

Elemental Accumulation in Kernels of the Maize Nested Association Mapping Panel Reveals Signals of Gene by Environment Interactions

Greg Ziegler¹, Philip J. Kear², Di Wu², Catherine Ziyomo³, Alexander E. Lipka⁴, Michael Gore², Owen Hoekenga⁴, Ivan Baxter^{1,3*}

¹USDA-ARS, Donald Danforth Plant Science Center, St Louis, MO 63132,

²Plant Breeding and Genetics Section, School of Integrated Plant Science, Cornell University, Ithaca, NY 14853,

³ Donald Danforth Plant Science Center, St Louis, MO 63132

⁴Department of Crop Sciences, University of Illinois, IL 61801

⁵Cayuga Genetics Consulting Group LLC, Ithaca NY 14850

*Corresponding author

E-mail: ivan.baxter@ars.usda.gov

Elemental accumulation in seeds is the product of a combination of environment and a wide variety of genetically controlled physiological processes. We measured the kernel elemental composition of the Nested Association Mapping (NAM) of maize (*Zea mays* L.) grown in 4 different environments. Analysis of variance revealed strong effects of genotype, environment and genotype by environment interactions. Using Joint-linkage mapping on a set of 7000 markers we identified 354 quantitative trait loci (QTL) across 20 elements, four environments and a combination of the environments. Leveraging 20 M SNPs derived from genome resequencing on the parents of the population, genome-wide association mapping studies (GWAS) detected 8573 loci. While most of the GWAS SNPs were located near genes not previously implicated in elemental regulation, several SNPs were located next to orthologs of well-characterized elemental regulation genes.

Keywords

Zea mays, joint-linkage mapping, genome-wide association study, ionomics

Introduction

Mineral elements are essential for all organisms to function, so over evolutionary time, all plants have developed and refined of processes to control their uptake, distribution and storage. To optimize fitness, organisms need to maximize uptake of limiting essential nutrients while reducing the uptake or sequestering chemically similar undesirable elements. This process is complicated by the wide variability of elemental availability, a function of parameters such as concentration, soil moisture, pH, cation-exchange-capacity, organic matter, and fertilizer applications [1]. The processes that respond to this variation in the environment to enable elemental uptake are poorly understood at the physiological and genetic level.

From its center of origin in Mexico, maize (*Zea mays* L.) has successfully adapted to many agro-ecological and environmental conditions due to its extensive genetic diversity [2]. Maize is now one of the most prevalent cereal crops in the world, with global production surpassing 1 billion tonnes in 2013 [3]. With a rich history as a genetic model, the maize community has created powerful genetic populations to use the genetic diversity to dissect the genetic control of many different traits. One such population is the maize nested association panel, derived from crossing 25 diverse parents to a single hub genotype and used to map loci for diverse traits such as flowering time, carotenoid concentration, and heat tolerance [4,5].

Ionomics, high throughput elemental profiling, has been a useful approach to understand elemental accumulation in both wild and crop plants. In addition to its use as a food source, the maize kernel is a defined step in the plants' development and provides a highly heritable gauge of the genetic and environmental influences on elemental acquisition and homeostasis [6]. In this study, we have analyzed the elemental composition of kernels of the Maize NAM population grown in four diverse environments to identify the genetic loci underlying the ionome.

Table 1. Site soil type and chemistry of NAM population planting locations in 2006

Loc	Climate	pH	Soil Type	Substrate	Agronomy	Coordinates**
FL	Subtropical	6.7	Sand-clay loam	Limestone dolomite	High OM, fertilized	25°30'14.87" N 80°30'18.77" W
PR	Trop	7.5	San Anton clay loam	Alluvial deposit	Wet / fertile High CEC	18° 0'11.17" N 66°31'8.43" W
NC	Temp	4.5	Norfolk loamy sand	Loamy marine dep	Fertilized	35°40'10.05" N 78°30'16.67" W
NY	Temp	6.7	Lima silt loam	Limestone calcareous	Fertilized	42°43'31.18" N 76°39'5.10" W

** Coordinates provided by © Google 2016 from Google Earth
Soil data provided by Natural Resource Conservation Service [7]

Methods

The Maize Diversity project self-pollinated ears of the Maize NAM in 2006 in North Carolina, New York, Puerto Rico and Florida as described in Buckler 2009 [4], and Brown [8]. Individual

kernels from these growouts were analyzed on the ionomics pipeline as described in Ziegler et al.

Initially we analyzed the grow-outs in separate experimental batches containing only a single grow-out, but for later runs we mixed samples from separate locations, allowing us to remove batch effects from the analysis of environmental variation. All locations analyzed had some samples mixed in a batch with samples from other locations. The ratio between samples in these mixed runs was used to determine the true elemental ratio across locations. The elemental analysis encompassed over 100 runs of the ICP-MS and 50,000 total samples. Of the 18,000 packets analyzed, over 16,000 had two or more kernels analyzed.

Outlier removal

Analytical outliers were removed from single-seed measurements using a method described in Davies and Gather [9]. Values were considered an outlier and removed from further analysis if the median absolute deviation (MAD) calculated based on the location and population each seed belongs to was greater than 10.

After outlier removal, an aggregate value for each analyte was calculated by taking the median value of the single-seed measurements for each line.

Transformation

To meet the normality assumption required for further statistical analyses, the Box-Cox power transformation [10] was used to determine an appropriate transformation for each phenotype. Only the boron, aluminum, and cobalt phenotypes required a transformation.

Calculation of BLUPs

Best linear unbiased predictors (BLUPs) of each ionomics compound for each line were predicted from a mixed model analysis across environments per Holland et al. [11] in ASReml version 3.0:

$$Y_{ijklmnop} = \mu + \text{family}_i + \text{RIL}(\text{family})_{ij} + \text{rep}_k + \text{pblock}(\text{rep})_{kl} + \text{block}(\text{rep}*\text{pblock})_{klm} + \text{wrrun}_n + \text{range}_o + \text{row}_p + \varepsilon_{ijklmnop} \quad (1)$$

where $Y_{ijklmnop}$ represents the phenotypic value of the p^{th} row of the o^{th} range of the n^{th} weighing robot run of the m^{th} block of the i^{th} pblock of the k^{th} rep of the j^{th} RIL of the i^{th} family, μ is the grand mean, family_i is the effect of the i^{th} family, $\text{RIL}(\text{family})_{ij}$ is the effect of the j^{th} RIL within the i^{th} family, rep_k is the effect of the k^{th} rep, $\text{pblock}(\text{rep})_{kl}$ is the effect of the l^{th} pblock within the k^{th} rep, $\text{block}(\text{rep}*\text{pblock})_{klm}$ is the effect of the m^{th} block within the l^{th} pblock and the k^{th} rep, wrrun_n is the effect of the n^{th} weighing robot run, range_o is the effect of the o^{th} range, row_p is the effect of the p^{th} row, and $\varepsilon_{ijklmnop}$ represents a random error term. A first-order autoregressive (AR1 \times AR1) correlation structure was used to account for spatial variation among the rows and ranges. All terms except for family_i and $\text{RIL}(\text{family})_{ij}$ are random.

Likelihood ratio tests (testing H_0 : model term is not significant) were conducted to remove all terms from the model that were not significant at $\alpha = 0.05$ [12]. The final model was then used to estimate BLUPs for each line and to estimate variance components.

Joint-Linkage Analysis

Joint-linkage analysis was run using the StepwiseJointLinkagePlugin in TASSEL version 3.0 [13] with ~7,000 SNPs obtained by genotyping-by-sequencing from Elshire et al. 2011 [14]. To determine an empirical entry/exit p-value thresholds that control the type I error rate at $\alpha = 0.05$, 1000 permutation tests were performed, in which the BLUP phenotype data were randomly shuffled within each NAM family before Joint Linkage analysis [15].

Genome-wide Association

Stepwise forward regression genome-wide association was performed using the NamGwasPlugin in TASSEL version 4.0 [13] and is based upon previous NAM-GWAS experiments [16–18]. Briefly, genome-wide association was on a chromosome-by-chromosome basis. To account for variance explained by QTL on other chromosomes, the phenotypes used were the residuals from each chromosome calculated from the joint-linkage model fit with all significant joint-linkage QTL except those on the given chromosome. Association analysis for each trait was performed 100 times by randomly sampling, without replacement, 80% of the lines from each population.

The input SNP dataset was 28.9 million SNPs obtained from Maize Hapmap1 [19], Maize Hapmap2 [20] and an additional ~800,000 putative copy-number variants (CNVs) from analysis of read depth counts in Hapmap2 [16,20]. These ~30 million markers were projected onto the entire 5,000 line NAM population using low-density markers obtained through genotyping-by-sequencing [14]. A cutoff p-value of $<1e-6$ was used for inclusion in the final model. SNP associations were considered significant if selected in more than 5 of the 100 models (Resample model inclusion probability (RMIP) ≥ 0.05 , [21]).

Data availability: All ionomic data, JL and GWAS results are available on the www.ionomicshub.org Data exchange. (Uploaded, 3/31/2016)

Results & Discussion

Elemental profiling of the nested association mapping panel kernels across four locations

We conducted an ionomics analysis of the publicly available nested association-mapping (NAM) panel across four locations chosen for their widely differing environments (Table 1). The NAM panel combines high levels of genetic diversity while providing sufficient power to detect QTL. NAM is composed of 25 F6 generation recombinant inbred line (RIL) families, each with ~200 members, that all share B73 as a common parent [22,23]. A total of >50,000 kernels were

analyzed for aluminum (Al), arsenic (As), boron (B), calcium (Ca), cadmium (Cd), cobalt (Co), copper (Cu), iron (Fe), potassium (K), magnesium (Mg), manganese (Mn), molybdenum (Mo), sodium (Na), nickel (Ni), phosphorus (P), rubidium (Rb), sulfur (S), selenium (Se), strontium (Sr) and zinc (Zn).

Element	Difference (%)
Magnesium	2
Potassium	3
Phosphorus	6
Sulfur	8
Boron	10
Arsenic	10
Manganese	12
Iron	18
Aluminum	21
Selenium	30
Zinc	38
Calcium	38
Copper	56
Molybdenum	57
Sodium	77
Cadmium	220
Nickel	255
Rubidium	263
Strontium	330

Table 2. Percent difference of the median concentrations for each element between the highest and lowest grow-out.

With a naive model including only location, and comparing >3500 packets from each location, all elements showed a significant effect ($p < 0.0001$) of location. The difference between the median concentrations across environments varied from Mg (2%) to Sr (330%) (excluding Co which had a large difference between environments which were likely due to analytical issues owing to its extremely low concentration in seeds)(Table 2 and Figure 1). Soil properties such as underlying substrate and pH, as well as agronomic practices, are predicted to have strong effects on elemental accumulation. However, even though the NC soil is 2.2 pH units lower than all the other locations it didn't have the highest levels of Fe or other acid soluble elements. This suggests that the simple soil parameters are not the major drivers of differences in elemental accumulation data collected between locations.

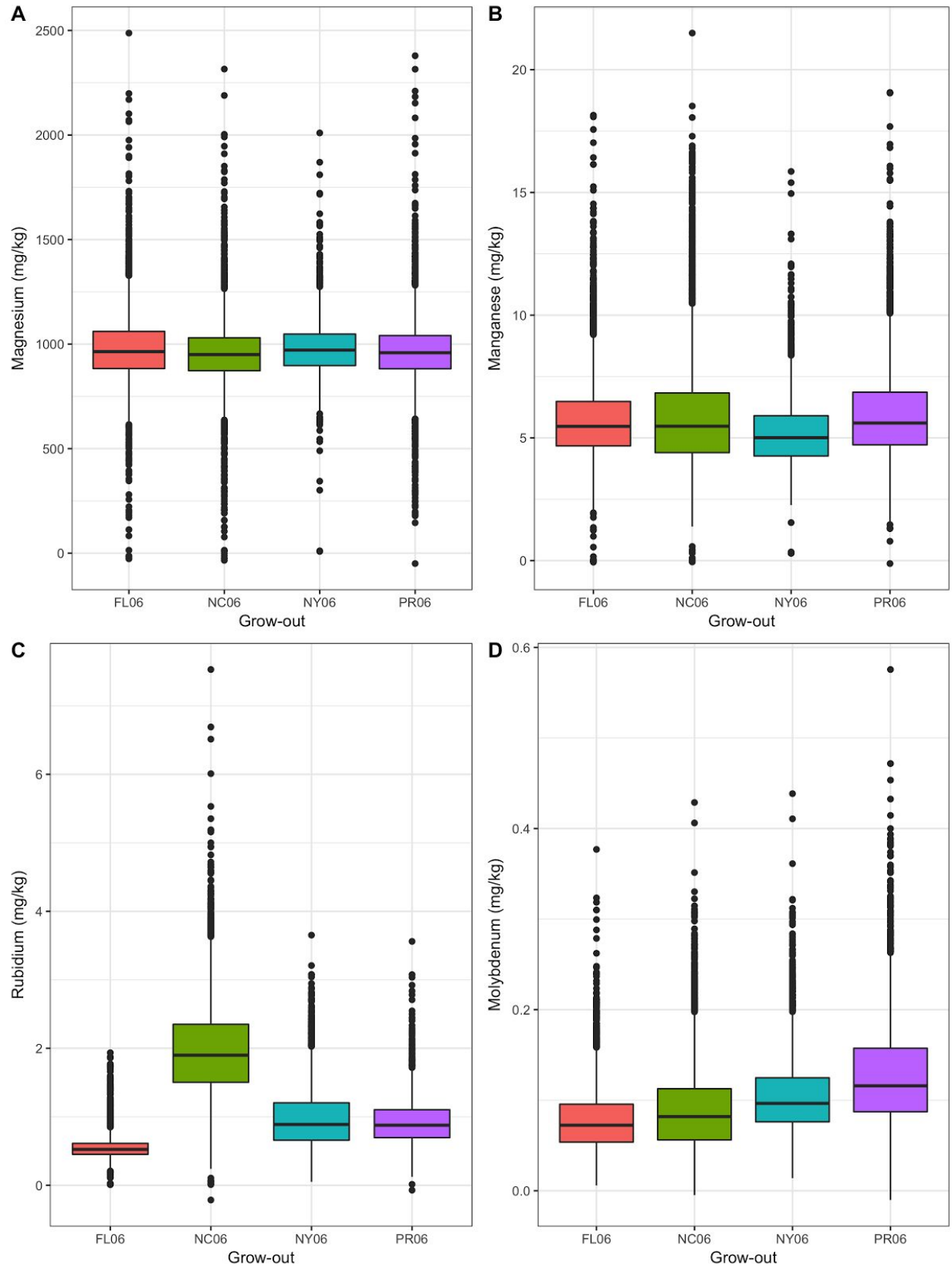


Figure 1. Concentrations of magnesium, manganese, rubidium, and molybdenum in NAM

RIL kernels, a subset of twenty elements analyzed, at four locations. Concentrations (ppm) magnesium, manganese, rubidium, and molybdenum in Florida (FL06; red), North Carolina (NC06; green), New York (NY06; blue) and Puerto Rico (PR06; purple) in 2006 from a sample size of approximately three maize kernels for each of ~3500 lines in each location analyzed using ICP-MS. The median (solid black horizontal line), 25%-75% quartile (colored boxes for location), higher/lower extreme concentration (solid black vertical line) and outliers (individual points).

Heritability

In order to partition the variance and estimate genotype values, we calculated BLUPs using the spatial checks in the field design and analytical checks. The nested design of the NAM population and the replicated field design across the four environments enabled a detailed statistical analysis of the sources of variation underlying the elemental traits. We performed linear mixed modeling of each element for each RIL across environments in ASReml. Significant terms were nested within a model for single and multiple environments. Models were fit separately for RIL families across environments, and each trait across environments. The final model was used to estimate best linear unbiased predictors (BLUPs) for each line and to estimate the components of phenotypic variance. The inclusion of the RIL family term, enables a robust calculation of $V_{G \times E}$ despite the single replicate plots of most lines in each location.

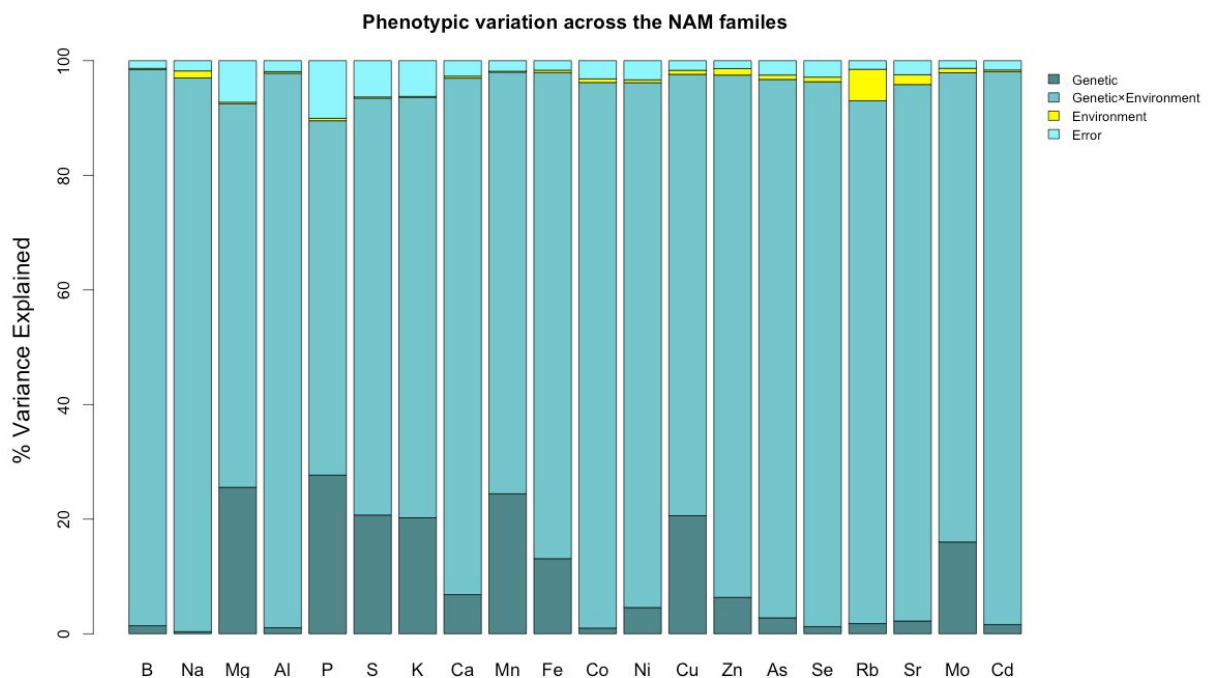


Figure 2: Partitioning of variance using the BLUP model.

Hung *et al* [24] introduced a combined model for the analysis of H^2 and variances in the NAM population using harmonic means across multiple locations [24]. When applied to the elemental dataset, four elements had a high H^2 across NAM families (> 0.60 ; Cu, Fe, Mn and Mo), 6 elements were moderate ($\geq 0.30 - 0.59$; Al, Ca, K, Mg and S), and 10 were low (≤ 0.29 ; As, B, Cd, Co, Na, Ni, P, Rb, Se and Sr) (Figure 2). These results suggest that the elemental profile is highly heritable within environments but strongly influenced by the environment in a genotype dependent manner. The large amount of genetic by environment variation we observed is in sharp contrast to the amount found by Pfeiffer et al and X using similar methods to study height and flowering time in the same populations. However a more appropriate comparison would be to carry out the same analysis on the exact same plots for the different traits, allowing a side by side comparison.

Joint Linkage Results

Stepwise joint-linkage regression was used for QTL detection across all 20 elements in the NAM population. QTL for each element were detected by using two methods: analyzing location data both individually and merged across locations using a best linear unbiased predictor (BLUP) model. In the four individual growouts, a total of 219 QTL were detected across 19 elements, with the exception of Se. Comparison of the individual location experiments revealed overlapping loci across multiple locations for a given element, clearly demonstrating that several QTL are derived from the same loci (Table 3).

El	All Locs QTL	FL	NC	NY	PR	Total Loc Specific QTL	multiloc QTL
Aluminum	1	1	0	0	1	2	0
Arsenic	2	0	1	2	0	3	0
Boron	3	0	0	1	0	1	0
Calcium	11	0	0	4	3	7	1
Cadmium	1	3	1	1	0	5	1
Copper	8	2	10	2	7	21	3
Iron	16	7	10	5	5	27	5
Potassium	5	0	1	2	0	3	1
Magnesium	12	4	5	6	2	17	3
Manganese	17	11	10	9	6	36	8
Molybdenum	12	5	7	5	7	24	5
Sodium	0	0	2	0	1	3	1
Nickel	6	0	4	3	4	11	2
Phosphorus	11	2	5	4	3	14	2
Rubidium	4	2	3	2	1	8	1
Sulfur	0	0	10	2	2	14	1
Selenium	1	0	0	0	0	0	0
Strontium	3	4	0	0	2	6	0
Zinc	9	2	7	2	1	12	1
Total	122	43	76	50	45	214	35

Table 3: Summary of results for a Joint Linkage Analysis on four NAM grow-outs and a combined BLUP analysis. All locs QTL: Number of QTL detected from the all locations BLUP analysis; Total Loc specific QTL: Number of QTL detected for each element in the four locations; multi-loc QTL: Number of QTL whose 95% confidence interval overlaps between two or more locations.

GWAS Results

The Joint-Linkage QTLs were used to account for loci on other chromosome as each chromosome was scanned for association with the 28M SNP markers that had been imputed on the full population. Using a cutoff of 5 iterations (out of 100) in the resampling approach, we identified a total of 8573 significant GWAS associations: 2923 in the all locations analysis and 5650 in the four locations specific datasets (Table 4). Counting overlap only as the identical SNP returned for the same element, we identified 29 loci that were found in multiple locations. Counting SNPs in LD with each other would result in many more SNPs being found in multiple locations.

El	All Locs QTL	FL	NC	NY	PR	Total Loc Specific QTL	multiloc QTL
Aluminum	176	79	0	0	88	167	0
Arsenic	182	0	274	54	0	328	0
Boron	108	0	0	31	0	31	0
Calcium	105	0	0	53	55	108	0
Cadmium	630	72	115	45	0	232	1
Copper	165	143	102	89	97	431	4
Iron	171	130	112	61	83	386	4
Potassium	130	0	253	47	0	300	1
Magnesium	153	89	84	78	84	335	0
Manganese	168	115	111	94	110	430	8
Molybdenum	154	124	102	123	140	489	6
Sodium	0	0	320	0	96	416	1
Nickel	99	0	100	78	144	322	2
Phosphorus	123	84	91	73	68	316	1
Rubidium	135	118	91	71	151	431	0
Sulfur	0	0	91	70	91	252	0
Selenium	162	0	0	0	0	0	0
Strontium	113	91	0	0	95	186	1
Zinc	149	52	99	221	118	490	0
Total	2923	1097	1945	1188	1420	5650	29

Table 4: Summary of results from GWAS analysis of four NAM grow-outs and a combined location BLUP analysis. All locs QTL: Number of QTL detected from the all locations BLUP analysis; Total Loc specific QTL: Total number of QTL detected for each element in the four locations; multi-loc QTL: Number of instance where the exact same QTL was returned in two or more locations.

High Confidence Candidate Genes

Visual inspection of the genes proximal to strong GWAS peaks revealed several genes whose orthologs are known to have roles in elemental accumulation (Table 5). These candidates suggest that the ionomics approach can be used with the NAM to identify maize genes important for elemental accumulation in the kernel.

Element	Chr	SNP bp	RMIP Score	Candidate gene	annotation	Orthologs
Mn	1	161430001		GRMZM2G028036	NRAMP	AT1G47240.1(ATNRAMP2,NRAMP2): NRAMP metal ion transporter 2
Mn	3	181680694		GRMZM2G014454	MTP	AT2G39450.1(ATMTP11,MTP11): Cation efflux family protein
Mo	1	244936001		GRMZM2G083156	Mot1	At2g25680 MOLYBDATE TRANSPORTER 1
Mo	5	174665257		GRMZM2G135816	CNX2	AT2G31955.1 (CNX2) cofactor of nitrate reductase and xanthine dehydrogenase 2
P	9	20313894		GRMZM2G444801	sulfate transporter	AT3G15990.1(SULTR3;4): sulfate transporter 3;4
Rb	2	61858998		GRMZM2G093826 & GRMZM2G395267	K transporter	AT4G13420.1(ATHAK5,HAK5): high affinity K+ transporter 5

Table 5. High Confidence genes identified under GWAS peaks

Manganese

Manganese is an essential plant element that is a cofactor in many crucial metalloenzymes, such as those in photosystem II [25]. Various methods of manganese transport and homeostasis in plants have been characterized [26,27]. Exploring the GWAS results for a model that combined the data from the four grow-out locations reveals a strong association (RMIP=0.98) for a SNP on chromosome 1 (Figure 3). In three of the four grow-outs, this SNP is strongly associated with manganese accumulation. The gene directly under this SNP is a putative natural resistance-associated macrophage protein (NRAMP2) metal ion transporter (Figure 3). The second highest SNP for the combined location GWAS is on chromosome 3 with an RMIP of 0.83 (Fig 3). This SNP is significant in all four of the locations and it falls directly in a gene annotated as manganese transporter protein 11 (MTP11) homolog. MTP11 is a member of the cation diffusion facilitator (CDF) family and is known to have roles in known to have roles in manganese tolerance and homeostasis [28,29].

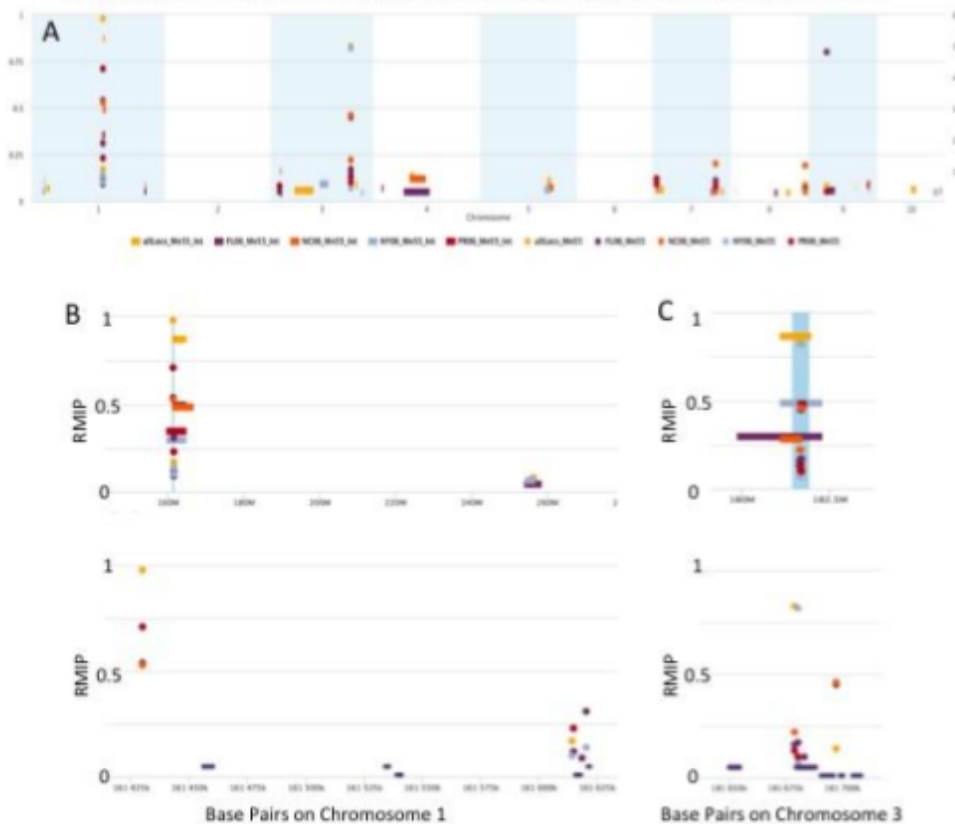


Figure 3. Multilocation hit for two candidate genes contributing to manganese accumulation. A. Whole genome view showing GWAS and Joint Linkage hits. B. Chromosome level (top) and Gene level (bottom) zoom of chromosome 1 peak. C. Chromosome level (top) and Gene level (bottom) zoom of chromosome 3 peak.

Molybdenum

We identified 13 unique QTL for Mo in the joint linkage analysis, two of which were extremely strong (F-test of 32 and 18 with a experiment wide significance cutoff < 3) (Figure 4). Candidate genes for Mo accumulation include those from the biosynthetic pathway of the molybdopterin cofactor and Mo transporters with orthologs in Arabidopsis. A biosynthetic gene and the maize ortholog of the Arabidopsis Mot1 transporter are located under the top two joint linkage peaks and thus represent strong candidates for the likely quantitative trait genes (QTGs).

Molybdenum Gene by Environment Interaction

The two top molybdenum loci showed a strong reciprocal gene by environment interaction (Figure 5). While both loci were significant in the joint linkage in all four environments (and the all locations), in FL, PR and NY, the Chr 1 loci had a substantially higher F statistic, while in NC the Chr. 5 loci was stronger than Chr 1. While we do not know what environmental parameters drove this difference, the growth environment is having a strong effect on the genetic cause of the phenotype.

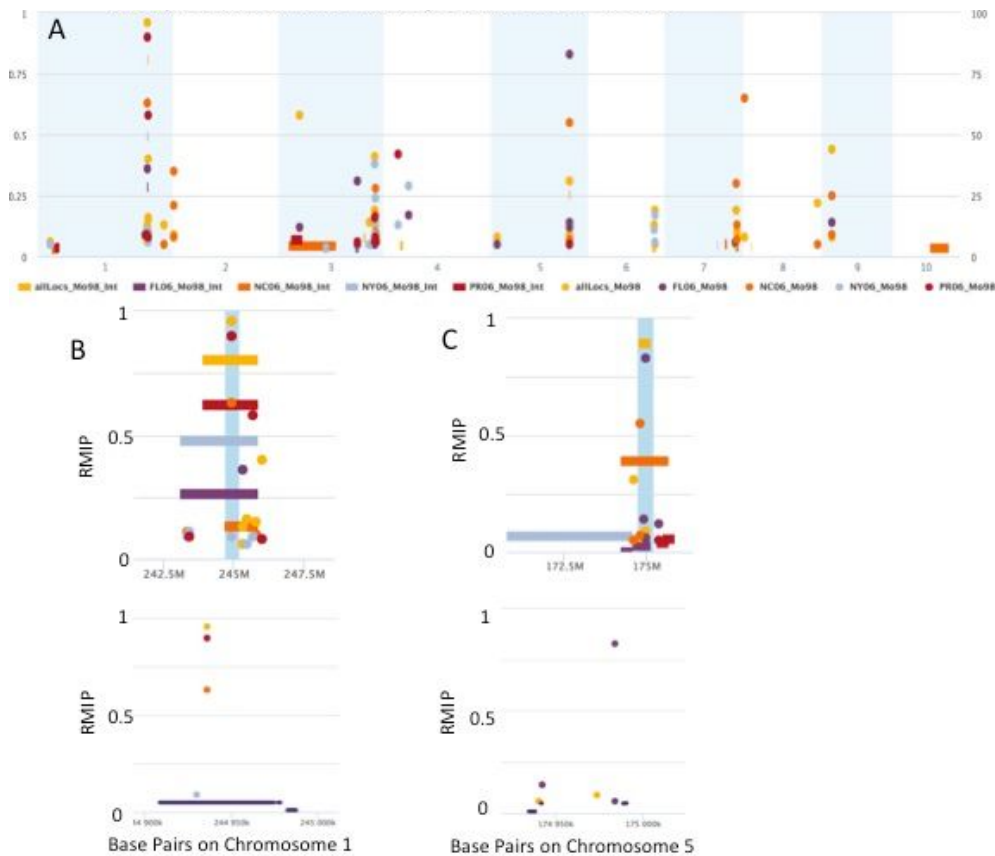


Figure 4. Manhattan plot showing Joint Linkage and GWAS results for molybdenum in 4 locations and 1 combined location. A. Whole genome view showing GWAS and Joint Linkage hits. B. Chromosome level (Top) and Gene Level (Bottom) zoom of peak over Mot1 on chromosome 1. C. Chromosome level (Top) and Gene Level (Bottom) zoom of peak over Mot1 on chromosome 5.

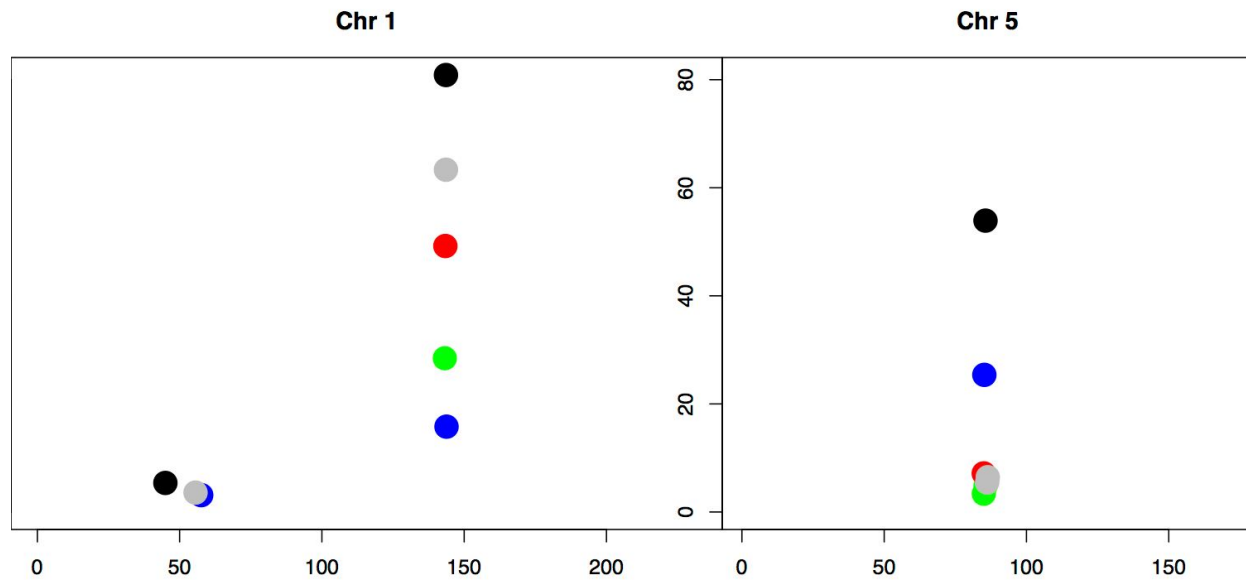


Figure 5. F-value of the two strongest Molybdenum loci in the four locations and the all locations model. Black (all locs), Green (FL), Blue (NC), Red (NY), and Grey (PR) denote the locations.

Phosphorus

A phosphorus QTL on chromosome 9 was returned for 3 out of the 4 locations (NC, NY, and FL) (Figure 6). A likely candidate gene for this peak is GRMZM2G444801 which is an ortholog of *low-phytic acid-1* (*lpa1*) in barley. In barley, mutations in *lpa1* heavily influence seed phosphorus composition [30]. Interestingly, in Maize, *lpa1* mutants have been characterized as altering the balance of phytic acid to inorganic phosphorus, but not the total phosphorus content [31]. However, from our results, it appears that diversity in *lpa1* does have some phenotypic effect on total phosphorus content.

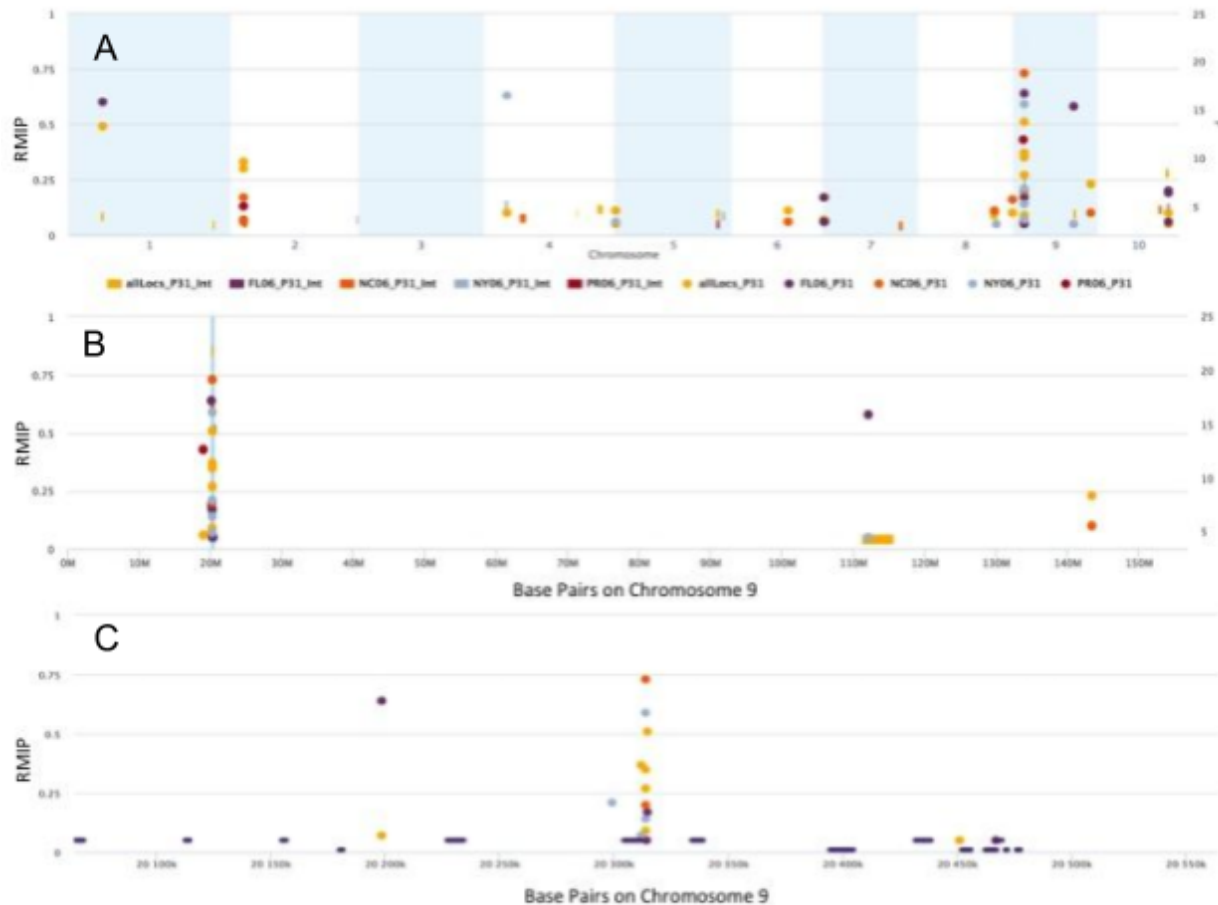
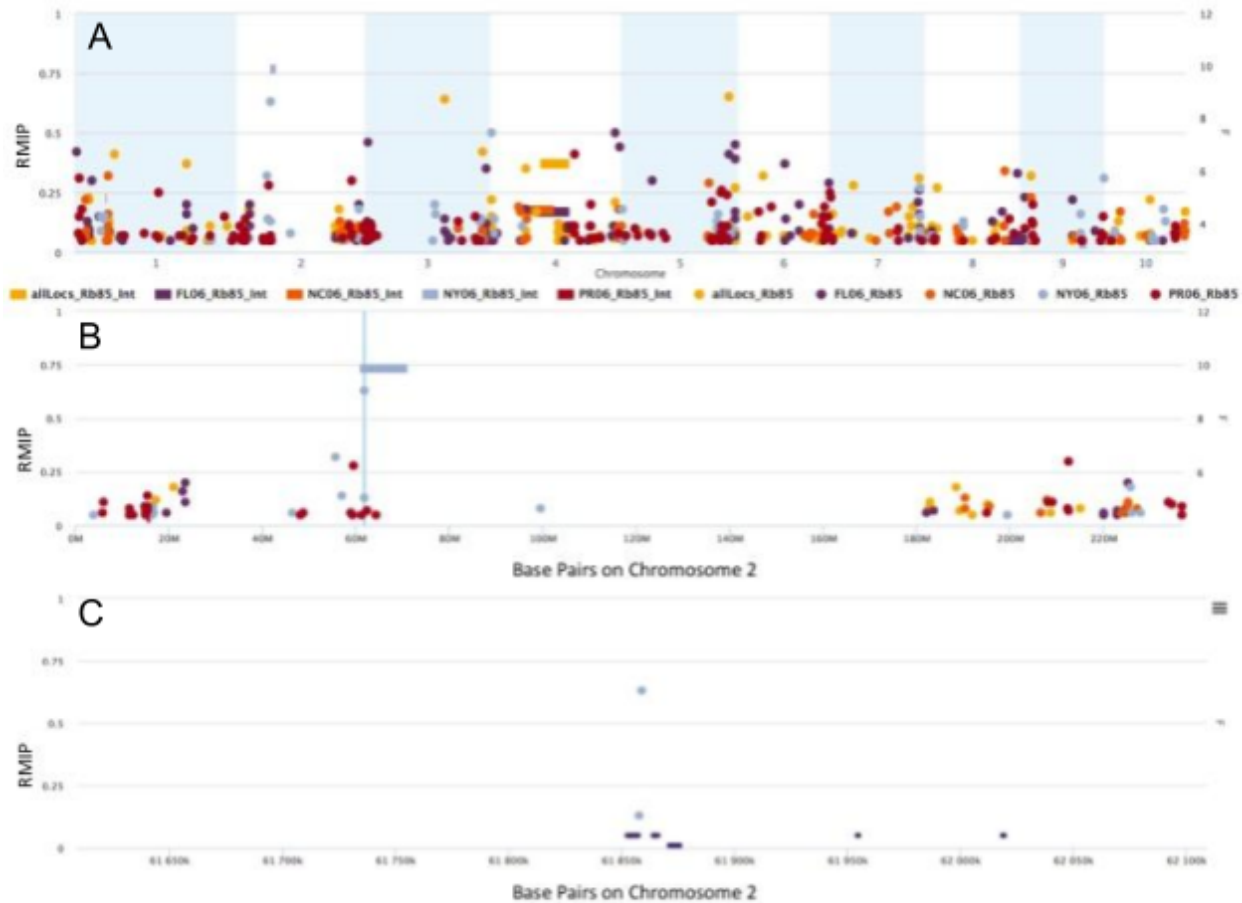


Figure 6. Manhattan plot of GWAS and Joint Linkage results for phosphorus. Displaying SNPs found in two or more locations. A. Genome-wide manhattan plot. B. Chromosome level zoom to peak on chromosome 9. C. Gene-level zoom to peak on chromosome 9.

Rubidium

The most significant hit across the 4 locations tested for rubidium in both the joint linkage and GWAS analysis was a SNP on chromosome 2. This SNP was only returned as significant in the New York growout. Interestingly, the New York location did not have the highest median accumulation of rubidium (Figure 1). This rubidium hit is very close to two orthologs of an *Arabidopsis* high-affinity potassium transporter (Figure 7). In *Arabidopsis*, this gene has been shown to play a role in potassium acquisition under low-potassium conditions [32]. In sunflowers, it has been shown that there is a change in rubidium uptake kinetics depending on potassium availability [33]. The unique interplay between environmental factors and genetics found in the New York NAM grow-out resulted in a significant difference in rubidium accumulation in plants.



Conclusion.

Here we have analyzed the elemental composition of over 50,000 kernels of the maize Nested association mapping population grown in four locations. We demonstrate that the elemental composition is heritable with a large genetic by environment interaction. Using a two step genetic mapping approach, we identified >300 loci controlling 19 of the 20 elements that we measured. Strong signal was observed around orthologs of known elemental accumulation genes. This data will be a rich resource for identification of the genes driving elemental accumulation in maize.

Acknowledgements: The authors would like to thank the Maize Diversity project for creating and sharing the resources used in this project. They would also like to thank Jason Wallace and Margaret Smith for technical assistance and Cyverse for computational support. The research was funded by U.S. National Science Foundation, Plant Genome Research Program (grant

#IOS 1126950 awarded to Ivan Baxter and Owen Hoekenga, IOS-0922493 (past) and IOS-1546657 (current) awarded to Michael Gore) and United States Department of Agriculture-Agricultural Research Service Intramural funds to Ivan Baxter. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Marschner H. Marschner's Mineral Nutrition of Higher Plants. Academic Press; 2011.
2. Gibson L, Benson G. Origin, History, and uses of corn (Zea mays). Iowa state university Department of Agronomy [En línea][citado 4 noviembre 2011]. 2002;
3. Food, Agriculture Organization of the United, Nations. FAOSTAT statistics database [Internet]. [Rome?]: FAO, 1998-; 1998. Available: <https://search.library.wisc.edu/catalog/999882363002121>
4. Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, et al. The genetic architecture of maize flowering time. *Science*. science.sciencemag.org; 2009;325: 714–718.
5. Frey F, Urbany C, Huettel B, Reinhardt R, Stich B. Heat Tolerance of Maize - Improving the Tolerance By Next Generation Plant Science. 2014.
6. Baxter IR, Ziegler G, Lahner B, Mickelbart MV, Foley R, Danku J, et al. Single-kernel ionic profiles are highly heritable indicators of genetic and environmental influences on elemental accumulation in maize grain (Zea mays). *PLoS One*. journals.plos.org; 2014;9: e87628.
7. United States Department of Agriculture. Soil Survey Staff, Natural Resources Conservation Service.
8. Brown PJ, Upadyayula N, Mahone GS, Tian F, Bradbury PJ, Myles S, et al. Distinct genetic architectures for male and female inflorescence traits of maize. *PLoS Genet*. journals.plos.org; 2011;7: e1002383.
9. Davies L, Gather U. The Identification of Multiple Outliers. *J Am Stat Assoc*. Taylor & Francis; 1993;88: 782–792.
10. Box GEP, Cox DR. An Analysis of Transformations. *J R Stat Soc Series B Stat Methodol*. [Royal Statistical Society, Wiley]; 1964;26: 211–252.

11. Holland JB, Nyquist WE. Estimating and interpreting heritability for plant breeding: an update. *Plant Breed.* books.google.com; 2003; Available: https://books.google.com/books?hl=en&lr=&id=oiwSW9l1SXQC&oi=fnd&pg=PA9&dq=Estimating+and+interpreting+heritability+for+plant+breeding:+An+update&ots=ocgTGkiVEI&sig=bHg7Rz1Q8R8dAQMNRNBpaM69Hf_c
12. Littell RC. SAS. Wiley StatsRef: Statistics Reference Online. John Wiley & Sons, Ltd; 2014.
13. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics.* academic.oup.com; 2007;23: 2633–2635.
14. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One.* journals.plos.org; 2011;6: e19379.
15. Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. *Genetics.* Genetics Soc America; 1994;138: 963–971.
16. Wallace JG, Bradbury PJ, Zhang N, Gibon Y, Stitt M, Buckler ES. Association mapping across numerous traits reveals patterns of functional variation in maize. *PLoS Genet.* journals.plos.org; 2014;10: e1004845.
17. Cook JP, McMullen MD, Holland JB, Tian F, Bradbury P, Ross-Ibarra J, et al. Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. *Plant Physiol. Am Soc Plant Biol*; 2012;158: 824–834.
18. Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, et al. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat Genet.* nature.com; 2011;43: 159–162.
19. Gore MA, Chia J-M, Elshire RJ, Sun Q, Ersoz ES, Hurwitz BL, et al. A first-generation haplotype map of maize. *Science.* science.sciencemag.org; 2009;326: 1115–1117.
20. Chia J-M, Song C, Bradbury PJ, Costich D, de Leon N, Doebley J, et al. Maize HapMap2 identifies extant variation from a genome in flux. *Nat Genet.* nature.com; 2012;44: 803–807.
21. Valdar W, Holmes CC, Mott R, Flint J. Mapping in structured populations by resample model averaging. *Genetics.* Genetics Soc America; 2009;182: 1263–1277.
22. Yu J, Holland JB, McMullen MD, Buckler ES. Genetic design and statistical power of nested association mapping in maize. *Genetics.* Genetics Soc America; 2008;178: 539–551.
23. Flint-Garcia SA, Thuillet A-C, Yu J, Pressoir G, Romero SM, Mitchell SE, et al. Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J. Wiley Online Library*; 2005;44: 1054–1064.
24. Hung H-Y, Shannon LM, Tian F, Bradbury PJ, Chen C, Flint-Garcia SA, et al. ZmCCT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *Proc Natl Acad Sci U S A. National Acad Sciences*; 2012;109: E1913–21.

25. Ferreira KN, Iverson TM, Maghlaoui K, Barber J, Iwata S. Architecture of the photosynthetic oxygen-evolving center. *Science*. science.sciencemag.org; 2004;303: 1831–1838.
26. Krämer U, Talke IN, Hanikenne M. Transition metal transport. *FEBS Lett*. Wiley Online Library; 2007;581: 2263–2272.
27. Mäser P, Thomine S, Schroeder JI, Ward JM, Hirschi K, Sze H, et al. Phylogenetic relationships within cation transporter families of Arabidopsis. *Plant Physiol. Am Soc Plant Biol*; 2001;126: 1646–1667.
28. Delhaize E, Gruber BD, Pittman JK, White RG, Leung H, Miao Y, et al. A role for the AtMTP11 gene of Arabidopsis in manganese transport and tolerance. *Plant J. Wiley Online Library*; 2007;51: 198–210.
29. Gustin JL, Zanis MJ, Salt DE. Structure and evolution of the plant cation diffusion facilitator family of ion transporters. *BMC Evol Biol*. bmcevolbiol.biomedcentral.com; 2011;11: 76.
30. Dorsch JA, Cook A, Young KA, Anderson JM, Bauman AT, Volkmann CJ, et al. Seed phosphorus and inositol phosphate phenotype of barley low phytic acid genotypes. *Phytochemistry*. Elsevier; 2003;62: 691–706.
31. Raboy V, Gerbasi PF, Young KA, Stoneberg SD, Pickett SG, Bauman AT, et al. Origin and seed phenotype of maize low phytic acid 1-1 and low phytic acid 2-1. *Plant Physiol. Am Soc Plant Biol*; 2000;124: 355–368.
32. Nieves-Cordones M, Alemán F, Martínez V, Rubio F. The Arabidopsis thaliana HAK5 K⁺ transporter is required for plant growth and K⁺ acquisition from low K⁺ solutions under saline conditions. *Mol Plant*. Elsevier; 2010;3: 326–333.
33. Benlloch-González M, García J, García-Mateo M-J, Benlloch M, Fournier J-M. K⁺ starvation causes a differential effect on shoot and root K⁺ (Rb⁺) uptake in sunflower plants. *Sci Hortic* . Elsevier; 2012;146: 153–158.

