals appears in white). Scale bars = 10  $\mu\text{m}$ .

30 mM, a two-fold increase in  $\text{Na}^+$  conductance was observed (Fig. 3). For  $\text{Na}^+_{\text{HAP3}}$ ,  $\text{Na}^+$  and  $\text{K}^+$  elicited currents were similar to water-injected controls with the conductance unaltered when external  $\text{Na}^+$  concentration was increased (Fig. 3), indicating that  $\text{Na}^+_{\text{HAP3}}$  was severely compromised in its ability to transport  $\text{Na}^+$  across the plasma membrane.

Publicly available data for HKT1;5 led us to focus on four single amino-acid residue changes as potentially causal for compromised transport (N57S, P189L, V416I, and S438N)<sup>21,29</sup>. We swapped these candidate amino acids residues individually into the compromised high  $\text{Na}^+_{\text{HAP3}}$  and quantified their impact on  $\text{Na}^+$  conductance by TEVC in the oocyte system. In comparison to high  $\text{Na}^+_{\text{HAP3}}$ ,  $\text{Na}^+_{\text{HAP3\_L189}}$  showed a  $\text{Na}^+$ -dependent conductance that was comparable to low  $\text{Na}^+_{\text{HAP1}}$  (Fig. 3). The reciprocal substitution (P189) into low  $\text{Na}^+_{\text{HAP1}}$  showed significantly reduced  $\text{Na}^+$  conductance, comparable to high  $\text{Na}^+_{\text{HAP3}}$  (Fig. 3). No other substitution converted a low  $\text{Na}^+$  haplotype to a high  $\text{Na}^+$  haplotype or vice versa; however, the  $\text{Na}^+_{\text{HAP1\_V416}}$  reduced but did not abolish the  $\text{Na}^+$ -dependent conductance of  $\text{Na}^+_{\text{HAP1}}$  (Supplementary Fig. 1).

These data support the conclusion that the naturally occurring P189 amino-acid residue in  $\text{Na}^+_{\text{HAP3}}$  compromises the function of HvHKT1;5, and that certain variants (e.g.  $\text{Na}^+_{\text{HAP1\_V416}}$ ) can also affect  $\text{Na}^+$  dependent conductance in TEVC experiments.

#### HvHKT1;5<sub>HAP3</sub> does not localise to the plasma membrane.

Consistent with their role in  $\text{Na}^+$  retrieval from the xylem sap, HKT1;5 proteins have been previously shown to localise specifically to the plasma membrane (PM)<sup>30</sup>. We were therefore interested in whether the observed functional variation had consequences for HvHKT1;5 subcellular localisation. We transiently co-expressed N-terminally GFP-tagged HvHKT1;5<sub>HAP3\_L189P</sub> variants with a plasma membrane (PM)-marker in *Nicotiana benthamiana* epidermal cells. Confocal imaging revealed that the low  $\text{Na}^+$  variant HvHKT1;5<sub>HAP3\_L189</sub> was almost exclusively localised at the PM (Fig. 3). However, the high  $\text{Na}^+$  HvHKT1;5<sub>HAP3\_P189</sub> did not colocalise with the PM-marker; the GFP signal was instead localised to internal cell structures (Fig. 3). Introduction of P189 into HvHKT1;5<sub>HAP1</sub> phenocopied the GFP-signal pattern of cells transformed with HvHKT1;5<sub>HAP3</sub> (Fig. 3). This GFP-signal pattern

in cells expressing HvHKT1;5 haplotypes harbouring P189 may suggest protein degradation.

To explore this further we constructed 3D molecular models of HvHKT1;5<sub>HAP3\_L189</sub> and HvHKT1;5<sub>HAP3\_P189</sub> in complex with Na<sup>+</sup> using the *B. subtilis* KtrB K<sup>+</sup> transporter (Protein Data Bank genotype 4J7C, chain I) as a template with K<sup>+</sup> substituted by Na<sup>+</sup><sup>31,32</sup> (Supplementary Fig. 2). In the structural models, detailed analysis of the micro-environments around  $\alpha$ -helices 4 and 5 revealed that the  $\alpha$ -helix 4 of low Na<sup>+</sup> allele HvHKT1;5<sub>HAP3\_L189</sub> established a network of four polar contacts at separations between 2.7 Å to 3.1 Å with A185, V186, Y192, and S193 neighbouring residues. However, these were not formed in high Na<sup>+</sup> HvHKT1;5<sub>HAP3\_P189</sub>, which only established two polar contacts at separations between 2.5 and 2.7 Å with S193. We observed a positive correlation between the structural characteristics of  $\alpha$ -helices 4 and 5 (trends in angles based on  $\alpha$ -helical planes), differences in Gibbs free energies of forward (P189L) and reverse (L189P) mutations, and the ability to produce Na<sup>+</sup> fluxes across oocyte membranes. Combined with our previous observations we hypothesise that P189 in HvHKT1;5 does affect protein structure, potentially triggering protein degradation prior to insertion into the plasma membrane thereby reducing Na<sup>+</sup> retrieval from the xylem with bulk flow ultimately elevating Na<sup>+</sup> in the grain.

**The impact of HvHKT1;5 haplotypes under salt stress.** While our original data were collected from plants grown under optimal conditions, most recent reports in the literature focus on the impact of variation at HKT1;5 on natural tolerance to growth in saline environments<sup>33–35</sup>. We therefore explored the impact of HvHKT1;5 variants on a range of phenotypic traits after growth in 0, 150, and 250 mM added NaCl (Fig. 4a). While we observed confounding between allele, haplotype, and line (see full analysis given in Supplementary dataset 1), we can nevertheless conclude that grain Na<sup>+</sup> content is influenced significantly by both allele (L189P) and treatment (NaCl); lines containing the functionally compromised P189 allele accumulate higher concentrations of Na<sup>+</sup> than lines containing L189 and show a larger difference between control and salt treatments. All lines with L189 accumulated less Na<sup>+</sup> in the grain than those with P189, with one genotype, Maris Mink, having especially high grain Na<sup>+</sup>. This line had the second highest concentration of grain Na<sup>+</sup> when grown as part of the GWAS panel (Supplementary Data 1). Despite the higher Na<sup>+</sup> content, no clear detrimental effect of P189 was observed on vegetative biomass (Fig. 4b). NaCl at 250 mM had a strong and consistent negative influence on total biomass across all lines regardless of HvHKT1;5 haplotype. There was no influence of L189P on grain K<sup>+</sup> content ( $P > 0.05$ ), although there was an effect of treatment on this trait ( $P = 0.02$ ) (Supplementary Fig. 3). As previous studies of HKT1;5 s generally focus on shoot Na<sup>+</sup>, we also examined shoots from the same plants. Again, we observed that an interaction between allele and treatment significantly influenced leaf Na<sup>+</sup> content ( $P < 0.001$ ), with lines containing L189 accumulating less Na<sup>+</sup> in leaf tissue than those with P189 (Supplementary Fig. 4, Supplementary Data 5), mirroring our observations in grain.

### HvHKT1;5<sub>HAP3</sub> frequency increases significantly in elite barley.

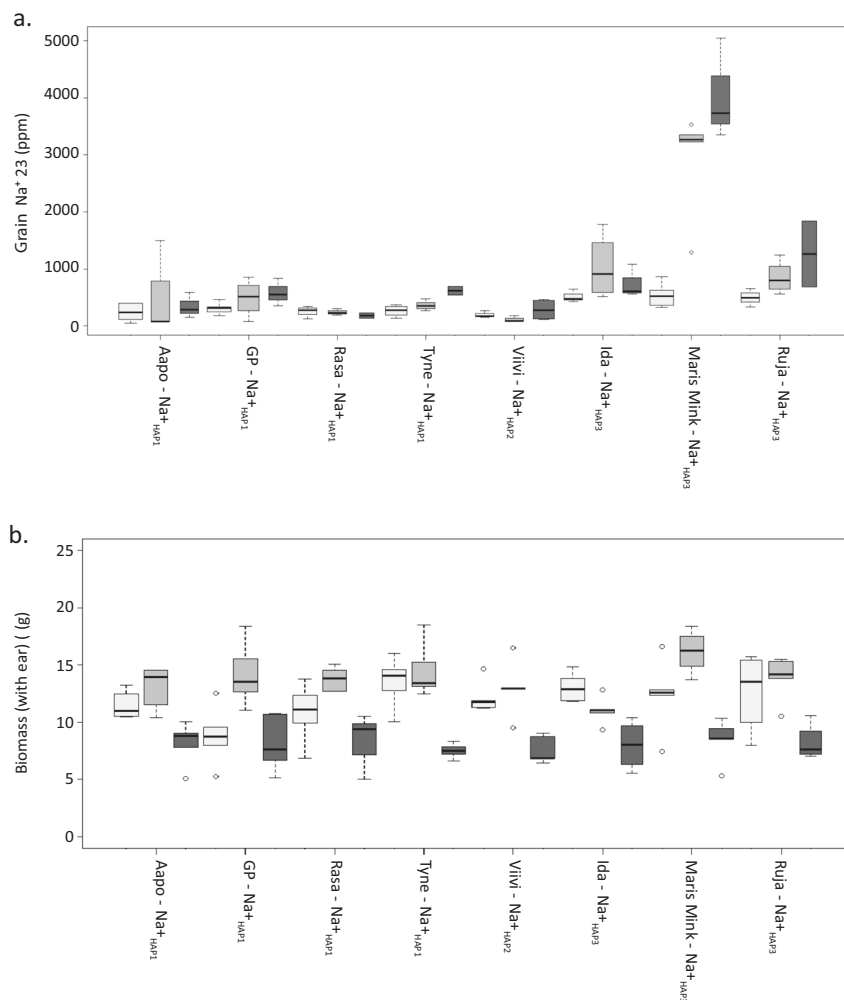
Previously observed associations between *AtHKT1* alleles and the environment<sup>36</sup> prompted us to explore whether barley HvHKT1;5 haplotypes had any obvious evolutionary or ecological significance. We identified and downloaded orthologs of *HKT1;5* and aligned the retrieved protein sequences with MUSCLE<sup>37</sup>. Based on the available sequences only barley genotypes contained the L189P substitution in HKT1;5 despite comparing amino-acid

sequences of 11 different species (Supplementary Fig. 5). To explore the origin and distribution of the P189 variant, we then PCR-sequenced HvHKT1;5 from a collection of 73 georeferenced wild barley (*H. spontaneum*) genotypes from the fertile crescent<sup>38</sup>. This revealed 19 additional nonsynonymous SNPs defining 27 haplotypes (Supplementary Data 3) that are distinct from those observed in the GWAS panel, with the exception of Na<sup>+</sup><sub>HAP3</sub> that was found in one genotype, FT064, originating from southern Israel (latitude = 31.35, longitude = 35.12) (Supplementary Figs. 6, 7). A maximum likelihood tree based on nonsynonymous SNPs in these *H. spontaneum* genotypes plus those in the elite cultivated barley genotypes revealed three discrete clades each containing one elite line haplotype (Supplementary Figure 7). We then genotyped the L189P polymorphism in 184 georeferenced landraces and found that 7 genotypes (<4%), mostly located in western Europe, contained the P189 substitution (Supplementary Figs. 6, 7). Strikingly, this frequency increased to 35% in the 131 elite genotypes used for GWAS (Supplementary Fig. 7). This increase, which occurs across all branches of the cultivated gene pool, is symptomatic of what would be expected for a locus that is currently in the process of undergoing directional selection.

### Discussion

Sodium accumulation is a complex trait that can have serious implications for plant performance and survival. By combining high density SNP array and ionic data collected from the grain of plants grown under nonsaline conditions, we identified haplotypes of HvHKT1;5 as the major genetic factor determining Na<sup>+</sup> content in contemporary two-row spring barley. *HKT1;5* has previously been implicated in conferring a degree of salinity tolerance in wheat and barley through a Na<sup>+</sup> exclusion mechanism<sup>23,24</sup>. In high Na<sup>+</sup> accumulating genotypes we found that a single SNP generating an L189P amino-acid residue substitution led to severely compromised HvHKT1;5 function. We propose this is likely due to a combination of protein structural changes leading to misfolding, aberrant subcellular localisation, and subsequent degradation, and is compounded by higher transcript abundance of the functional alleles<sup>21</sup>. Intriguingly, a recent study<sup>39</sup> in bread wheat identified the exact same nucleotide substitution found in the high Na<sup>+</sup> accumulating HvHKT1;5<sub>HAP3</sub> in the landrace Mocho de Espiga Branca, TaHKT1;5 (L190P). Like HvHKT1;5 when this allele is expressed in oocytes, elicited currents were similar to water-injected controls with the conductance unaltered when external Na<sup>+</sup> concentration was increased. Similarly, this TaHKT1;5 (L190P) was also localised to internal cell structures. In bread wheat the TaHKT1;5 (L190P) was associated with lower Na<sup>+</sup> retrieval from the xylem and high shoot Na<sup>+</sup> accumulation.

After growing representatives of each haplotype under saline conditions, our observations align with previous findings in barley<sup>19,20</sup>, and in rice<sup>25</sup> in response to short-term NaCl stress. However, in the latter, longer term stress (21 days of 40 mM or 80 mM NaCl) led to a 72% decrease in biomass in an *OsHKT1;5* expression mutant compared to wild-type. Consistent with these findings, when Munns et al.<sup>33</sup> backcrossed a functional *Nax2* locus (*TmHKT1;5 A*) from *T. monococcum* into commercial Durum wheat they observed a grain yield increase of 25% compared to the control when grown on saline soils. Both studies parallel the relationship observed between Arabidopsis *AtHKT1* allele and seed number from plants grown under saline conditions; wild-type Arabidopsis Col-0 produced seeds when exposed to moderate salt stress while an *athkt1* knockout mutant was virtually sterile<sup>34</sup>. Together they suggest that HKT1's are critically important for maintaining fitness under saline conditions. Our



**Fig. 4** Influence of L189P polymorphism in *HvHKT1;5* on grain  $\text{Na}^+$  accumulation and other morphological traits. **a** Mature grain  $\text{Na}^+$  content of barley genotypes exposed to 0, 150, and 250 mM NaCl at the fourth leaf stage of development. **b** Above ground biomass (including the ear of grain) after barley genotypes were exposed to different concentrations of NaCl at the fourth leaf stage of development. For both **a** and **b** *HvHKT1;5* haplotype is shown alongside name of genotype, low grain  $\text{Na}^+$ <sub>HAP1</sub>,  $\text{Na}^+$ <sub>HAP2</sub>, and high grain  $\text{Na}^+$ <sub>HAP3</sub>. White bars indicate 0 mM of NaCl, light grey bars indicate 150 mM NaCl, and dark grey indicates 250 mM NaCl added to plants. For the boxplots the horizontal bar of the boxplot shows the median, the box delineates the first and third quartile, and the whiskers show  $\pm 1.5 \times \text{IQR}$ ,  $n = 5$  biologically independent samples.

data, supported by the recent evidence that shoot  $\text{Na}^+$  accumulation in certain bread wheat genotypes is not negatively associated with plant salinity tolerance<sup>39,40</sup>, question this conclusion for barley. Here we show that biomass yield is maintained in lines such as Maris Mink (*HvHKT1;5*<sub>HAP3</sub>) that have levels of grain (and leaf)  $\text{Na}^+$  content that would be expected to have a significant negative impact. We conclude that alternative or additional mechanisms must be involved in  $\text{Na}^+$  tolerance in barley and that  $\text{Na}^+$  exclusion by *HvHKT1;5* per se may be a relatively minor player. This discrepancy could potentially arise because—unlike barley—*Arabidopsis*, rice, and Durum wheat are all particularly sensitive to saline conditions, pointing to fundamental differences in the roles of *HKT1;5* between salt-tolerant and salt-sensitive plants. Recently, *HvHKT1;5* RNAi knockdown lines generated in cv.GP ( $\text{Na}^+$ <sub>HAP1</sub><sub>L189</sub>), were shown to exhibit an increase in shoot biomass compared to wild-type at increasing NaCl concentrations<sup>41</sup>. Somewhat controversially, the authors hypothesise that *HvHKT1;5* translocates  $\text{Na}^+$  from the root to the shoot, and the increase in biomass in the RNAi lines is due to a reduction in translocation of  $\text{Na}^+$  due to decreased expression of *HvHKT1;5*. However, the data presented here, in another recent study of *HvHKT1;5*<sup>21</sup>, and in reports from several different crop

species<sup>10–12</sup>, indicate that the presence of functionally compromised alleles of *HvHKT1;5* leads to an increase in shoot and grain sodium content due to a reduced ability to exclude  $\text{Na}^+$  from the plant.

The observed increase in the frequency of high  $\text{Na}^+$  *HvHKT1;5*<sub>HAP3</sub> in breeding germplasm from NW Europe returns us to the possible role of  $\text{Na}^+$  as a functional micronutrient in agriculture and its beneficial effects on plant growth and development, particularly in low  $\text{K}^+$  environments. While the NW European growing environment is largely devoid of saline soils,  $\text{K}^+$  deficiency is widespread both in soils with a low clay content and where the annual removal of  $\text{K}^+$  by high yielding varieties is greater than the amount of  $\text{K}^+$  applied. In such situations the higher  $\text{Na}^+$  content of *HvHKT1;5*<sub>HAP3</sub> could provide a physiological advantage, for example through use of  $\text{Na}^+$  as a substitute for  $\text{K}^+$  in a range of metabolic functions or as a free and abundant osmolyte to reduce leaf water potential, increase (or maintain) transpiration and photosynthesis, and ultimately impact yield. Assuming the latter hypotheses are correct, we cautiously speculate that the increase in frequency of *HvHKT1;5*<sub>HAP3</sub> may reflect ongoing positive selection during breeding due to it providing a selective advantage. However, given it is not yet near

fixation in elite genotypes, distinguishing this hypothesis from alternatives such as indirect selection due to linkage drag with another positive trait (i.e. hitchhiking) clearly remains to be tested.

Overall, we conclude that natural allelic variation at *HvHKT1;5* has a strong influence over grain and shoot  $\text{Na}^+$  homeostasis in barley in both nonsaline and saline environments. A single SNP causing an L189P amino-acid residue substitution likely results in a change in protein structure, that leads to aberrant subcellular localisation, loss of capacity to transport  $\text{Na}^+$  and consequent reduction in capacity to remove  $\text{Na}^+$  from the transpiration stream leading to elevated concentrations in the shoots and grain. When grown under salt stress conditions we observed no negative consequences of the exceptionally high concentrations of  $\text{Na}^+$  found in genotypes containing *HvHKT1;5*<sub>HAP3</sub> on a range of life history traits, most notably biomass. This questions the widely accepted role of *HvHKT1;5* driven  $\text{Na}^+$  exclusion in the root as the dominant player in salinity tolerance in barley. Given the identical amino-acid residue substitution (L190P) in *TaHKT1;5D* was recently identified in the bread wheat genotype, Mocho de Espiga Branca, that similarly exhibits atypically high shoot  $\text{Na}^+$  content and significantly reduced  $\text{Na}^+$  conductance in TEVC experiments<sup>39</sup>, we provide a remarkable example of parallel evolution in two economically important species with potential value in plant breeding.

## Methods

**Phenotypic characterisation of grain sodium for GWAS.** A collection of 131 contemporary European two-rowed spring barley genotypes, for the purposes of this study our GWAS panel, were grown in a polytunnel in Dundee, Scotland, using standard barley soil and growth conditions. These are as follows: peat = 1.2 m<sup>3</sup>, sand = 100 l, osmocote exact start = 1.5 kg, osmocote Exact Mini/3–4 month/10–12 month = 3.5 kg, 2.5 kg each of Lime, Ca, and Mg, celcote = 0.5 kg, perlite = 100 l, and intercept/exemptor = 280 g/390 g.

We screened the grain sodium concentration of the GWAS panel using Inductively Coupled Plasma Mass Spectrometry (ICP-MS). Barley grains were transferred into Pyrex test tubes (single grain per tube) and weighted. Samples were predigested overnight at room temperature with 1 mL trace metal grade nitric acid Primar Plus (Fisher Chemicals) spiked with indium internal standard followed by digestion in dry block heaters (DigiPREP MS, SCP Science; QMX Laboratories, Essex, UK) at 115 °C for 4 h. Then, 1 mL of hydrogen peroxide (Primar, for trace metal analysis, Fisher Chemicals) was added and samples were digested in dry block heater at 115 °C for 2 h. After cooling down, the digests were diluted to 10 mL with 18.2 MΩcm Milli-Q Direct water (Merck Millipore) and elemental analysis was performed using PerkinElmer NexION 2000 ICP-MS equipped with Elemental Scientific Inc. autosampler, in the collision mode (He). Twenty-one elements (Li, B, Na, Mg, P, S, K, Ca, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Rb, Sr, Mo, Cd, and Pb) were monitored. The isotopes 23 and 39 were measured for Na and K, respectively. Liquid reference material composed of the pooled digested samples was prepared before the beginning of the sample run and was used throughout the whole samples run. It was run after every ninth sample in all ICP-MS sample sets to correct for variation between and within ICP-MS analysis runs. The calibration standards (with indium internal standard and blanks) were prepared from single element standards solutions (Inorganic Ventures; Essex Scientific Laboratory Supplies Ltd, Essex, UK). Sample concentrations were calculated using external calibration methods within the instrument software. Further data processing was performed in Microsoft Excel. For each genotype Na was measured in five single grains per sample digested, equating to five biological replicates. From these data BLUPs were predicted using GenStat (15th edition).

**DNA extraction, 50k iSelect genotyping, and GWAS.** DNA from 7-day-old leaves was extracted for all genotypes using the QIAamp kit (Qiagen) on the QIAcube HT (Qiagen) using default settings. All samples were genotyped using the 50k iSelect SNP array as described in<sup>18</sup>. GWAS was carried out on adjusted variety means using the EMMA algorithm, a kinship matrix derived using Van raden, and setting the PCA to true in GAPIT<sup>42</sup> with R version 3.5.2<sup>43</sup>. False discovery rates as calculated in GAPIT using Benjamini-Hochberg (1995) FDR-controlling procedure were used to account for the large number of SNPs used in this analysis. We anchored regions of the genome, which were significantly associated with  $\text{Na}^+$  content to the physical map of the barley sequence to provide annotations for genes within these regions<sup>22,28</sup>. Linkage disequilibrium (LD) was calculated for regions of the genome containing significant associations between pairs of markers using a sliding window of 500 markers and a threshold of  $R^2 < 0.2$  using Tassel v5<sup>44</sup> to allow us to identify local blocks of LD, facilitating a more precise delimitation of

QTL regions. We anchored regions of the genome containing markers that passed the FDR to the physical map and then expanded this region using local LD derived from genome-wide LD analysis as described above.

**Resequencing *HvHKT1;5* and sequence alignment.** For the 131 genotypes of the GWAS panel we PCR amplified and Sanger sequenced the coding sequence of *HvHKT1;5*. We used the primers listed in Supplementary Data 4 to resequence this gene. DNA was amplified and cleaned up prior to Sanger sequencing on an ABI3100 capillary sequencer using reaction mixes and conditions described<sup>45</sup>. Sequences were aligned in Geneious version 9.0.2 (Biomatters Ltd). Haplotype networks were produced using PopART version 1.7<sup>46</sup>. Orthologs of *HvHKT1;5* were identified using the blastx function at NCBI, and the sequences retrieved aligned in Geneious version 9.0.2 (Biomatters Ltd) using MUSCLE with default settings.

**RNA extraction and cDNA synthesis.** Materials detailed in Fig. 2a were sampled, snap frozen and stored at –80 °C for RNA extraction. The root tissue was ground to fine powder on 2010 Geno/Grinder® (SPEX SamplePrep) at 1200 RPM for 30 seconds, and RNA was extracted from the tissue powder by using Direct-Zol RNA MiniPrep (Zymo Research) according to the manufacturer's protocol. Final elution was performed with 40 µL DNA/RNAase-Free water supplied with the kit and the eluted RNA was subsequently quantified using ND-1000 Spectrophotometer (NanoDrop Technologies). cDNA synthesis was then performed on 500 ng RNA by using High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific) according to the manufacturer's instruction in a 20 µL reaction and stored at –20 °C until use.

**RNA extraction and qPCR of *HvHKT1;5*.** Root and shoot tissue from 12-day-old roots and shoots were collected in sections as described<sup>47</sup> from genotypes grown in the same polytunnel as described above, in the same conditions for RNA extraction. Each of the four Biological reps consisted of tissue collected from 15 individual plants. cDNA was synthesised using RNA to cDNA EcoDry™ Premix (Double Primed) (Takara) using standard conditions and used for qPCR. qPCR and the analysis of the subsequent data was carried out as described<sup>47</sup> using 3 housekeeping genes,  $\alpha$ -tubulin, GAPDH, and HSP70. Primer sequences and annealing temperatures are provided in Supplementary Data 4.

**Characterisation of diversity of *HKT1;5*.** Species orthologs of *HKT1;5* were identified using the blastx function at NCBI, and the protein sequences retrieved were aligned in Geneious version 9.0.2 (Biomatters Ltd) using MUSCLE with default settings. For the 73 *H. spontaneum* genotypes, DNA was extracted, *HvHKT1;5* amplified and Sanger sequenced as described above. For the landraces, the L189P SNP was genotyped in 184 georeferenced genotypes<sup>38</sup> using primer pair 3 in Supplementary Data 4 and PCR-sequencing conditions described above. The geolocation data for these genotypes is available<sup>37</sup>.

**In situ PCR.** Barley roots in situ PCR was followed by Athman et al.<sup>48</sup> with the following modifications. Root cross sections (from maturation zone) were 60 µm obtained using Vibrating Microtome 7000 Model 7000smz-2 (Campden Instruments Ltd.). Thermocycling conditions for the PCR were: initial denaturation at 98 °C for 30 seconds, 35 cycles of 98 °C for 10 seconds, 59 °C (for *HvHKT1;5*), or 57 °C (for *Hv18S*) for 30 seconds, 72 °C for 10 seconds, and a final extension at 72 °C for 10 min. Gene specific primers for *HvHKT1;5* and *Hv18S* (positive control) are shown in Supplementary Data 4.

**Characterisation of *HvHKT1;5* in oocytes.** Methods for functional characterisation of *HvHKT1;5* variants in *Xenopus laevis* oocytes were as described previously<sup>30,32</sup>. Haplotype and engineered variants of *HvHKT1;5* were synthesised by GenScript (Piscataway, NJ, USA) and fragments were inserted into a gateway enabled pGEMHE vector. Nucleotides encoding *HvHKT1;5* N57S, P189L, V416I and S438N were modified by site-directed mutagenesis PCR using Phusion® High-Fidelity DNA Polymerase (New England Biolabs, Massachusetts, USA). pGEMHE constructs were linearized using *sbfl* (New England Biolabs, Massachusetts, USA) followed by ethanol precipitation. Complimentary RNA (cRNA) was transcribed using the Ambion mMESSAGE mMACHINE kit (Life Technologies, Carlsbad, CA, USA), 23 ng of cRNA (in 46 nL) or equal volumes of RNA-free water were injected into oocytes, followed by an incubation in ND96 for 24–48 h before recording. Membrane currents were recorded in the HMg solution (6 mM MgCl<sub>2</sub>, 1.8 mM CaCl<sub>2</sub>, 10 mM MES and pH 6.5 adjusted with a TRIS base) ±Na<sup>+</sup> glutamate and/or K<sup>+</sup> glutamate as indicated. All solution osmolarities were adjusted using mannitol at 220–240 mOsmol kg<sup>-1</sup><sup>131,49</sup>

**Transient expression of *HvHKT1;5* in *Nicotiana benthamiana*.** Transient expression of fluorescent fusion proteins was performed as described in detail<sup>26</sup>. In brief, *HKT1;5* coding sequences were recombined into *pMDC43* to generate N-terminally GFP-tagged proteins. For colocalisation studies, nCBL1-RFP was used as a PM-marker<sup>49</sup>. All constructs were transformed into *Agrobacterium tumefaciens* strain Agl-1. Agroinfiltration was performed on fully expanded leaves of 4- to

6-week-old *Nicotiana benthamiana* plants. After 2 days, leaf sections were imaged using a Nikon A1R Confocal Laser-Scanning Microscope equipped with a 633-water objective lens and NIS-Elements C software (Nikon Corporation). Excitation/emission conditions were GFP (488 nm/500–550 nm) and RFP (561 nm/570–620 nm).

### Construction of 3D molecular models of HvHKT1;5<sub>HAP3\_L189</sub> and HvHKT1;5<sub>HAP3\_P189</sub> in complex with Na<sup>+</sup>.

The most suitable template for cereal HKT1;5 transporter proteins was the *B. subtilis* KtrB K<sup>+</sup> transporter (Protein Data Bank genotype 4J7C, chain I) as previously identified. In KtrB, K<sup>+</sup> was substituted by Na<sup>+</sup> during modelling of all HKT1;5 proteins. 3D models of HvHKT1;5<sub>HAP3\_L189</sub> and HvHKT1;5<sub>HAP3\_P189</sub> in complex with Na<sup>+</sup> were generated in Modeller 9v19<sup>50</sup> as described previously<sup>51</sup> incorporating Na<sup>+</sup> ionic radii<sup>52</sup> taken from the CHARMM force field<sup>53</sup>, on the Linux station running the Ubuntu 12.04 operating system. Best scoring models (from an ensemble of 50) were selected based on the combination of Modeller Objective Function<sup>54</sup> and Discrete Optimised Protein Energy term<sup>55</sup> PROCHECK<sup>56</sup>, ProSa 2003<sup>57</sup> and FoldX<sup>58</sup>. Structural images were generated in the PyMOL Molecular Graphics System V1.8.2.0 (Schrödinger LLC, Portland, OR, USA). Calculations of angles between selected  $\alpha$ -helices in HvHKT1;5 models were executed in Chimera<sup>59</sup> and evaluations of differences ( $\Delta\Delta G = \Delta G_{mut} - \Delta G_{wt}$ ) in Gibbs free energies was performed with FoldX<sup>58</sup>. Sequence conservation patterns were analysed with ConSurf<sup>60,61</sup> based on 3D models of HvHKT1;5 transporters.

Evaluations of stereo-chemical parameters indicated that the template and HvHKT1;5 models had satisfactory parameters as indicated by Ramachandran plots with two residues positioned in disallowed regions, corresponding to 0.5% of all residues, except of G and P. Average G-factors (measures of correctness of dihedral angles and main-chain covalent bonds) of the template, and HvHKT1;5<sub>HAP3\_L189</sub> and HvHKT1;5<sub>HAP3\_P189</sub> models, calculated by PROCHECK (0.06, -0.07, and -0.07, respectively), and ProSa 2003 z-scores (measures of C <sup>$\beta$</sup> -C <sup>$\beta$</sup>  pair interactions of -9.0, -5.6, and -6, respectively), indicated that template and modelled structures had favourable conformational energies.

**Plant material and growth conditions.** Eight barley genotypes (*Hordeum vulgare*) with variation in the key functional SNP (P189L) in HvHKT1;5 were grown for screening leaf and grain Na<sup>+</sup> accumulation under salt stress conditions. This set of genotypes consist of at least three representative barley genotypes of each allele of P189L in HvHKT1;5 characterised in the elite germ in plasm within this study, cv. Golden Promise (GP) (Na<sup>+</sup><sub>HAP1\_L189</sub>), Aapo (Na<sup>+</sup><sub>HAP1\_L189</sub>), Rasa (Na<sup>+</sup><sub>HAP1\_L189</sub>), and Tyne (Na<sup>+</sup><sub>HAP1\_L189</sub>), Viivi (Na<sup>+</sup><sub>HAP2\_L189</sub>), Ida (Na<sup>+</sup><sub>HAP3\_P189</sub>), Maris Mink (Na<sup>+</sup><sub>HAP3\_P189</sub>), Rujā (Na<sup>+</sup><sub>HAP3\_P189</sub>). Three germinated seeds from each genotype were sown in a 10 × 10 cm pot filled with a standard cereal compost mix as described above. Eight replicates were sown per genotype in a randomized design. Every 24 pots were randomized in a plastic gravel tray (56 × 40 × 4 cm), and in total 12 trays were placed in the glasshouse under long-day conditions (light:dark, 16 h:8 h, 18 °C:14 °C). At sowing, the soil moisture and weight of all pots ranged between 31–35% (w/w) and between ~380–385 g. The three seedlings in each pot were thinned to one at the emergence of 2nd leaf.

Before applying the salt treatment soil moisture content in all pots was controlled to around 25% (w/w) to provide larger NaCl uptake capacity. At the emergence of the 4th leaf, a single salt treatment (150 or 250 mM NaCl) was applied directly into each tray in a 2 L volume. The same volume of water was added to the control trays. The fully expanded 5th leaf was harvested for Na<sup>+</sup> content analysis using ICP-MS as described above to evaluate the strength of salt treatment (Supplementary Data 5). Plants were harvested at maturity and the Na<sup>+</sup> contents of grains and 5th leaves of five of the eight replicates quantified by ICP-MS. Detailed description of the analysis of the resulting data is included in Supplementary Data 1.

**Phylogenetic analysis of HvHKT1;5.** Phylogenetic analysis was conducted using MEGA 7.0.14<sup>62</sup>. For all four datasets we first did a Model selection using maximum likelihood. Sites containing gaps in more than 5% of the data were removed from the analysis. For the elite cultivated barley and the landraces, the GTR + G + I model was selected (General Time Reversible model, Gamma distributed with Invariant Sites). For the *H. spontaneum* genotypes a HKY + G model was selected (Hasegawa–Kishino–Yano model, Gamma distributed). For the *H. spontaneum* tree with only nonsynonymous SNPs (Figure 7A) the Jukes–Cantor model was selected. The phylogenetic tree was done using Maximum Likelihood method with 100 bootstrap replications and the above described models accordingly. Bootstrap values above 60% (represented as 0.6) are shown on the tree.

**Statistics and reproducibility.** Analysis of the phenotypic data collected after plants were exposed to 0, 150, or 250 mM was carried out using ASReml-R and Genstat. Significance was tested at the 5% level, applying either Tukeys or Bonferroni least significant difference to test for differences between treatment levels. A detailed description of the analysis of this dataset is provided in the Supplementary dataset section of the Supplementary Information.

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

### Data availability

NCBI accession numbers for Sanger sequence data corresponding to HvHKT1;5 which have no restrictions are available in Supplementary Data 1 and 3. Genotyping data used for GWAS and all other relevant data not provided in this manuscript already are available from the corresponding authors upon request.

Received: 15 January 2020; Accepted: 28 April 2020;

Published online: 22 May 2020

### References

- Subbarao, G. V., Ito, O., Berry, W. L. & Wheeler R. M. Sodium—a functional plant nutrient. *Crit. Rev. Plant Sci.* **22**(5), 391–416. (2003).
- Flowers, T. J. & Colmer, T. D. Plant salt tolerance: adaptations in halophytes. *Ann. Bot.* **115**(3), 327–331 (2015).
- Lehr, J. J. The importance of sodium for plant nutrition I. *Soil Sci.* **52**(3), 237–244 (1941).
- Lehr, J. J. Sodium as a plant nutrient. *J. Sci. Food Agr.* **4**, 460–471 (1953).
- Mullison, W. R. & Mullison, E. Growth responses of barley seedlings in relation to potassium and sodium nutrition. *Plant Physiol.* **17**(4), 632–644 (1942).
- Harmer, P. M. & Benne, E. J. Sodium as a crop nutrient. *Soil Sci.* **60**(2), 137–148 (1945).
- Cope, J. T., Bradford R. & Peech, M. Effect of sodium fertilization on yield cation content of some field crops. *Soil Sci.* **76**(1), 65–74 (1953).
- Lehr, J. J. & Wybenga, J. M. Exploratory pot experiments on sensitiveness of different crops to sodium: C. Flax. *Plant Soil.* **6**(3), 251–261 (1955).
- Truog, E., Berger, K. C. & Attoe, O. J. Response of nine economic plants to fertilization with sodium. *Soil Sci.* **76**(1), 41–50 (1953).
- Flowers T. J. & Läuchli A. *Inorganic Plant Nutrition*. (Springer Verlag, Berlin, 1983).
- Kronzucker, H. J., Coskun, D., Schulze, L. M., Wong, J. R. & Britto, D. Sodium as nutrient and toxicant. *Plant Soil* **369**(1), 1–23 (2013).
- Tester, M. & Davenport, R. Na<sup>+</sup> tolerance and Na<sup>+</sup> transport in higher plants. *Ann. Bot.* **91**(5), 503–527 (2003).
- Munns, R. & Tester, M. Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.* **59**, 651–681 (2008).
- Roy, S. J., Negrão, S. & Tester, M. Salt resistant crop plants. *Curr. Opin. Biotech.* **26**, 115–124 (2014).
- Munns, R., Passioura, J. B., Colmer, T. D. & Byrt C. S. Osmotic adjustment and energy limitations to plant growth in saline soil. *N. Phytol.* **225**, 1091–1096 (2019).
- Darrier, B. et al. A comparison of mainstream genotyping platforms for the evaluation and use of barley genetic resources. *Front. Plant Sci.* **10**, 544 (2019).
- 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**, 1388–1392 (2012).
- Bayer, M. et al. Development and evaluation of a barley 50k iSelect SNP array. *Front. Plant Sci.* **8**, 1792 (2017).
- Long, N. V. et al. Association mapping of salt tolerance in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **126**(9), 2335–2351 (2013).
- Nguyen, V. L. et al. Identification of quantitative trait loci for ion homeostasis and salt tolerance in barley (*Hordeum vulgare* L.). *Mol. Breeding* **31**, 137–152 (2013).
- van Bezouw, R. H. M. et al. Shoot sodium exclusion in salt stressed barley (*Hordeum vulgare* L.) is determined by allele specific increased expression of HKT1;5. *J. Plant Physiol.* **241**, 153029 (2019).
- Mascher, M. et al. A chromosome conformation capture ordered sequence of the barley genome. *Nature* **544**, 427 (2017).
- Hazzouri, K. M. et al. Mapping of HKT1;5 gene in barley using GWAS approach and its implication in salt tolerance mechanism. *Front. Plant Sci.* **9**, 156 (2018).
- Byrt, C. S. et al. HKT1;5-like cation transporters linked to Na<sup>+</sup> exclusion loci in wheat, Nax2 and Kna1. *Plant Physiol.* **143**, 1918–1928 (2007).
- Kobayashi, N. I. et al. OsHKT1;5 mediates Na<sup>+</sup> exclusion in the vasculature to protect leaf blades and reproductive tissues from salt toxicity in rice. *Plant J.* **91**, 657–670 (2017).
- Henderson, S. W. et al. Functional differences in transport properties of natural HKT1;1 variants influence shoot Na<sup>+</sup> exclusion in grapevine rootstocks. *N. Phytol.* **217**, 1113–1127 (2018).
- Zhang, M. et al. A retrotransposon in an HKT1 family sodium transporter causes variation of leaf Na<sup>+</sup> exclusion and salt tolerance in maize. *N. Phytol.* **217**, 1161–1176 (2018).



28. Schreiber, M. et al. A genome assembly of the barley ‘transformation reference’ cultivar Golden Promise. *bioRxiv* G3: Genes[Genomes]Genetics <https://doi.org/10.1534/g3.119.401010> (2020).
29. Amarasinghe, S. L. et al. Transcriptional variation is associated with differences in shoot sodium accumulation in distinct barley varieties. *Environ. Exp. Bot.* **166**, 103812 (2019).
30. Byrt, C. S. et al. The Na<sup>+</sup> transporter, TaHKT1;5-D, limits shoot Na<sup>+</sup> accumulation in bread wheat. *Plant J.* **80**, 516–526 (2014).
31. Vieira-Pires, R. S., Szollosi, A. & Morais-Cabral, J. H. The structure of the KtrAB potassium transporter. *Nature* **496**, 323–328 (2013).
32. Xu, B., Hrmova, M. & Gilliland, M. High affinity Na<sup>+</sup> transport by wheat HKT1;5 is blocked by K<sup>+</sup>. *bioRxiv* <https://doi.org/10.1101/28045> (2018).
33. Munns, R. et al. Wheat grain yield on saline soils is improved by an ancestral Na<sup>+</sup> transporter gene. *Nat. Biotechnol.* **30**, 360 (2012).
34. An, D. et al. AtHKT1 drives adaptation of *Arabidopsis thaliana* to salinity by reducing floral sodium content. *PLoS Genet.* **13**, e1007086 (2017).
35. Busoms, S. et al. Fluctuating selection on migrant adaptive sodium transporter alleles in coastal *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* **115**, E12443–E12452 (2018).
36. Baxter, I. et al. A coastal cline in sodium accumulation in *Arabidopsis thaliana* is driven by natural variation of the sodium transporter AtHKT1;1. *PLoS Genet.* **6**(11), e1001193 (2010).
37. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
38. Russell, J. et al. Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nat. Genet.* **48**, 1024–1030 (2016).
39. Borjigin C. et al., A single nucleotide substitution in TaHKT1;5-D controls shoot Na<sup>+</sup> accumulation in bread wheat. *bioRxiv* <https://doi.org/10.1101/2020.01.21.909887> (2020).
40. Genc, Y. et al. Bread wheat with high salinity and sodicity tolerance. *Front. Plant Sci.* **10**, 1280 (2019).
41. Huang, L. et al. The HKT transporter HvHKT1;5 negatively regulates salt tolerance. *Plant Physiol.* **182**(1), 84–96 (2020).
42. Kang, H. M. et al. Efficient control of population structure in model organism association mapping. *Genetics* **178**, 1709–1723 (2008).
43. R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing* Vienna, Austria. <https://www.R-project.org/> (2018).
44. Bradbury P. J. et al. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**(19), 2633–2635 (2007).
45. Houston, K. et al. Analysis of the barley bract suppression gene Trd1. *Theor. Appl. Genet.* **125**, 33–45 (2012).
46. Leigh, J. W. & Bryant, D. popart: full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
47. Burton, R. A., Shirley, N. J., King, B. J., Harvey, A. J. & Fincher, G. B. The Cesa gene family of Barley. Quantitative analysis of transcripts reveals two groups of co-expressed genes. *Plant Physiol.* **134**, 224–236 (2004).
48. Athman, A. et al. Protocol: a fast and simple in situ PCR method for localising gene expression in plant tissue. *Plant Methods* **10**, 29 (2014).
49. Batistić, O., Sorek, N., Schultke, S., Yalovsky, S. & Kudla, J. Dual fatty acyl modification determines the localization and plasma membrane targeting of CBL/CIPK Ca<sup>2+</sup> signaling complexes in *Arabidopsis*. *Plant Cell* **20**, 1346–1362 (2008).
50. Šali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
51. Cotsaftis, O., Plett, D., Shirley, N., Tester, M. & Hrmova, M. A two-staged model of Na<sup>+</sup> exclusion in rice explained by 3D modeling of HKT transporters and alternative splicing. *PLoS ONE* **7**, e39865 (2012).
52. Waters, S., Gilliland, M. & Hrmova, M. Plant high-affinity potassium (HKT) transporters involved in salinity tolerance: structural insights to probe differences in ion selectivity. *Int. J. Mol. Sci.* **14**, 7660–7680 (2013).
53. Brooks, B. R. et al. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009).
54. Shen, M. Y. & Sali, A. Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507–2524 (2006).
55. Eswar, N., Eramian, D., Webb, B., Shen, M. Y. & Sali, A. Protein structure modelling with MODELLER. *Methods Mol. Biol.* **426**, 145–159 (2008).
56. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
57. Sippl, M. J. Recognition of errors in three-dimensional structures of proteins. *Proteins* **17**, 355–362 (1993).
58. Schymkowitz, J. et al. The FoldX web server: an online force field. *Nucleic Acids Res.* **33**, 382–388 (2005).
59. Pettersen, E. F. et al. UCSF Chimera— $\alpha$  visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**(13), 1605–1612 (2004).
60. Landau, M. et al. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* **33**, 299–302 (2005).
61. Celniker, G. et al. ConSurf: using evolutionary data to raise testable hypotheses about protein function. *Isr. J. Chem.* **53**, 199–206 (2013).
62. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).

## Acknowledgements

R.W., M.S., and M.M. acknowledge support from ERC project 669182 ‘SHUFFLE’ to R.W. K.H., P.S., M.B., J.R., and R.W. acknowledge the Rural & Environment Science & Analytical Services Division of the Scottish Government. C.B., J.Q., and Y.Q. acknowledge support from Rutherford Fund Strategic Partner Grants 2018 - Award Reference: RF-2018-30 to C.H. and R.W. S.R. is grateful for financial assistance from The Australian Research Council Industrial Transformation Research Hub for Wheat in a Hot and Dry Climate (IH130200027), the Grains Research and Development Corporation (ACP00009) and The Waite Research Institute, University of Adelaide. C.B. acknowledges support from the Grains Research and Development Corporation (GRDC) and the Australian Research Council (FT180100476). We are thankful to the Australian Research Council for funding through CE140100008 to M.G. (J.Q. and Y.Q.), FT180100476 to C.S.B. and DE160100804 to S.W. (A.S.). D.E.S. acknowledges support from BBSRC Grant BB/L000113/1. MH acknowledges financial support from the Huaiyin Normal University, China.

## Author contributions

R.W., D.E.S., M.G., C.B., K.H., J.Q., J.R., and S.R. designed experiments. K.H., J.Q., S.W., Y.Q., P.S., A.S., M.M., and P.F. carried out experiments. K.H., M.H., M.S., H.O., and M.B. analysed data. The manuscript was written by K.H. and R.W. with contributions from all other authors.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s42003-020-0990-5>.

**Correspondence** and requests for materials should be addressed to M.G., D.E.S. or R.W.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© Crown 2020