This is a repository copy of *Identification of candidate genes for calcium and magnesium accumulation in Brassica napus L. by association genetics*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/123821/

Version: Accepted Version

Article:

Alcock, T. D., Havlickova, Lenka orcid.org/0000-0002-5874-8615, He, Zhesi orcid.org/0000-0001-8335-9876 et al. (4 more authors) (2017) Identification of candidate genes for calcium and magnesium accumulation in Brassica napus L. by association genetics. Frontiers in Plant Science. ISSN: 1664-462X

Identification of candidate genes for calcium and magnesium accumulation in *Brassica napus* L. by association genetics

Running title: Calcium and magnesium accumulation candidates

Thomas D Alcock[1], Lenka Havlickova[2], Zhesi He[2], Ian Bancroft[2], Philip J White[3, 4], Martin R Broadley[1], Neil S Graham[1*]

[1]Plant and Crop Sciences Division, Sutton Bonington Campus, University of Nottingham, Loughborough, Leicestershire, UK.
[2]Department of Biology, University of York, Heslington, York, UK.
[3]The James Hutton Institute, Invergowrie, Dundee, UK.
[4]Distinguished Scientist Fellowship Program, King Saud University, Riyadh, Kingdom of Saudi Arabia.

*Correspondence:
Neil S Graham
neil.graham@nottingham.ac.uk

Abstract

Calcium (Ca) and magnesium (Mg) are essential plant nutrients and vital for human and animal nutrition. Biofortification of crops has previously been suggested to alleviate widespread human Ca and Mg deficiencies. In this study, new candidate genes influencing the leaf accumulation of Ca and Mg were identified in young *Brassica napus* plants using associative transcriptomics of ionomics datasets. A total of 247 and 166 SNP markers were associated with leaf Ca and Mg concentration, respectively, after false discovery rate correction and removal of SNPs with low second allele frequency. Gene expression markers at similar positions were also associated with leaf Ca and Mg concentration, including loci on chromosomes A10 and C2, within which lie previously identified transporter genes *ACA8* and *MGT7*. Further candidate genes were selected from seven loci and the mineral composition of whole *Arabidopsis thaliana* shoots were characterised from lines mutated in orthologous genes. Four and two mutant lines had reduced shoot Ca and Mg concentration, respectively, compared to wild type plants. Three of these mutations were found to have tissue specific effects; notably reduced silique Ca in all three such mutant lines. This knowledge could be applied in targeted breeding, with the possibility of increasing Ca and Mg in plant tissue for improving human and livestock nutrition.

1 Introduction

Calcium (Ca) and magnesium (Mg) are essential plant nutrients and vital for human and animal nutrition (Broadley and White, 2010; White and Brown, 2010). In plants, most Ca is extracellular, where it is a key strengthening component in cell walls (Grusak *et al.*, 2016). It also has an important role in plant-cell signalling. Calcium enters root cells through a variety of Ca^{2+}-permeable cation channels (Karley and White, 2009; White and Broadley 2003; White, 2015). The opening of these channels must be tightly controlled, as changes in cytosolic Ca^{2+} concentrations coordinate numerous developmental and environmental stress responses

50 (McAinsh and Pittman, 2009). The accumulation of Ca at tissue- and cellular-levels is
51 dependent on the expression of transport proteins (Conn and Gilliham, 2010; Rios *et al*., 2012).
52 After uptake from soil by roots, Ca travels via either apoplastic or symplastic pathways to the
53 xylem, through which, in the form of either Ca^{2+} or complexed with organic acids, it is
54 transported to the shoot. Calcium is immobile in the phloem, and as such, tissues with low
55 transpiration rates (including fruits, seeds and tubers) often have low Ca concentrations (Karley
56 and White, 2009). Among plant nutrients, Ca is required in relatively large amounts. However,
57 concentrations vary amongst taxa, typically ranging from ~0.1 % to 4.4 % dry matter (Broadley
58 *et al*., 2003). Calcium deficiencies are relatively rare in field-grown crops, but can occur in
59 crops grown in acidic or leaching prone soils. Where Ca supply is insufficient to meet growth
60 requirements, costly symptoms can ensue. For instance, fruits lacking in Ca are prone to
61 cracking, as a direct result of weakness in the cell wall (White, 2015).

62

63 Magnesium is essential for photosynthesis, forming the central atom of chlorophyll molecules.
64 It also has a key role in protein synthesis by functioning as a bridging element for the
65 aggregation of ribosome subunits, as well as in photophosphorylation and generation of
66 reactive oxygen species in plants (Cakmak and Yazici, 2010). Magnesium is taken up by roots
67 as Mg^{2+}. Control of influx across the plasma membrane is dominated by members of the
68 MGT/MRS2 family of transport proteins and potentially Mg^{2+}-permeable cation channels
69 (Karley and White, 2009; Lenz *et al*., 2013). One member of the MGT gene family in
70 *Arabidopsis thaliana*, *MAGNESIUM TRANSPORTER 1* (*MGT1*), encodes a protein localised
71 to the plasma membrane (Li *et al*., 2001), suggesting its importance in the import and/or export
72 of Mg in cells. Like Ca, Mg is transported from root to shoot cells through the xylem either as
73 Mg^{2+} or complexed with organic acids. However, Mg is a phloem-mobile element, and is
74 readily translocated to fruit, seeds and tubers (White and Broadley, 2009). Shoot Mg
75 concentrations are typically lower than shoot Ca concentrations across plant taxa, and vary
76 between ~0.1 % to ~1.0 % dry matter (White *et al*., 2015).

77

78 In humans and animals, Ca is associated with the formation and metabolism of bone as well as
79 being crucial for mediating vascular contraction and vasodilation, muscle function, nerve
80 transmission, intracellular signalling and hormonal secretion (Catharine, 2011). Based on food
81 supply data, it was estimated that half of the population worldwide was at risk of Ca deficiency
82 in 2011, with significant deficiency risks across all continents (Kumssa *et al*., 2015a).
83 Magnesium is needed for over 300 biochemical reactions. It helps to maintain muscle function,
84 prevents an irregular heartbeat, and is involved in protein synthesis (Yardley, 2009). Based on
85 food supply data, <1% of the global population appeared to be at risk of dietary Mg deficiency
86 in 2011 (Kumssa *et al*., 2015b). However, these data do not account for inhibitors of Mg
87 adsorption, household waste, or distribution within countries and it is likely that significant
88 deficiency risks exist within some populations. Magnesium deficiency risks are also likely to
89 be greater in higher-income groups consuming processed foods, because Mg is among the
90 nutrients commonly lost in processing (Broadley and White, 2010; Kumssa *et al*., 2015b;
91 Swaminathan, 2003). Biofortification of crops has been previously suggested as a suitable
92 approach for alleviating human deficiencies in a number of mineral nutrients, including Ca and
93 Mg (White and Broadley, 2009; White and Broadley, 2005).

94

95 Previous analyses of variation in mineral concentrations across a wide range of plant species
96 have shown that tissue Ca and Mg concentrations are inherently high in Brassicaceae compared
97 most other taxa (Broadley *et al*., 2004; White *et al*., 2015). These traits have proven to be
98 heritable in *Brassica oleracea* (Broadley *et al*., 2008), *B. rapa* (Graham *et al*., 2014), and *B.*
99 *napus* (Thomas *et al*., 2016). Thus, *Brassica* spp. are potentially good targets for understanding

genetic bases of leaf Ca and Mg accumulation, and for potentially increasing dietary intakes of Ca and Mg in humans and animals. Expression quantitative trait locus (eQTL) analyses in *B. rapa* previously led to the discovery of Ca responsive genes which may prove useful in marker-assisted selection for increased Ca concentration in shoot tissue (Graham *et al*., 2014). These include orthologues of *A. thaliana* Ca^{2+} transporter genes *CATION EXCHANGER 1* (*CAX1*) and *AUTOINHIBITED CA^{2+} ATPASE, ISOFORM 8* (*ACA8*), and subsequent work showed that allelic variants of the former gene in *B. rapa* influenced Ca accumulation. *B. napus* includes oilseed types, swedes and fodder crops, and is widely cultivated globally. It is an amphi-diploid species that likely originated from multiple spontaneous hybridisations between *B. rapa* (A genome; turnip rape) and *B. oleracea* (C genome; cabbage, kale) and contains a full set of chromosomes from each (Chalhoub *et al*., 2014; Iniguez-Luy and Federico, 2011). This complexity has previously hindered the genetic study of this and other polyploid crops. However, recent and ongoing advances in sequencing and genome mapping technologies have allowed the rapid genotyping of multiple accessions at a fraction of the cost of older technologies. This has improved the feasibility of using a large diversity population over traditional mapping populations in genetic studies of crop species (Trick *et al*., 2009).

Associative transcriptomics (Harper *et al*., 2012) focusses on the analysis of transcribed sequences (mRNA-seq) across diversity populations to identify high-resolution loci influencing complex traits. An advantage using of RNA over DNA sequences in association studies is the ability to develop markers based on both single-nucleotide polymorphisms (SNPs) and transcript abundance (gene-expression markers; GEMs; Harper *et al*., 2012). Gene expression levels may be particularly important in the control of traits in polyploid species in which gene duplication may have led to unequal expression (Adams *et al*., 2003). Associative transcriptomics has been recently used in *B. napus* to identify genes underlying control of seed glucosinolate content (Harper *et al*., 2012; Lu *et al*., 2014) and anion homeostasis (Koprivova *et al*., 2014). The former two studies utilised panels of 84 and 101 genotypes, respectively. Despite the relatively small population sizes, a number of loci associated with seed glucosinolate concentrations were identified. Most notable associations include loci containing orthologues of *A. thaliana HIGH ALIPHATIC GLUCOSINOLATE 1* and *3* (*HAG1* and *HAG3*), known to regulate aliphatic glucosinolate biosynthesis (Sønderby *et al*., 2010). Koprivova *et al*. (2014), also made use of the panel of 84 genotypes and identified a number of loci associated with leaf nitrate, phosphate and sulphate. Within these loci were a number of clear candidate genes, including a calcium-activated chloride channel previously shown to control nitrate levels in *A. thaliana* (De Angeli *et al*., 2006) which was associated with leaf nitrate concentration and a hypothetical phosphate/phosphoenolpyruvate translocator associated with leaf phosphate concentration.

Leaf Ca and Mg concentrations were previously characterised in a diversity population of ~400 genotypes of *B. napus* in a broad-spectrum mineral analysis (Thomas *et al*., 2016). This population is likely to capture most of the species-wide variation, comprising oilseed, swede and fodder types. In this study, we perform associative analyses on this data using transcriptome sequences from 383 genotypes to identify genes influencing Ca and Mg accumulation. Candidate genes could be applied in marker assisted breeding in this and other *Brassica* crops, with the possibility of improving nutrient use efficiency of the crop and increasing available nutrients in edible plant tissue for improving human and livestock nutrition.

2 Materials and methods

150 **2.1 Characterisation of leaf Ca and Mg concentration**

151

152 This study used the Renewable Industrial Products from Rapeseed (RIPR) diversity population
153 of inbred lines of *Brassica napus* genotypes (Thomas *et al.*, 2016). These were developed from
154 the ERANET-ASSYST consortium diversity population (Bus *et al.*, 2014; Bus *et al.*, 2011;
155 Körber *et al.*, 2015; Körber *et al.*, 2012) with further lines included. A subset of 383 genotypes
156 were selected, comprising 160 winter-, 127 spring-, and seven semiwinter-oilseed rape (OSR),
157 35 swede, 15 winter fodder, and 39 exotic/unspecified habits. These were previously
158 characterised for leaf mineral concentrations by inductively coupled plasma-mass spectrometry
159 (ICP-MS) of polytunnel-grown plants sampled at the rosette stage (typically 6-8 true leaves
160 showing; Thomas *et al.*, 2016). The full leaf mineral dataset is available at the Brassica
161 Information Portal (BIP; https://bip.earlham.ac.uk/; The Earlham Institute, Norwich, UK) and
162 at doi:10.5281/zenodo.59937.

163

164 **2.2 Associative analyses**

165

166 **2.2.1 Transcriptome sequencing and population structure analysis**

167

168 Extraction of RNA, quality checking and Illumina transcriptome sequencing were carried out
169 as described by He *et al*. (2016). Tissue samples for RNA extraction were prepared from second
170 true leaves, harvested when they reached ~3 cm in diameter. RNA-seq data from each accession
171 was mapped using methods described by Bancroft *et al.* (2011) and Higgins *et al.* (2012) onto
172 ordered *Brassica* A and C genome-based pan-transcriptomes developed by He *et al.* (2015).
173 Transcriptome sequencing was performed by the Earlham Institute (formerly The Genome
174 Analysis Centre; Norwich, UK). Across the 383 accession panel, 46,307 single SNPs and
175 309,229 hemi-SNPs were detected and scored of which 256,397 SNPs had a population second
176 allele frequency (saf) > 0.01. Transcript abundance was quantified and normalised as reads per
177 kb per million aligned reads (RPKM) for each accession for 116,098 coding DNA sequence
178 (CDS) models of the pan-transcriptome reference. Significant expression (mean >0.4 RPKM)
179 was detected for 53,889 CDS models. Inference of population structure by Q-matrix was
180 obtained by Population Structure Inference using Kernel-PCA and Optimization (PSIKO;
181 Popescu *et al.*, 2014). A heatmap illustrating the relatedness of all genotypes in this study can
182 be found in Supplementary Figure 1. Transcriptome sequences are deposited within the
183 Sequence Read Archive (Leinonen *et al.*, 2011) under accession number PRJNA309367.

184

185 **2.2.2 Associative transcriptomics**

186

187 Associative transcriptomics was performed using SNPs, Q-matrix and trait data in a
188 compressed mixed linear model approach (Zhang *et al.*, 2010) implemented in the GAPIT R
189 package (Lipka *et al.*, 2012) in R 3.2.0 (R Core Team, 2015). The association analysis between
190 gene expression markers (GEMs) and traits was performed by using fixed effect linear
191 modelling in R with RPKM values and Q-matrix data as the explanatory variables and trait
192 score the response variable, with scripts developed by Harper *et al.* (2012). Coefficients of
193 determination (R^2), constants and significance values were calculated for each regression.
194 Manhattan plots were generated using graph functions in R. SNPs with low second allele
195 frequency (<0.01) were filtered from the dataset prior to generating plots. In total 256,397 SNPs
196 and 53,889 GEMs were plotted. False Discovery Rate (FDR; Benjamini and Hochberg, 1995)
197 and Bonferroni (Dunn, 1961) corrections were used to set significance thresholds at *P<0.05*.
198 Due to sequence similarity between *B. napus* A and C genomes, assignment to a specific
199 genome was not possible for all SNP markers; such markers are plotted in grey and appear in

200 both positions on Manhattan plots. See Supplementary Figures 2-5 for quartile-quartile (QQ)
201 plots of each association analysis.
202
2.2.3 Candidate gene identification

205 Ordered pan-transcriptome data based on *Brassica* A and C genomes from *B. rapa*, *B. napus*,
206 and *B. oleracea* CDS gene models (He *et al.*, 2015) were used to identify candidate genes.
207 Candidate genes were selected based on *Arabidopsis thaliana* annotated functions of *Brassica*
208 orthologues within estimated linkage disequilibrium decay of significantly associated markers
209 (around 1-2 cM on average; Ecke *et al.*, 2010). Further information relating to candidate gene
210 predicted function was obtained from genome browsers comprising sequences of *B. rapa* (A
211 genome, Chiifu-401-41; Wang *et al.*, 2011) and *B. oleracea* (C genome, TO1000DH3; Parkin
212 *et al.*, 2014) at Ensembl Plants (Kersey *et al.*, 2016). *A. thaliana* functional information were
213 obtained from The Arabidopsis Information Resource (TAIR; Huala *et al.*, 2001). Further
214 resources used to aid with selection of candidates included *A. thaliana* gene expression data at
215 The Bio-Analytic Resource for Plant Biology (Waese and Provart 2017) and ionomic data at
216 the Purdue Ionomics Information Management System (PIIMS; Baxter *et al.*, 2007).
217
2.3 Experiments using *Arabidopsis thaliana* mutants

2.3.1 Plant material and genotyping

222 Seed of 15 *Arabidopsis thaliana* mutant lines representing 10 candidate genes were acquired
223 from the Nottingham Arabidopsis Stock Centre (Nottingham, UK). These comprised SALK
224 (Alonso *et al.*, 2003) and SAIL (Sessions *et al.*, 2002) T-DNA lines and are summarised in
225 Table 1. *Arabidopsis thaliana* ecotype Columbia-0 (Col-0) was used as the wild type control
226 in all experiments. Plants were by genotyped for homozygous T-DNA insertions by
227 conventional PCR. Genotyping primers are summarised in Table 1. Left border primers used
228 were SALK LBb1 and SAIL LB1 for SALK and SAIL lines respectively.
229
2.3.2 Preliminary phenotyping

232 Seeds from homozygous mutant lines and Col-0 were sterilized in bleach, then washed in H_2O
233 and 70% ethanol prior to sowing on plates containing 1% agar containing 0.4 g L^{-1} Hoagland's
234 solution (Hoagland and Arnon 1950; ¼ strength). Plates were stored in darkness at 4°C for 24
235 h, and then moved to a controlled environment growth chamber set to 23°C (~30 W m^{-2}
236 continuous light). After seven days, plants were transferred to pots containing Levington M3
237 compost (ICL Specialty Fertilizers, Ipswich, Suffolk, UK) plus T34 biocontrol (Fargro Ltd,
238 Arundel, West Sussex, UK) and placed on flow benches in a glasshouse with 18°C heating,
239 venting at 20°C, with 16 hour supplementary lighting (76 W m^{-2}). Flow bench automatic
240 irrigation operated once daily. After 10 days of establishment, six plants of each line were
241 chosen randomly and transferred to individual wells in 16 well trays in a six block, using a one-
242 way randomised design generated in GenStat (17th edition; VSN International, 2014) in which
243 plants of each line were represented once per block and randomised within each block
244 (Supplementary Table 1). In total, each line was represented six times. ARACON systems
245 (Betatech BVBA, Gent, Belgium) were used to keep plants separate. At mid-flowering, whole
246 shoots were harvested by cutting them below the rosette. Shoots were dried at 50°C for at least
247 two days, and then crushed by hand within paper bags. Shoot subsamples (~0.10 g DW) were
248 digested using a microwave system comprising a Multiwave 3000 platform with a 48-vessel
249 MF50 rotor (Anton Paar GmbH, Graz, Austria). Digestion vessels were perfluoroalkoxy (PFA)

liner material and polyethylethylketone (PEEK) pressure jackets (Anton Paar GmbH). Leaf material was digested in 2 mL 70% Trace Analysis Grade HNO_3, 1 mL Milli-Q water (18.2 MΩ cm; Fisher Scientific UK Ltd, Loughborough, UK), and 1 mL H_2O_2 with microwave settings as follows: power = 1400 W, temp = 140 °C, pressure = 2 MPa, time = 45 min. Two operational blanks and duplicate samples of certified reference material (CRM; Tomato SRM 1573a, NIST, Gaithersburg, MD, USA) were included in each digestion run. Following digestion, each tube was made up to a final volume of 15 mL by adding 11 mL Milli-Q water and transferred to a 25 mL universal tube (Sarstedt Ltd., Nümbrecht, Germany) and stored at room temperature. Leaf digestates were diluted 1-in-5 using Milli-Q water prior to broad-spectrum elemental analysis by ICP-MS as described previously (Thomas *et al.*, 2016). For each data-point, an element-specific operational blank concentration (mean of each ICP-MS run) was subtracted. Data were then multiplied by initial sample volume, divided by the initial dry mass of plant material, and converted to mg element kg^{-1} of dry leaf or seed material. The CRM Ca and Mg recovery averaged 99 and 89% respectively.

2.3.4 Tissue partitioning experiment

Based on results from preliminary phenotyping, lines At2g13610.2, At5g07320.2 and At5g48650.2 were found to have significantly lower shoot Ca or Mg concentrations than wild type plants and hence were selected for further characterisation. Individual seed from these lines and Col-0 were sown into 12 well trays containing Levington M3 compost plus T34 biocontrol and placed on flow benches in a glasshouse with 18°C heating, venting at 20°C, with 16 hour supplementary lighting (76 W m^{-2}). Flow bench automatic-irrigation operated once daily. After successful establishment, 12 plants per genotype were selected randomly and transplanted into individual 9 cm pots. These were arranged in a 12 block, one-way randomised design generated in GenStat in which each genotype was represented once per block and genotypes randomised within each block (Supplementary Table 2). At mid-flowering (40 days after sowing), entire shoots were harvested. Shoots were partitioned into rosette leaves, stem, stem leaves, and siliques. Tissue samples were dried at 50°C for six days, and then samples from plants in blocks 1-4, 5-8, and 9-12 were pooled into the four genotypes to ensure enough sample was available for mineral analysis. Pooled samples were crushed by hand, and then microwave digested prior to mineral analysis by ICP-MS as described above. Digestates were diluted 1-in-10 prior to mineral analysis. The recovery of Ca and Mg from the CRM averaged 96 and 88% respectively.

2.3.5 Statistical analyses

Data from experiments using *A. thaliana* mutants were analysed using one-way ANOVA in GenStat (17th edition; VSN International, 2014) with block design included in the model. Tissue Ca and Mg concentration data were analysed separately in each case, and tissue types were analysed using separate ANOVA tests in the tissue partitioning experiment. For the preliminary phenotyping experiment, six replicate plants were analysed for each genotype. For the tissue partitioning experiment, three samples, each comprising pooled samples from four replicate plants, were analysed for each genotype. Means of different *A. thaliana* genotypes were compared using Least Significant Difference (LSD) functions in GenStat with differences considered significant at *P<0.05*. Further LSD tests were conducted at *P<0.01* and *P<0.001* levels.

3 Results

3.1 Variation in leaf Ca and Mg concentration in the RIPR diversity population

The leaf concentrations of 21 mineral elements including Ca and Mg in the RIPR diversity population were previously determined by Thomas *et al.* (2016). Leaf Ca concentrations varied over 3-fold across the population, from 5,838 mg kg^{-1} to 18,752 mg kg^{-1}. Leaf Mg concentrations were of a similar order of magnitude and varied over 2-fold, from 5,118 mg kg^{-1} to 13,429 mg kg^{-1}. The frequency distribution of these two elements approximated a normal distribution (Supplementary Figure 6). Leaf Ca and Mg concentrations were among the highest, positively correlated elements measured across genotypes and tissues, with an *r* value of 0.87 (*P<0.001*). Leaf Ca and Mg concentrations varied between crop type, with higher concentrations of both elements in leaves of spring and semiwinter OSR than in winter OSR, winter fodder, and swede types.

3.2 Associative transcriptomics suggest flowering time regulators are important markers for leaf Ca and Mg concentrations

To identify candidate loci, SNPs and GEMs were used separately in analyses. A total of 1295 and eight SNPs were found to be significantly associated with *B. napus* leaf Ca concentration after FDR and Bonferroni corrections, respectively. After removing SNPs with low second allele frequency, this was reduced to 247 and five SNPs respectively across all chromosomes. Visually determined association peaks on Manhattan plots were observed on chromosomes A3, A6, A7, A10, C2, C3 and C9 (Fig 1A). The most well defined peak was located on chromosome A10 and contained four out of the five SNPs above the Bonferroni corrected significance threshold (*P=0.05*). The fifth SNP above this threshold fell in a peak on chromosome C9, in a region known to share sequence homology with parts of chromosome A10 (Chalhoub *et al.*, 2014). A total of 5557 and 141 GEMs were identified as significantly associated with leaf Ca concentration after FDR and Bonferroni corrections respectively (Fig 1B). Notable peaks were observed on chromosomes A2 and C2. Single, associated GEMs were found at similar locations to SNP association peaks on chromosomes A3 and C2. The *A. thaliana* orthologue of *B. napus* genes corresponding to both these GEMs is At5g10140, which encodes FLOWERING LOCUS C (FLC), a transcription factor important for controlling flowering time (Michaels and Amasino 1999). A further associated GEM was found in a region of chromosome A10, close to a SNP peak associated with leaf Ca concentration. Other single GEMs associated with leaf Ca concentration were observed on chromosomes A5, A6, C4 and C6. *B. napus* genes corresponding to these GEMs on chromosomes A5 and C4 are orthologous to *A. thaliana* At2g45660, which encodes SUPPRESSOR OF OVEREXPRESSION OF CO 1 (SOC1), another flowering time regulator (Lee *et al.*, 2000).

The SNP and GEM associations for leaf Mg concentration were similar to those described for leaf Ca concentration. A total of 1012 and one significant SNP(s) were found after FDR and Bonferroni corrections, respectively. After removing SNPs with low second allele frequency, this was reduced to 166 and zero SNPs respectively across all chromosomes. Of these 166 SNPs, 86 were identical to SNPs identified as significantly associated with leaf Ca concentration after FDR correction and removal of SNPs with low second allele frequency, indicating the potential of similar mechanisms to partly regulate accumulation. Visually determined association peaks largely co-localised to those associated with leaf Ca concentration, specifically on chromosomes A3, A7, A10, C2, C3 and C9 (Fig 1C). The most well defined peak was again on chromosome A10, and as before, a region on chromosome C9 with sequence homology to this region also contained associated SNPs. An association peak on chromosome A2 was also particularly well defined, containing 15 SNPs above the FDR

350 corrected significance threshold. A total of 12973 and 1489 GEMs were identified as
351 significantly associated with leaf Mg concentration after FDR and Bonferroni corrections
352 respectively across all chromosomes (Fig 1D). Of these, 5160 and 131 were also significantly
353 associated with leaf Ca concentration after FDR and Bonferroni corrections respectively.
354 Notable peaks were again observed on chromosomes A2 and C2. The most highly associated
355 GEMs on A3 and C2 were identical to those associated with leaf Ca concentration which
356 correspond to *A. thaliana FLC*, and an associated GEM on C4 is identical to the GEM
357 associated with leaf Ca concentration corresponding to *SOC1*.

3.3 Genes encoding previously identified Ca and Mg transporters are within linkage
disequilibrium of highly associated markers for leaf Ca and Mg concentration

362 Linkage disequilibrium (LD) describes the non-random association of alleles at different loci
363 (Slatkin, 2008). Genes located physically near to each are generally inherited together, and
364 hence are often in very strong LD. It is therefore feasible that any number of genes within LD
365 of SNPs significantly associated with a trait may be controlling such associations. Based on
366 associative transcriptomics results, seven loci were focussed on for the identification of
367 candidate genes. These comprised regions of chromosomes A2, A3, A5, A6, A10, C2 and C4.
368 A total of 17 *B. napus* candidate genes orthologous to 15 *A. thaliana* genes are summarised in
369 Table 2. Four candidate genes were selected based on direct GEM hits as described above.
370 These are Cab002472.4 and BnaC02g00490D (on chromosomes A3 and C2 respectively)
371 encoding orthologues of *A. thaliana* At5g10140 (*FLC*), and Cab025356.1 and Bo4g024850.1
372 (on chromosomes A5 and C4 respectively) encoding orthologues of *A. thaliana* At2g45660
373 (*SOC1*). One and two candidate genes on chromosomes A2 and C2 respectively are
374 orthologous to *A. thaliana MAGNESIUM TRANSPORTER 7* (*MGT7/MRS2-7*). This was
375 previously characterised in *Arabidopsis thaliana* as an Mg transporter important for Mg uptake
376 at low external concentrations (Gebert *et al.*, 2009). A further candidate gene on chromosome
377 A10 is orthologous to *A. thaliana AUTOINHIBITED CA^{2+} -ATPASE, ISOFORM 8* (*ACA8*).
378 This was previously characterised as a plasma membrane-localised Ca^{2+} transporting ATPase
379 in *A. thaliana* (Bonza *et al.* 2000) and was identified as Ca responsive in *B. rapa* (Graham *et*
380 *al.*, 2014). The functions of the remaining nine candidate genes were selected based on
381 sequence homology and annotations of *A. thaliana* orthologues and are either uncharacterised,
382 or have not previously been experimentally shown to be involved in plant Ca or Mg
383 accumulation (Table 2). These and At2g45660 (*SOC1*) were used for the selection of *A.*
384 *thaliana* mutants.

3.4 Four mutant *A. thaliana* lines have reduced shoot Ca and Mg concentration compared
to wild type and effects are tissue specific

389 Shoot Ca concentrations in a preliminary *A. thaliana* phenotyping experiment varied three-fold
390 between individual plants, from 8,684 to 26,387 mg kg^{-1} dry weight (DW; Supplementary
391 Table 3). Much of this variation was observed within genotypes, with the largest variation
392 observed in lines At5g7320.1 and At5g08670.1. Four mutant lines had significantly lower
393 mean shoot Ca concentrations than wild type plants. These were At2g13610.2 (*P<0.05*),
394 At5g07320.2 (*P<0.01*), At5g08670.1 (*P<0.05*) and At5g48650.2 (*P<0.01*; Fig 2A). Leaf Mg
395 concentrations varied less, with two-fold variation from 8,189 to 16,186 mg kg^{-1} DW observed
396 between individual plants. Much of this variation was between genotypes. Two mutant lines
397 had lower mean shoot Mg concentration than wild type plants. These were At2g13610.1
398 (*P<0.05*) and At5g48650.2 (*P<0.05*; Fig 2B). Based on these data, lines At2g13610.2,

399 At5g07320.2 and At5g48650.2 were selected for characterisation of tissue specific leaf Ca and
400 Mg concentration.
401
402 Calcium concentrations varied over eight-fold between tissues and pooled tissue samples,
403 ranging from 4,998 mg kg^{-1} DW in stems to 40,536 mg kg^{-1} DW in cauline leaves (Fig 3A-D,
404 Supplementary Table 4). Cauline and rosette leaf Ca concentrations were similar, ranging from
405 32,966 mg kg^{-1} DW in rosette leaves to 40,536 mg kg^{-1} DW in cauline leaves. Mean silique Ca
406 concentrations were lower in lines At2g13610.2, At5g07320.2 and At5g48650.2 than wild type
407 plants (*P<0.01*; Fig 3B). Mean stem Ca concentrations were lower in lines At5g07320.2 and
408 At5g48650.2 than wild type plants (*P<0.05*; Fig 3C). Mean stem leaf Ca concentration was
409 lower in line At5g48650.2 than wild type plants (*P<0.05*; Fig 3D). Mg concentrations varied
410 over nine-fold between tissues and pooled samples, ranging from 2,608 mg kg^{-1} DW in stems
411 to 23,999 mg kg^{-1} DW in rosette leaves (Fig 4A-D). Cauline leaf and rosette leaf Mg
412 concentrations had a similar range, from 18,026 to 21,328 mg kg^{-1} DW and from 19,240 to
413 23,999 mg kg^{-1} DW respectively. Lines At2g13610.2 and At5g48650.2 had lower mean silique
414 Mg than wild type plants (*P<0.05*; Fig 4B). Line At5g48650.2 also had lower mean stem Mg
415 than wild type plants. Finally, and comparable with results from cauline leaf Ca concentration
416 analysis, mean cauline leaf Mg concentration was lower in line At5g48650.2 than in wild type
417 plants (*P<0.05*; Fig 4D).
418
419 In summary, data from *A. thaliana* experiments identified four and two mutant lines with lower
420 shoot Ca and Mg concentrations than wild type plants respectively and three of these mutations
421 have tissue specific phenotypes. The main tissue specific effects were observed in silique
422 tissue, with lower silique Ca concentrations in all three mutant lines investigated in the tissue
423 partitioning experiment.
424
425 **4 Discussion**
426
427 **4.1 SNP based association analyses identify novel and confirm pre-determined candidate**
428 **loci for leaf Ca and Mg concentrations**
429
430 Leaf Ca concentration was highly associated with loci on chromosomes A3, A6, A7, A10, C2,
431 C3 and C9 (Fig 1A). Similar loci were associated with leaf Mg concentration, specifically in
432 regions of chromosomes A3, A7, A10, C2, C3 and C9 (Fig 1C). The most highly associated
433 SNP for leaf Ca concentration was located on chromosome A10 (Fig 1A). This co-localises
434 with associated markers on C9 and markers on A10 and C9 for leaf Mg concentration. Co-
435 localisation of association peaks and associated markers for both mineral elements is
436 unsurprising, as leaf Ca and Mg concentration data used in this study were very highly
437 correlated (*r* = 0.87, *P<0.001*; Thomas *et al*., 2016) and may reflect the relative lack of
438 selectivity between these and other group II elements during accumulation within the plant
439 (White, 2001). Such correlations between shoot Ca and Mg concentration have been previously
440 shown in *B. oleracea* (Broadley *et al.*, 2008) and a number of other angiosperm species
441 (Broadley *et al.*, 2004). Bus *et al.* (2014) previously investigated the genetic control of shoot
442 ionome traits across 505 lines of *B. napus* using 3,910 SNPs in association analyses. Results
443 showed two associations at a locus on chromosome C9 for shoot Ca and Mg concentration with
444 a further association on chromosome C7 for shoot Ca concentration. The detection of an
445 association locus on C9 is consistent with co-localised associations identified in this study.
446 These results are also consistent with earlier findings by Broadley *et al.* (2008) who identified
447 significant QTL for shoot Ca and Mg in *B. oleracea* on chromosomes C2, C6, C7, C8, and C9.
448 Together, these results indicate the importance of loci on chromosomes A10 and C9 for Ca and

449 Mg accumulation. The QTL identified for shoot Mg in *B. oleracea* on chromosome C2 by
450 Broadley *et al.* (2008) is also consistent with findings in the present study that a locus on C2 is
451 highly associated with leaf Mg concentration in *B. napus*. However, further work is required
452 to confirm whether the loci are in close proximity to one another. To our knowledge, the
453 remaining loci identified in this study have not been previously identified as important QTL
454 for leaf Ca and Mg concentration in *Brassica* spp.

4.2 FLC and SOC1 GEM associations may be linked to variation in leaf Ca and Mg
concentrations between spring and winter *B. napus* types in the RIPR panel

459 GEM analyses associated markers corresponding to *FLC* and *SOC1* with both leaf Ca and Mg
460 concentrations. In *A. thaliana*, FLC is a repressor of flowering (Michaels and Amasino 1999)
461 and it has been previously shown that FLC transcript concentration correlates with
462 vernalisation requirements (Sheldon *et al.* 2000). Expression levels of *SOC1* also correlate with
463 flowering time in *A. thaliana*; in lines which flower later, *SOC1* expression is very low (Lee *et*
464 *al.*, 2000). It is thought that *SOC1* expression is repressed by *FLC*, indicating the tight
465 regulatory links between these genes and flowering time. Leaf Ca and Mg concentration data
466 used in this study were obtained from analysis of plants in the RIPR panel (Thomas *et al.*,
467 2016) which includes a large number of spring and winter *B. napus* varieties. Thomas *et al.*
468 (2016) observed differences in leaf Ca and Mg concentrations between these types, with higher
469 mean concentrations of both leaf Ca and Mg in spring OSR compared to winter OSR and winter
470 fodder types. Since winter OSR types are generally considered to have longer vernalisation
471 requirements than spring types, it is possible that the association of *FLC* and *SOC1* with leaf
472 Ca and Mg concentration observed in this study was a result of differences in vernalisation
473 requirement between these groups rather than direct genetic control of Ca and Mg uptake. It is
474 worth noting that the association of GEMs with a trait does not indicate the causative
475 polymorphism/s, only genes in which expression level is associated with variation in the trait.
476 The causative polymorphism/s may lie in the promotor sequence of such genes, or localise
477 somewhere upstream in the pathway. Hence, in the case of the flowering time genes identified
478 here, it is unclear whether or not the observed associations with leaf Ca and Mg concentration
479 are directly caused by changes in expression of *FLC* and *SOC1*. Despite this, their expression
480 appears to be a suitable marker for the concentrations of these elements in *B. napus*. Further to
481 this, the concentrations of a number of other mineral elements measured in the study of Thomas
482 *et al.* (2016) were found to vary between crop types with typically different flowering times
483 and vernalisation requirements. Most notably, leaf concentrations of Mo, Na, P and S were
484 higher in spring OSR than winter OSR types. This suggests that flowering time, or the upstream
485 mechanisms leading to changes in flowering time, has an effect on the concentrations of a
486 number of nutrients in *B. napus*, though the pathway/s that lead to these differences remain
487 unclear.

4.3 ACA8 and MGT7 are among genes within linkage disequilibrium of associated loci

491 Identification of high-resolution loci influencing leaf Ca and Mg concentrations enabled locus-
492 specific exploration of the *Brassica* pan-transcriptomes and other genome resources for
493 candidate genes within LD of SNPs. LD is especially relevant to the efficacy of associative
494 transcriptomics in the absence of a marker in a trait-controlling gene. LD decays relatively
495 quickly in *B. napus* (Harper *et al.*, 2012; Ecke *et al.*, 2010), and this helps to reduce the number
496 of possibilities when searching for candidate genes. However, in this study, typically hundreds
497 of genes were still within previously estimated LD decay (around 1-2 cM on average; Ecke *et*
498 *al.*, 2010) of most candidate loci. Fortunately, well annotated browsers of *Brassica* A and C

499 genomes are available at Ensembl Plants (Kersey *et al.*, 2016), which enabled rapid
500 identification of nearby genes in the reference sequences with links to functional annotation of
501 *A. thaliana* orthologues.

503 Most notable genes identified using this workflow include an orthologue of *A. thaliana ACA8*
504 near markers associated with leaf Ca on chromosome A10 and two orthologues of *A. thaliana*
505 *MGT7* near markers associated with leaf Mg on chromosomes A2 and C2. *ACA8* encodes a
506 Ca^{2+} transporting ATP-ase localised to the plasma membrane (Bonza *et al.*, 2000). A *B. rapa*
507 orthologue of *A. thaliana ACA8* was previously identified under an eQTL hot spot on
508 chromosome A3 (Graham *et al.*, 2014). The eQTL associated with this gene was defined as
509 Ca-responsive, i.e. the direction of the eQTL changed under high Ca supply. The *A. thaliana*
510 orthologue of *ACA8* was further investigated *in silico* in the same study using publically
511 available phenotypic data at the PIIMs database (Baxter *et al.*, 2007). This led to the
512 identification of *ACA8* T-DNA knockout mutants with greater shoot Ca concentrations than
513 control plants in over 50% of mutant samples, indicating the ability of this gene to influence
514 Ca accumulation in *Brassica. MGT7* is a member of the MGT/MRS2 Mg transport family. This
515 was previously characterised as a key transporter for Mg uptake at low external Mg
516 concentrations by Gebert *et al.* (2009). *Arabidopsis thaliana* T-DNA knockout mutants were
517 severely retarded in development when grown at low external Mg concentrations, but were
518 visually unaffected when grown at higher external Mg concentrations. Both *ACA8* and *MGT7*
519 are very promising candidate genes for the control of Ca and Mg accumulation in *B. napus*.
520 The presence of these genes within LD of highly associated SNPs demonstrates the
521 effectiveness of associative transcriptomics in candidate gene identification. Since *ACA8* and
522 *MGT7* knockout mutants had previously been characterised in *A. thaliana*, they were not
523 included in further experiments in this study.

4.4 *Arabidopsis thaliana* mutant phenotyping reveals new candidates for Ca and Mg accumulation

528 The preliminary *A. thaliana* phenotyping experiment identified four mutant lines with lower
529 shoot Ca concentrations and two with lower shoot Mg concentrations than wild type plants.
530 The most notable of these was At5g48650.2, the only line in which both shoot Ca and Mg was
531 affected. The gene mutated in this line encodes NUCLEAR TRANSPORT FACTOR 2
532 (NTF2). This protein is proposed to function in the import of RAN, a multifunctional GTPase
533 involved in nucleocytoplasmic transport (Zhao *et al.*, 2006). It is the first time that it has been
534 characterised with a shoot Ca and Mg phenotype in *A. thaliana*. Further investigation of this
535 line showed that it had lower Ca and Mg concentration than wild type plants in all tissues
536 except rosette leaves, suggesting it could be a promising candidate for manipulating the
537 translocation of Ca and Mg to specific tissues in crop plants. At5g07320 encodes the ATP-
538 Mg/Pi transporter APC3. Despite being annotated as an Mg/Pi transporter, mutants in this gene
539 were only found to have reduced shoot Ca concentration. Tissue specific characterisation of
540 this line showed silique and stem Ca concentrations were lower than wild type plants. This
541 suggests that, at least in these conditions, the effects of the mutation are limited. However,
542 effects at different external Ca or Mg concentrations might be different. A further line mutated
543 in a gene encoding an ABC transporter was found to have both lower Ca and Mg concentrations
544 in siliques compared to wild type plants. Identifying candidate genes controlling silique
545 nutrient traits is particularly important in *B. napus*, which is mostly grown for the harvest of
546 seeds which have a secondary use in animal feed. All *A. thaliana* experiments in this study
547 took place using a high-nutrient compost. This could have masked the phenotypes of mutations
548 in a number of candidate genes which may have been able to maintain normal Ca and Mg

549 concentrations due to sufficient soil concentrations. In addition, it is possible that the mutations
550 characterised here would show greater defects in shoot Ca and Mg concentrations when grown
551 in nutrient limiting conditions. As well as this, all plants in both *A. thaliana* experiments were
552 harvested at a single growth stage and other phenotypes might be seen at other growth stages.
553 Despite this, four candidate genes analysed here have proven to be potential targets for altering
554 Ca and Mg concentrations in *B. napus*. These are orthologues of the *Arabidopsis thaliana* genes
555 At2g13610, At5g07320, At5g08670 and At5g48650.
556
557 **5 Summary and potential applications**
558
559 In this study, we have identified a number of genetic loci associated with leaf Ca and Mg
560 concentration in *B. napus*. Within these loci, several novel candidate genes together with genes
561 previously shown to influence or respond to Ca and Mg concentrations in this and closely
562 related *Brassica* spp. were localised. Most well defined loci included regions on chromosomes
563 A2, A10, C2 and C9, close to the known Ca and Mg transporters *ACA8* and *MGT7*.
564 Experiments in *A. thaliana* T-DNA knockouts confirmed that a further four candidate genes
565 influence shoot Ca and Mg concentrations. This study used *B. napus* associative
566 transcriptomics followed by an *A. thaliana* T-DNA knockout workflow to identify and test
567 candidate genes quickly and efficiently. Due to similar phylogeny, genes characterised here in
568 *A. thaliana* are likely to have additive effects in *B. napus*. However, further study of candidate
569 genes in *B. napus* is required to confirm *A. thaliana* gene functions observed here and in
570 previous studies are conserved. Both *ACA8* and *MGT7* are good targets for this, especially
571 since *ACA8* has previously exhibited Ca-responsiveness in *B. rapa* (Graham *et al.*, 2014), and
572 since the effects of mutations in *A. thaliana MGT7* are so marked. Selection of *B. napus*
573 genotypes with different alleles of target genes may lead to improved ability to grow in the
574 presence of low soil Ca or Mg concentrations. The development of high Ca and Mg
575 accumulating lines in edible portions of *Brassica* spp. also has the potential to reduce nutrient
576 deficiencies in humans and livestock across the world.
577
578 **Conflict of interest statement**
579
580 The authors declare that they have no competing interests.
581
582 **Author contributions**
583
584 MB, IB, PW, TA and NG conceived the project and contributed to experimental design. TA
585 analysed associative transcriptomics data and performed and analysed *A. thaliana* mutant
586 experiments. LH and ZH prepared functional genotypes and performed associative
587 transcriptomics. TA and NG wrote the manuscript. All authors contributed to and have read
588 and approved the final version of the manuscript.
589

602
603 **Supplementary material**
604
605 Supplementary Table 1: Block design of preliminary *Arabidopsis thaliana* phenotyping
606 experiment.
607 Supplementary Table 2: Block design of *Arabidopsis thaliana* tissue partitioning phenotyping
608 experiment.
609 Supplementary Table 3: Shoot Ca and Mg concentration data from preliminary *Arabidopsis*
610 *thaliana* phenotyping experiment.
611 Supplementary Table 4: Shoot Ca and Mg concentration data from *Arabidopsis thaliana* tissue
612 partitioning phenotyping experiment.

613
614 **References**
615

616 Adams KL, Cronn R, Percifield R, Wendel JF. 2003. Genes duplicated by polyploidy show
617   unequal contributions to the transcriptome and organ-specific reciprocal silencing.
618   *Proceeding of the National Academy of Sciences of the United States of America* **100**:
619   4649-54.
620 Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, *et al*. 2003. Genome-Wide
621   Insertional Mutagenesis of *Arabidopsis thaliana*. *Science* **301**: 653-657.
622 Bancroft I, Morgan C, Fraser F, Higgins J, Wells R, Clissold L, *et al*. 2011. Dissecting the
623   genome of the polyploid crop oilseed rape by transcriptome sequencing. *Nature*
624   *biotechnology* **29**: 762-766.
625 Baxter I, Ouzzani M, Orcun S, Kennedy B, Jandhyla SS, Salt DE. 2007. Purdue Ionomics
626   Information Management System. An Integrated Functional Genomics Platform. *Plant*
627   *Physiology* **143**: 600-611.
628 Benjamini Y. and Hochberg Y. 1995. Controlling the false discovery rate: a practical and
629   powerful approach to mutliple testing. *Journal of the Royal Statistical Society: Series B*
630   **57**: 289–300.
631 Bonza MC, Morandini P, Luoni L, Geisler M, Palmgren MG and De Michelis MI. 2000. *At-*
632   ACA8 Encodes a Plasma Membrane-Localized Calcium-ATPase of Arabidopsis with a
633   Calmodulin-Binding Domain at the N Terminus. *Plant Physiology* **123**: 1495-1506.
634 Broadley MR, Bowen HC, Cotterill HL, Hammond JP, Meacham MC, Mead A *et al*. 2003.
635   Variation in the shoot calcium content of angiosperms. *Journal of Experimental Botany*
636   **54**: 1431-1446.
637 Broadley MR, Bowen HC, Cotterill HL, Hammond JP, Meacham MC, Mead A *et al*. 2004.
638   Phylogenetic variation in the shoot mineral concentration of angiosperms. *Journal of*
639   *Experimental Botany* **55**: 321-336.
640 Broadley MR, Hammond JP, King GJ, Astley D, Bowen HC, Meacham MC, *et al*. 2008. Shoot
641   calcium (Ca) and magnesium (Mg) concentrations differ between subtaxa, are highly
642   heritable, and associate with potentially pleiotropic loci in *Brassica oleracea*. *Plant*
643   *Physiology* **146**: 1707-1720.
644 Broadley MR and White PJ. 2010. Eats roots and leaves. Can edible horticultural crops address
645   dietary calcium, magnesium and potassium deficiencies? *The Proceedings of the*
646   *Nutrition Society* **69**: 601-612.
647 Bus A, Körber N, Snowdon R, Stich B. 2011. Patterns of molecular variation in a species-wide
648   germplasm set of *Brassica napus*. *Theoretical and Applied Genetics* **123**: 1413–1423.

649   Bus A, Körber N, Parkin IAP, Samans B, Snowdon RJ, Li J, *et al*. 2014. Species- and genome-
650   wide dissection of the shoot ionome in *Brassica napus* and its relationship to seedling
651   development. *Frontiers in Plant Science* **5**: 485. doi: 10.3389/fpls.2014.00485.
652   Cakmak I and Yazici AM. 2010. Magnesium: A Forgotten Element in Crop Production. *Better*
653   *Crops* **94**: 23-25.
654   Catharine RA. 2011. Dietary Reference Intakes for Calcium and Vitamin D. National
655   Academic Press. Washington DC, USA. Pages 35-74.
656   Chalhoub B, Denoeud F, Liu S, Parkin IAP, Tang H, Wang X, *et al.* 2014. Early allopolyploid
657   evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* **345**: 950-953.
658   Conn S and Gilliham M. 2010. Comparative physiology of elemental distributions in plants.
659   *Annals of Botany* **105**: 1081–1102.
660   De Angeli A, Monachello D, Ephritikhine G, Frachisse JM, Thomine S, Gambale F, *et al*. 2006.
661   The nitrate/proton antiporter AtCLCa mediates nitrate accumulation in plant vacuoles.
662   *Nature* **442**: 939–942.
663   Dunn OJ. 1961. Multiple comparison among means. *Journal of the American Statistical*
664   *Association* **56**: 52–64.
665   Ecke W, Clemens R, Honsdorf N, Becker HC. 2010. Extent and structure of linkage
666   disequilibrium in canola quality winter rapeseed (*Brassica napus* L.). *Theoretical and*
667   *Applied Genetics* **120**: 921-931.
668   Gebert M, Meschenmoser K, Svidová S, Weghuber J, Schweyen R, Eifler K, *et al*. 2009. A
669   Root-Expressed Magnesium Transporter of the *MRS2/MGT* Gene Family in *Arabidopsis*
670   *thaliana* Allows for Growth in Low-Mg^{2+} Environments. *The Plant Cell* **12**: 4018-4030.
671   Graham NS, Hammond JP, Lysenko A, Mayes S, Ó Lochlainn S, Blasco B, *et al*. 2014.
672   Genetical and Comparative Genomics of *Brassica* under Altered Ca Supply Identifies
673   *Arabidopsis* Ca-Transporter Orthologs. *The Plant Cell* **26**: 2818-2830.
674   Grusak MA, Broadley MR, White PJ. 2016. "Plant Macro- and Micronutrient Minerals," in:
675   *eLS*. John Wiley & Sons, Ltd: Chichester. doi: 10.1002/9780470015902.a0001306.pub2
676   Harper AL, Trick M, Higgins J, Fraser F, Clissold L, Wells R, *et al*. 2012. Associative
677   transcriptomics of traits in the polyploid crop species Brassica napus. *Nature*
678   *biotechnology* **30**: 798-802.
679   He Z, Wang L, Harper AL, Havlickova L, Pradhan AK, Parkin IAP *et al*. 2016. Extensive
680   homoeologous genome exchanges in allopolyploid crops revealed by mRNAseq-based
681   visualization. *Plant Biotechnology Journal* **15**: 594-604. doi:10.1111/pbi.12657.
682   He Z, Cheng F, Li Y, Wang X, Parkin IA, Chalhoub B, *et al*. 2015. Construction of Brassica
683   A and C genome-based ordered pan-transcriptomes for use in rapeseed genomic research.
684   *Data in brief* **4**: 357-362.
685   Higgins J, Magusin A, Trick M, Fraser F, Bancroft, I. 2012. Use of mRNA-seq to discriminate
686   contributions to the transcriptome from the constituent genomes of the polyploid crop
687   species *Brassica napus*. *BMC Genomics* **13**: 247.
688   Hoagland DR and Arnon DI. 1950. *The Water-Culture Method for Growing Plants without*
689   *Soil.* Berkeley: College of Agriculture, University of California.
690   Huala E, Dickerman A, Garcia-Hernandez M, Weems D, Reiser L, LaFond F *et al*. 2001. The
691   Arabidopsis Information Resource (TAIR): A comprehensive database and web-based
692   information retrieval, analysis, and visualization system for a model plant. *Nucleic Acids*
693   *Research* **29**: 102-105.
694   Iniguez-Luy FL and Federico ML. 2011. "The Genetics of *Brassica napus*," in *Genetics and*
695   *Genomics of the Brassicaceae*, *Plant Genetics and Genomics: Crops and Models,*
696   *Volume* 9, eds Schmidt, R. and Bancroft, I. (New York, NY: Springer New York), 291-
697   322.

Karley AJ and White PJ. 2009. Moving cationic minerals to edible tissues: Potassium, magnesium, calcium. *Current Opinion in Plant Biology* **12**: 291-298.

Kersey PJ, Allen JE, Armean I, Boddu S, Bolt BJ, Carvalho-Silva D, *et al.* 2016. Ensembl Genomes 2016: more genomes, more complexity. *Nucleic Acids Research* **44**: D574-D580. doi: 10.1093/nar/gkv1209.

Körber N, Wittkop B, Bus A, Friedt W, Snowdon RJ, Stich B. 2012. Seedling development in a *Brassica napus* diversity set and its relationship to agronomic performance. *Theoretical and Applied Genetics* **125**: 1275–87.

Körber N, Bus A, Li J, Higgins J, Bancroft I, Higgins EE, *et al.* 2015. Seedling development traits in Brassica napus examined by gene expression analysis and association mapping. *BMC Plant Biology* **15**:136.

Kumssa DB, Joy EJM, Ander EL, Watts MJ, Young SD, Walker S, *et al.* 2015a. Dietary calcium and zinc deficiency risks are decreasing but remain prevalent. *Scientific Reports* **5**: 10974.

Kumssa DB, Joy EJM, Ander EL, Watts MJ, Young SD, Rosanoff A, *et al.* 2015b. Global magnesium supply in the food chain. *Crop & Pasture Science* **66**: 1278-1289.

Koprivova A, Harper AL, Trick M, Bancroft I, Kopriva S. 2014. Dissection of the control of anion homeostasis by associative transcriptomics in *Brassica napus*. *Plant Physiology* **166**: 442-450.

Lee H, Suh SS, Park E, Cho E, Ahn JH, Kim SG, *et al.* 2000. The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis. *Genes & Development* **14**: 2366-2376.

Leinonen R, Sugawara H and Shumway M. 2011. The Sequence Read Archive. *Nucleic Acid Research* **39**: D19-21. doi: 10.1093/nar/gkq1019.

Lenz H, Dombinov V, Dreistein J, Reinhard MR, Gebert M, Knoop V. 2013. Magnesium Deficiency Phenotypes Upon Multiple Knockout of *Arabidopsis thaliana* MRS2 Clade B Genes Can be Ameliorated by Concomitantly Reduced Calcium Supply. *Plant and Cell Physiology* **54**: 1118-1131. doi: 10.1093/pcp/pct062.

Li L, Tutone AF, Drummond RSM, Gardner RC and Luan S. 2001. A Novel Family of Magnesium Transport Genes in Arabidopsis. *The Plant Cell* **13**: 2761-2775.

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, *et al.* 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics* **28**: 2397-2399.

Lu G, Harper AL, Trick M, Morgan C, Fraser F, O'Neill C, *et al.* 2014. Associative Transcriptomics Study Dissects the Genetic Architecture of Seed Glucosinolate Content in *Brassica napus*. *DNA Research* **21**: 613-625.

McAinsh MR and Pittman JK. 2009. Shaping the calcium signature. *New Phytologist* **181**: 275-294.

Michaels SD and Amasino RM. 1999. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *The Plant Cell* **11**: 949-956.

Parkin IA, Koh C, Tang H, Robinson SJ, Kagale S, Clarke WE, *et al.* 2014. Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid Brassica oleracea. *Genome Biology* **15**: R77.

Popescu AA, Harper AL, Trick M, Bancroft I, Huber KT. 2014. A Novel and Fast Approach for Population Structure Inference using Kernel-PCA and optimisation (PSIKO). *Genetics* **198**: 1421-31.

R Core Team. 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org.

Rios JJ, O´Lochlainn S, Devonshire J, Graham NS, Hammond JP, King GJ, *et al.* 2012. Distribution of calcium (Ca) and magnesium (Mg) in the leaves of *Brassica rapa* under varying exogenous Ca and Mg supply. *Annals of Botany* **109**: 1081-1089.

748 Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D. 2002. A High-Throughput
749   Arabidopsis Reverse Genetics System. *The Plant Cell* **14**: 2985-2994.
750 Sheldon CC, Rouse DT, Finnegan EJ, Peacock WJ and Dennis ES. 2000. The molecular basis
751   of vernalization: The central role of FLOWERING LOCUS C (FLC). *Proceeding of the*
752   *National Academy of Sciences of the United States of America* **28**: 3753-3758.
753 Slatkin, M. 2008. Linkage disequilibrium — understanding the evolutionary past and mapping
754   the medical future. *Nature Reviews. Genetics* **9**: 477–485. doi: 10.1038/nrg2361.
755 Sønderby IE, Burow M, Rowe HC, Kliebenstein DJ, Halkier BA. 2010. A complex interplay
756   of three R2R3 MYB transcription factors determines the profile of aliphatic
757   glucosinolates in *Arabidopsis*. *Plant Physiology* **153**: 348-363.
758 Swaminathan R. 2003. Magnesium Metabolism and its Disorders. *The Clinical Biochemist*
759   *Reviews* **24**: 47-66
760 Thomas CL, Alcock TD, Graham NS, Hayden R, Matterson S, Wilson L, *et al*. 2016. Root
761   morphology and seed and leaf ionomic traits in a Brassica napus L. diversity panel show
762   wide phenotypic variation and are characteristic of crop habit. *BMC Plant Biology* **16**:
763   214. doi: 10.1186/s12870-016-0902-5.
764 Trick M, Long Y, Meng J, Bancroft I. 2009. Single nucleotide polymorphism (SNP) discovery
765   in the polyploidy *Brassica napus* using Solexa transcriptome sequencing. *Plant*
766   *Biotechnology Journal* **7**: 334-346.
767 VSN International. 2014. GenStat for Windows 17th Edition. VSN International, Hemel
768   Hempstead, UK.
769 Waese J. and Provart NJ. 2017. "The Bio-Analytic Resource for Plant Biology," in *Plant*
770   *Genomics Databases: Methods and Protocols*, ed. ADJ van Dijk (New York, NY:
771   Springer New York), 119-48.
772 Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, *et al*. 2011. The genome of the mesopolyploid
773   crop species *Brassica rapa*. *Nature Genetics* **43**: 1035-1039.
774 White PJ. 2001. The pathways of calcium movement to the xylem. *Journal of Experimental*
775   *Botany* **52**: 891–899.
776 White PJ. 2015. "Calcium," in *A Handbook of Plant Nutrition, Second Edition*, eds. AV Barker
777   and DJ Pilbeam (Boca Raton, FL: CRC Press), 165-198.
778 White PJ and Broadley MR. 2003. Calcium in Plants. *Annals of Botany* **92**: 487-511.
779 White PJ and Broadley MR. 2005. Biofortifying crops with essential mineral elements. *Trends*
780   *in Plant Science* **10**: 586-593.
781 White PJ and Broadley MR. 2009. Biofortification of crops with seven mineral elements often
782   lacking in human diets – iron, zinc, copper, calcium, magnesium, selenium and iodine.
783   *New Phytologist* **182**: 49-84.
784 White PJ and Brown PH. 2010. Plant nutrition for sustainable development and global health.
785   *Annals of Botany* **105**: 1073-1080.
786 White PJ, Bowen HC, Farley E, Shaw EK, Thompson JA, Wright G, *et al*. 2015. Phylogenetic
787   effects on shoot magnesium concentration. Crop & Pasture Science 66: 1241-1248. doi:
788   10.1071/CP14228.
789 Yardley AW ed. 2009. *Dietary Magnesium: New Research*. New York: Nova Science
790   Publishers, Inc.
791 Zhang Z, Ersoz E, Lai C, Todhunter RJ, Tiwari HK, Gore MA, *et al*. 2010. Mixed linear model
792   approach adapted for genome-wide association studies. *Nature Genetics* **42**: 355-360.
793 Zhao Q, Leung S, Corbett AH, and Meier I. 2006. Identification and Characterization of the
794   Arabidopsis Orthologs of Nuclear Transport Factor 2, the Nuclear Import Factor of Ran.
795   *Plant Physiology* **140**: 869–878.
796

797 Table 1. Summary of *Arabidopsis thaliana* T-DNA insertion lines acquired for characterisation
798 including primers used for genotyping. NASC stock code shown for reordering.
799

Line Name	SALK/SAIL code	NASC stock code	Forward primer	Reverse primer
At2g05120.1	SALK_119762	N619762	TTCTGGAGAAACAAGGTCCAA	ATGGCAGCAAGTTTTTCACC
At2g13610.1	SALK_074250	N574250	CGATTTGCCGAAAAGAAAAA	GTTTCCTCCACCGTAAGCAA
At2g13610.2	SALK_074250C	N681303	CGATTTGCCGAAAAGAAAAA	GTTTCCTCCACCGTAAGCAA
At2g45660.1	SALK_138131	N638131	GGTTCTTCCTTTCGCAGAGA	CCACAAAAGGCCAATCAAAT
At5g03960.1	SALK_138382	N638382	TGGTTGAGGAAGCAAGAAGG	TGTGCTCTGCCTCCTTTGTA
At5g06530.1	SALK_024391	N524391	TTCCCCAAAGGTATCGATTCTA	TCGAACAACTGGGATTGACA
At5g06530.2	SALK_076250	N576250	TTCCCCAAAGGTATCGATTCTA	CGGGCATTTGATAGCACTTT
At5g07320.1	SALK_037517	N537517	CGCTGCATATGAAACGCTAA	TCAATGATCGCAACAAAACAA
At5g07320.2	SALK_037517C	N683966	CGCTGCATATGAAACGCTAA	CCATAAAAATATATGTCCCAATTTCA
At5g08670.1	SALK_083107	N583107	CGATGTTCCCAACATTTGAA	AACAGAGACCGGCGAGACTA
At5g10520.1	SALK_019299	N519299	TATTTCATGCACGGCATTGT	GGGTTGGAAATGTGGAAGAA
At5g10520.2	SALK_053754	N553754	CCGTTTCGTCTTCTCACCAT	ACATGGTGAGGCCAGTTCTC
At5g14040.1	SALK_105845	N605845	CCCTTACTTTTCGGAGCATTC	TTGCACTTGACGAGATCGAG
At5g48650.1	SALK_027468	N527468	GCGGTAGCTGAGGGTACATC	CCACCATCAAGCCAAAGACT
At5g48650.2	SAIL_64_G12	N803057	GCCCAATAGGCAAACAAATG	AAGTCTGGGACCAACAATGG

800
801

802 Table 2. Summary of candidate genes selected from associative transcriptomics outputs.
803 Putative functions obtained from The Arabidopsis Information Resource (TAIR; Huala *et al.*
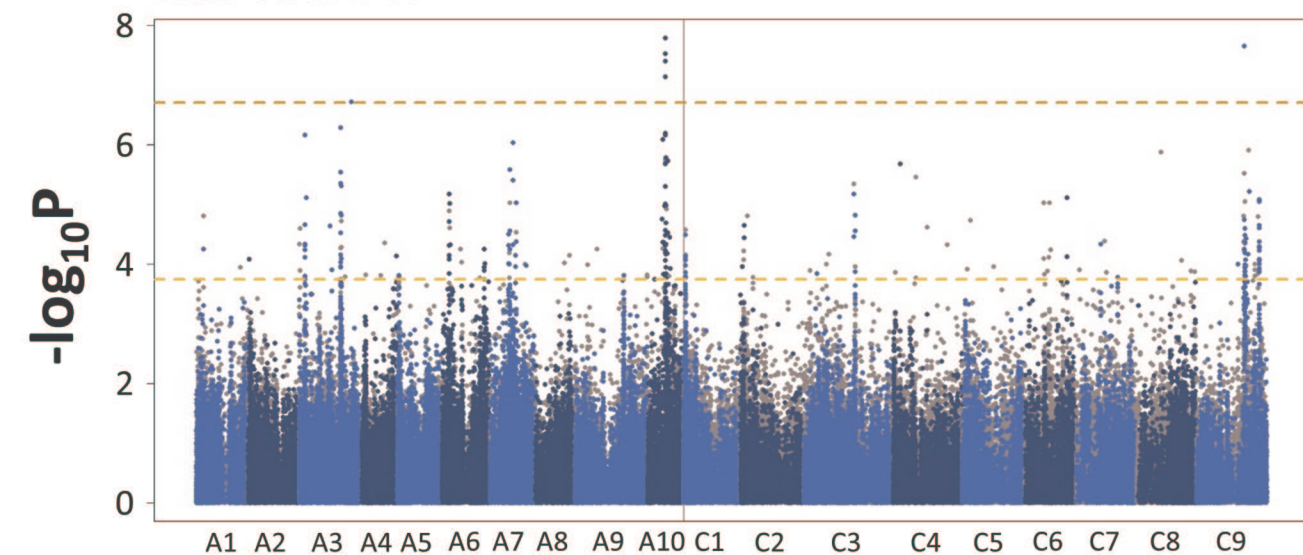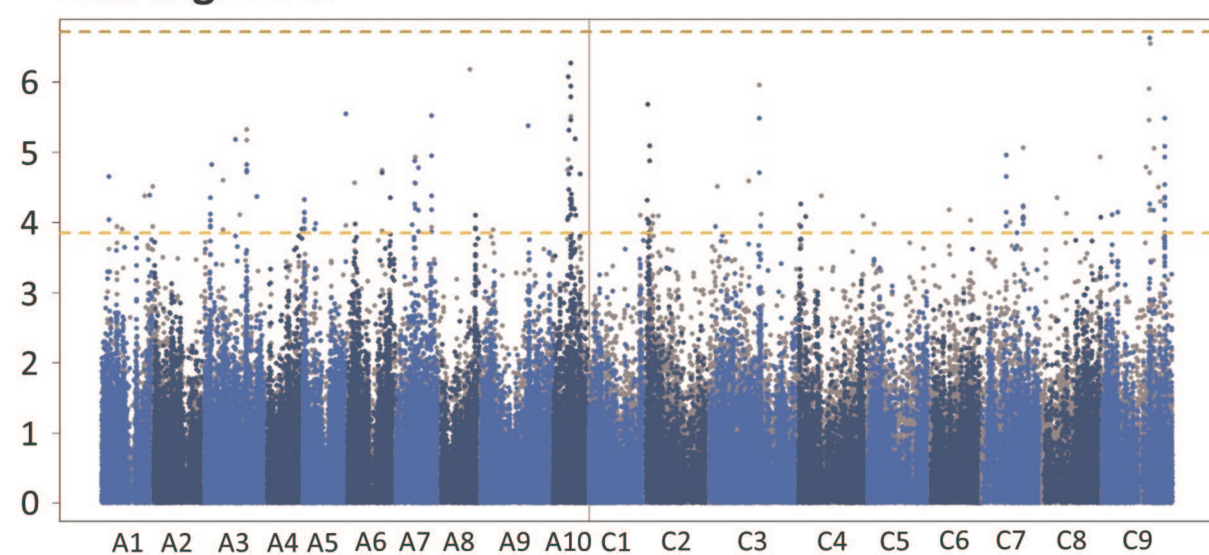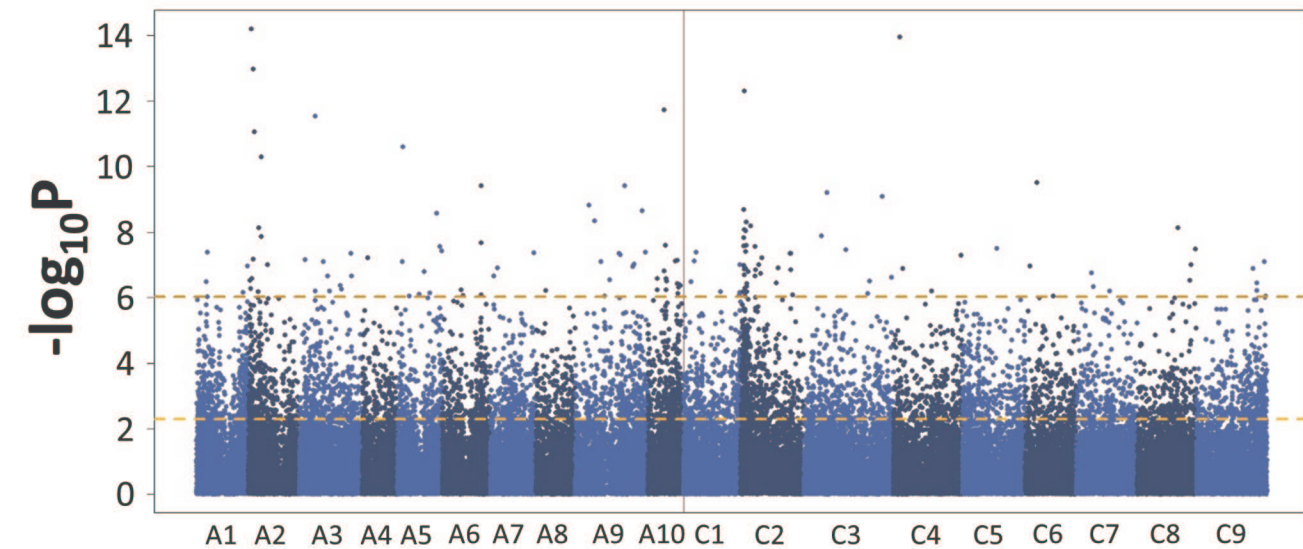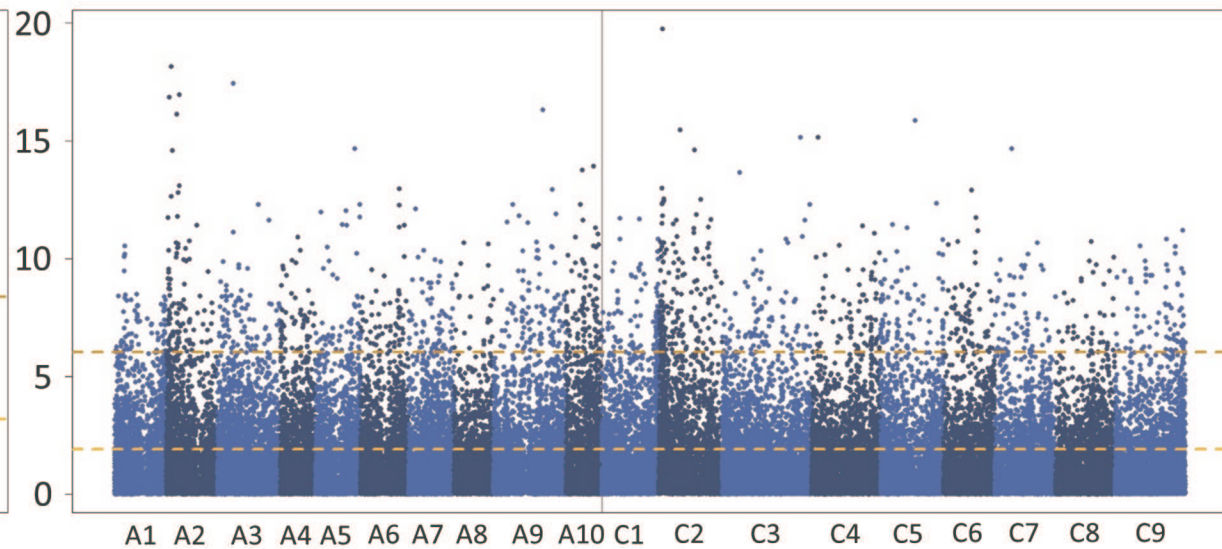804 2001).

805

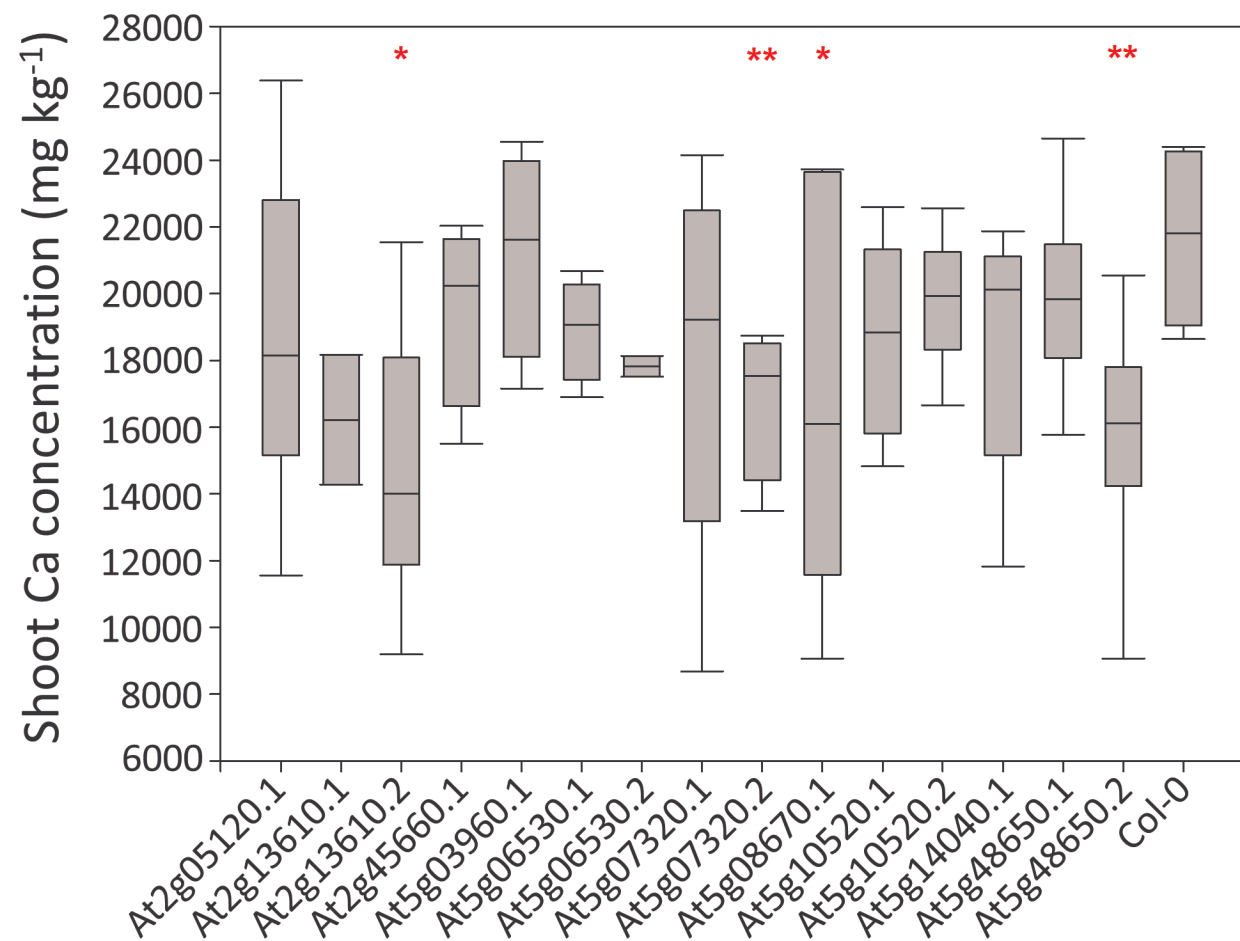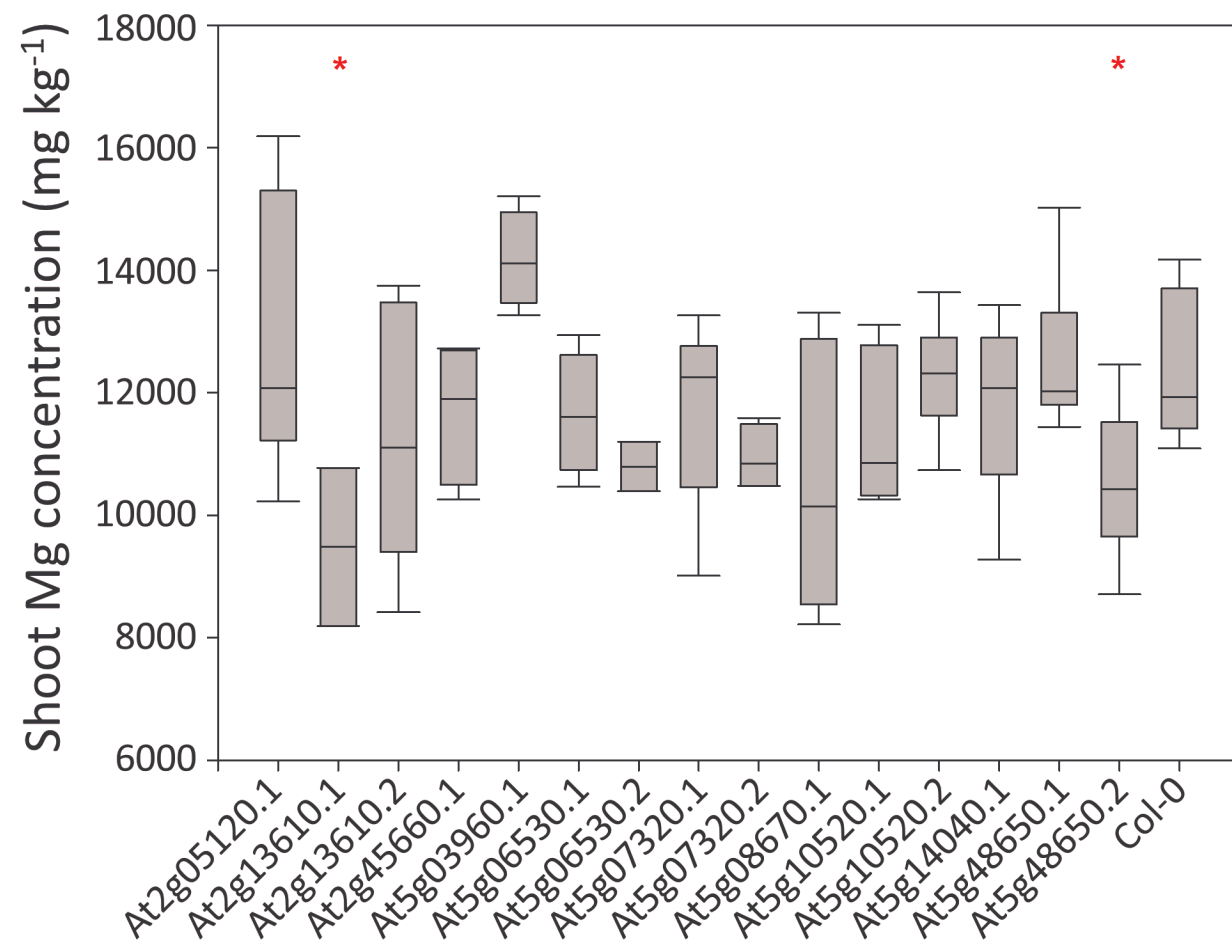| Candidate gene	Chromosome	*A. thaliana* orthologue	Putative function
Cab036107.1	A2	At5g09690	Mg transporter - MGT7
Cab039480.1	A2	At5g03960	IQ-domain - IQD12
Cab002472.4	A3	At5g10140	Flowering Locus - FLC
Cab001235.1	A3	At2g05120	Nucleoporin - NUP133
Cab001274.1	A3	At2g13610	ABC transporter - ABCG5
Cab025356.1	A5	At2g45660	Suppressor of overexpression of CO - SOC1
Cab007043.1	A6	At5g48650	Nuclear transport factor - NTF2
Cab017470.1	A10	At5g57110	Ca transporting ATPase - ACA8
BnaC02g00490D	C2	At5g10140	Flowering Locus - FLC
Bo2g007260.1	C2	At5g06530	ABC transporter - ABCG22
Bo2g008580.1	C2	At5g07320	ATP-Mg/Pi transporter - APC3
Bo2g009200.1	C2	At5g08670	Mitochondrial ATP synthase beta-subunit
Bo2g009480.1	C2	At5g09710	Mg transporter - MGT7
Bo2g009490.1	C2	At5g09690	Mg transporter - MGT7
Bo2g009910.1	C2	At5g10520	ROP Binding Protein Kinase - RBK1
Bo2g011650.1	C2	At5g14040	Phosphate transporter - PHT3;1
Bo4g024850.1	C4	At2g45660	Suppressor of overexpression of CO - SOC1

806

807 Fig 1. $-\log_{10}P$ values of SNPs and GEMs associated with leaf Ca concentration (panels A and
808 B respectively) and leaf Mg concentration (panels C and D respectively) in order of markers
809 within the *B. napus* pan-transcriptome. Upper, gold, dashed line represents Bonferroni
810 corrected significance threshold; lower, yellow, dashed line represents FDR corrected
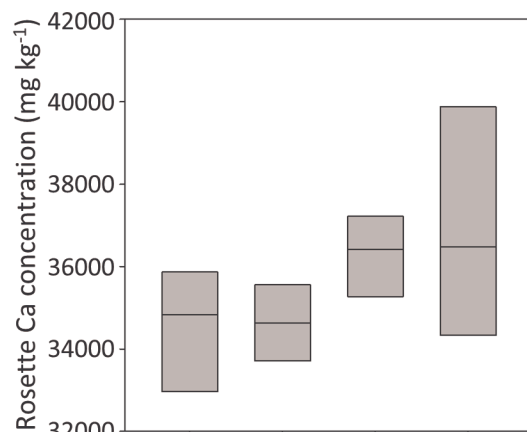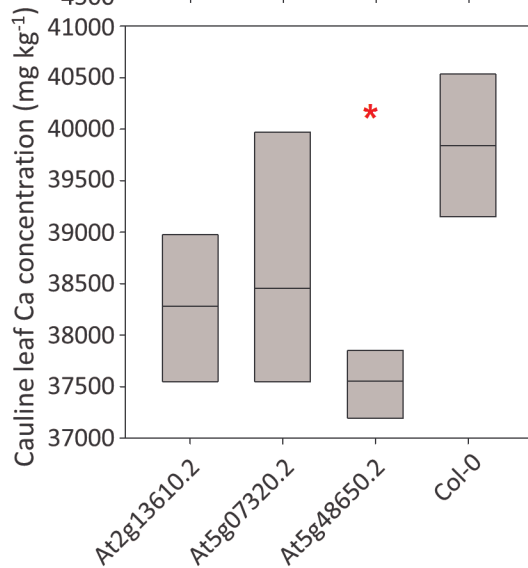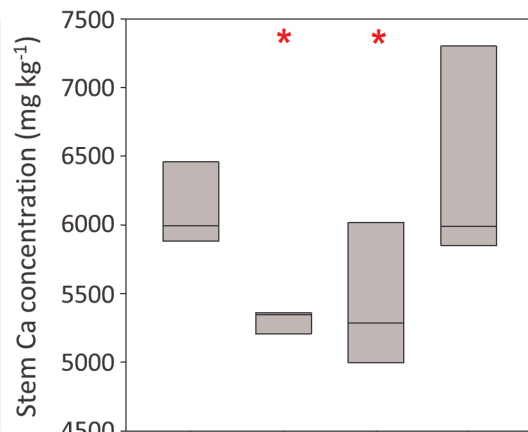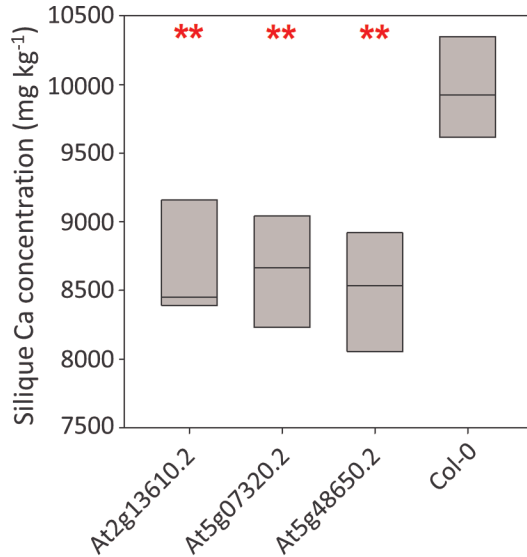811 significance threshold (*P=0.05*).

812

813 Fig 2. Shoot Ca (panel A) and Mg (panel B) concentrations across 15 mutant *A. thaliana* lines
814 and wild type (Col-0) plants. Boxes represent the mid two quartiles with the median drawn;
815 whiskers are the 95% confidence limits. Single and double stars above boxes represent
816 significance at *P<0.05* and *P<0.01* respectively compared to wild type (Col-0) plants.
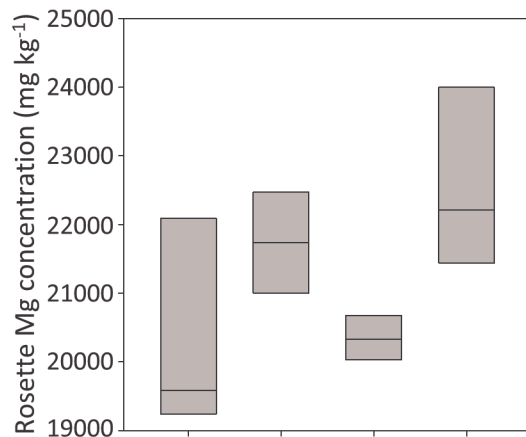
817

818 Fig 3. Rosette leaf (panel A), silique (panel B), stem (panel C) and cauline leaf (panel D) Ca
819 concentrations across three mutant *A. thaliana* lines and wild type (Col-0) plants. Boxes
820 represent full range of values with the median drawn. Single and double stars above boxes
821 represent significance at *P<0.05* and *P<0.01* respectively compared to wild type (Col-0)
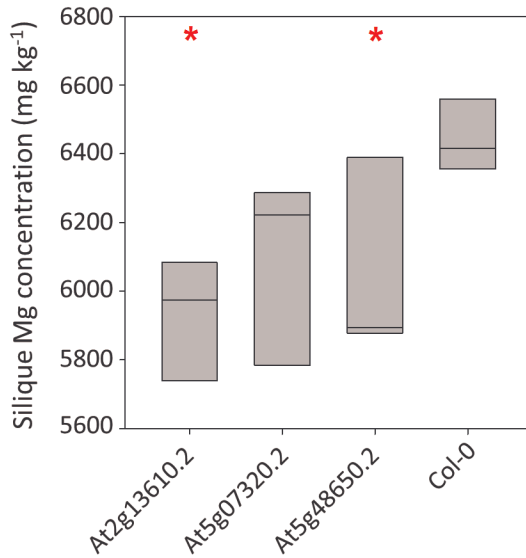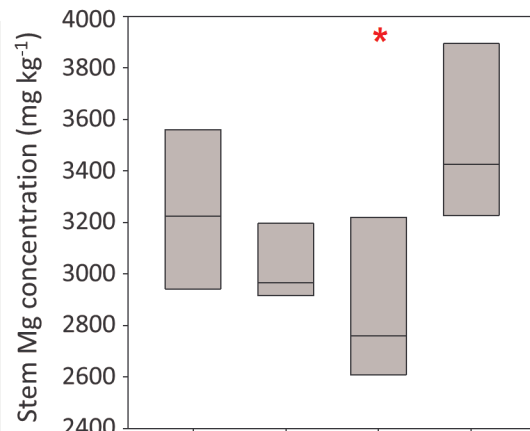822 plants.

823

824 Fig 4. Rosette leaf (panel A), silique (panel B), stem (panel C) and cauline leaf (panel D) Mg
825 concentrations across three mutant *A. thaliana* lines and wild type (Col-0) plants. Boxes
826 represent full range of values with the median drawn. Single and double stars above boxes
827 represent significance at *P<0.05* and *P<0.01* respectively compared to wild type (Col-0)
828 plants.

A Leaf Ca SNPs:

C Leaf Mg SNPs:

B Leaf Ca GEMs:

D Leaf Mg GEMs:

Markers arranged in *B. napus* chromosomal order

A

Shoot Ca concentration (mg kg^{-1})

*   *   **   *   **

At2g05120.1, At2g13610.1, At2g13610.2, At2g45660.1, At5g03960.1, At5g06530.1, At5g06530.2, At5g07320.1, At5g07320.2, At5g08670.1, At5g10520.1, At5g10520.2, At5g14040.1, At5g48650.1, At5g48650.2, Col-0

B

Shoot Mg concentration (mg kg^{-1})

*   *

At2g05120.1, At2g13610.1, At2g13610.2, At2g45660.1, At5g03960.1, At5g06530.1, At5g06530.2, At5g07320.1, At5g07320.2, At5g08670.1, At5g10520.1, At5g10520.2, At5g14040.1, At5g48650.1, At5g48650.2, Col-0