nature genetics

Supplementary information

https://doi.org/10.1038/s41588-024-01761-3

Seedling root system adaptation to water availability during maize domestication and global expansion

In the format provided by the authors and unedited

Supplementary Materials

This PDF file includes:

Supplementary Materials and Methods

Supplementary Figs. 1 to 17

Supplementary Materials and Methods

Maize mutants and population resources

Two near isogenic background lines (Ia5125 and IL101t) that differ in the presence of the recessive maize mutation *sugary1* (*su1*) and *shrunken2* (*sh2*) that enhance kernel sweetness[21] were used to test the effect of kernel starch and sugar levels on seminal roots. The maize mutant *rootless concerning crown and seminal roots* (*rtcs*) is defective in seminal root initiation during embryogenesis[84].

An introgression library (IL) containing 68 BC_5F_5 near isogenic lines was surveyed to understand the effect of a donor source genome in a common recipient genetic background using the Northern Flint traditional variety Gaspé Flint and the reference line B73 as donor and recipient genotypes, respectively[22].

A Mexican traditional varieties Multi-parent Advanced Generation InterCrosses (MAGIC) population was analyzed to explore local adaptation and functional diversity in Mexican native maize. Eight traditional varieties accessions were selected to cover major ecological groups[85] and biogeographic regions[86]. Five of the traditional varieties donors were represented by heterogeneous open-pollinated varieties (OPVs) provided by CIMMYT. Three of the donors of traditional varieties were partially inbred stocks, previously derived from OPVs[87]. The eight donors (A-H) were crossed in an 8-way MAGIC[88] scheme: first, four F_1-hybrids were generated (AxB; CxD; ExF; GxH), then 4-way hybrids ([AxB] x [CxD]; [ExF] x [GxH]), and then an 8-way mix. A single individual was used per F_1-cross, ensuring that only one haplotype was captured per donor. The 8-way mix was intermated for three generations by sibling cross before three rounds of self-pollination produced partially inbred S3 families. S3 families were bulked by sibling mating and tissue pooled from 5 S3 individuals for genotyping using the Illumina 50K chip. Analysis of genotyping data for parents, F_1 plants and derived families allowed the definition of the 8 captured parental haplotypes across ~20k polymorphic SNPs and calling of the same sites in the MAGIC families.

Imaging of maize root system in soil by MRI

To visualize root systems in soil, representative genotypes with different seminal root traits including teosinte I.A.12 (Ames 21793), the traditional varieties Navajo tribe (PI 311229) and Guatemala110 (PI 490825) and the modern inbred lines C30 (Ames 26815) and CML289 (Ames 32336) were investigated via MRI. In brief, the seeds were placed in a Petri dish on wet filter paper. The Petri dish was sealed with parafilm and stored lightproof for 24 h in the growth chamber (16 °C/20 °C night/day temperature, 14 h light per day, 60% relative humidity) to induce germination. Seeds were subsequently planted in pots with 8 cm inner diameter and 40 cm height using field soil (Sp2.1, Landwirtschaftliche Untersuchungs- und Forschungsanstalt Lufa, Speyer, Germany). Soil moisture was kept at 40% of the maximal water holding capacity (water content 8.9%$_{m/m}$). Lighting was provided by 400 W HPI (Philips, Hamburg, Germany) and 400 W SON-T lamps (Philips, Hamburg, Germany), which alternated every 2 h with 5 min overlap, giving PAR intensities between 350 and 450 µmol m^{-2} s^{-1} at canopy level with improved light spectrum compared to just one lamp type. The MRI images were acquired on a 4.7 T vertical magnet equipped with a Varian console[89] utilizing a birdcage RF coil with a 10 cm diameter. A multislice spin echo sequence was used with the following sequence parameters: 0.5 mm resolution, 1 mm slice thickness, 9.6 cm field of view, TE = 9 ms, TR = 2.85 s, Bandwidth = 156 kHz, two averages. Using this protocol, roots down to 300 µm diameter were visualized[89]. The Varian console was replaced by a MR Solutions (Guildford, UK) console. Due to the new console and the different soil used slight adjustments to the sequence parameters were necessary (TR = 2.8s, Bandwidth = 400 kHz, four averages). Measurements were performed on 9, 17 and 24 days after planting.

Non-invasive imaging of maize embryos by MRI

To explore whether embryo size is associated with the number of seminal roots, we applied a non-destructive MRI imaging method to visualize and calculate the volume of the embryo for 50 representative maize traditional varieties. In brief, MRI measurements of the dry maize kernels were performed on an Avance Neo 500 MHz Super Wide Bore NMR-spectrometer (Bruker Biospin) equipped with a 760 mT/m gradient system and a 1H probe head. The Spin Echo (SE) pulse sequence[90] with a repetition time of 1000 ms and echo time of 2.2 ms was used for the simultaneous 3D-analysis of up to 50 maize kernels. The field of view (FOV) was set to 53 mm × 45 mm × 45 mm with a matrix image size of 118 mm × 100 mm × 100 mm. Excitation and refocusing pulses with bandwidth of 3750 Hz were

applied to measure the total MR signal of the samples. To optimize the signal-to-noise ratio, the datasets were averaged 32 times leading to a total measurement of 88 h 53 min. Image processing and volumetric analysis were performed using MATLAB software (vR2019b, The MathWorks, Natick, MA, USA) and AMIRA software (Amira3D 2022.1, ThermoFisher Scientific, Inc., Schwerte, Germany).

Seed phenotyping of the US Ames inbred panel

To explore the relationship between seed size and seminal root number, we customized a high-throughput seed phenotyping method combined with image analysis using LemnaGrid (v1.0) analysis software (https://www.lemnatec.com/analysing-seeds). In brief, dry maize seeds of the Ames panel were arranged manually on a blue background and scanned by a high-resolution scanner (Epson 14000, Epson, Meerbusch, Germany). The area of embryo and endosperm from resulting images were detected and quantified by LemnaTec (Aachen, Germany) analytical software. Ten kernels per accession were analyzed via image processing and the fractions of the embryo and endosperm in each kernel were separately calculated. The size and ratio of the embryo to the whole seed area were correlated with the number of seminal roots across the whole Ames panel.

Root system architecture phenotyping using maize traditional varieties

To obtain architectural root traits from traditional varieties across different geographic regions, we used 218 US maize traditional varieties accessions and germinated them for three days in a paper roll system[45]. We randomly selected these varieties based on different classifications of geographic regions in the US and Canada and by considering the seminal root number variation (Supplementary Table 7). The equally germinated seeds were then transferred to the rhizobox system (60 cm × 40 cm × 40 cm) with 25 slides (45 cm × 35 cm) supplied with 10 l of deionized water. Each germinated seed was fixed on top of the blue paper (30 cm × 20 cm) attached on each slide. Additional germination paper was used to cover the germinated seed for maintaining the humidity of the whole system. The deionized water was supplied after 5-day growth and the seedlings were harvested after 9 days grown in this system. The fully developed root system on each blue background was scanned with an Epson scanner 14000 (Meerbusch, Germany), and the resulting images were analyzed accordingly.

Modelling of root hydraulic properties in maize traditional varieties

To understand how seminal root number associates with soil-plant water transmission, sixty-six maize traditional varieties from the subset of 218 US traditional varieties were grown under a photoperiod of 12 h, day/night temperature of 22 °C and relative humidity average of 50%. These sixty-six traditional maize varieties are representative of different geographic regions in the US and have different numbers of seminal roots (Supplementary Table 8). Plants were grown in 0.5L pots filled with 385 g of sandy soil. The soil was sieved before filling the pot with particles size of 1 mm to remove aggregates and to ensure homogeneous soil packing. The soil-filled pots were watered and the seeds were potted with one plant per pot in the three replicates for each genotype. The soil was covered with plastic beads to minimize soil evaporation. The plants were watered continuously. Transpiration measurement started after 10-15 days. Transpiration rates were obtained gravimetrically by weighing the pots at three-hour intervals. At the end of the measurement the leaf area of each plant was measured using the leaf area meter instrument LI3000C (LI-COR bioscience GmbH, Bad Homburg, Germany) by scanning each leaf individually.

A soil-plant hydraulic model was used to investigate the effects of differences in seminal roots on transpiration reduction[26,91-93]. The model calculates water fluxes across the soil-plant-atmosphere continuum considering series of resistances. The fitting parameter is the effective root length that is active in water uptake, which makes this model ideal to test differences in seminal roots and their length (Supplementary Table 14).

We used a simplified model that calculates series of resistances between the bulk soil, soil-root interface, and through the root to the leaf xylem. Water flow in the soil follows Darcy-Buckingham's equation, ignoring gravity:

$$q = -K_s(\psi_{soil})\frac{\partial \psi_{soil}}{\partial r} \qquad \text{(Eqn S1)}$$

Where q is the water flux (cm s^{-1}), K_s is the soil hydraulic conductivity (cm^2 s^{-1} hPa^{-1}), which is the function of the soil matric potential ψ_{soil} (hPa, 1 hPa ≈ 1 cm), r is the radial distance (cm), and $\frac{\partial \psi_{soil}}{\partial r}$ is the gradient

in the matric potential. Note that K_s has units of (cm s^{-1}) when the soil matric potential is expressed in heads unit.

The boundary conditions were expressed as follows:

$$q(r_0) = \frac{E}{2\pi r_0 L} \ , \qquad \text{(Eqn. S2)}$$

$$q(r_b) = 0 \ , \qquad \text{(Eqn. S3)}$$

where r_0 and r_b are the root radius and the exterior radius of soil around the root (cm), E is the transpiration rate ($cm^3\ s^{-1}$), L is the active root length in water uptake (cm). r_b is determined by L and V, the volume of the column (cm^3), according to:

$$r_b = \sqrt{\frac{V}{\pi L}} \qquad \text{(Eqn. S4)}$$

The soil hydraulic conductivity was determined according to the Brooks and Corey (1966) model[94]:

$$K_s(\psi_{soil}) = K_{sat} \left(\frac{\psi_{soil}}{\psi_0}\right)^{\tau}, \qquad \text{(Eqn. S5)}$$

where K_{sat} is the saturated hydraulic conductivity of the soil (cm s^{-1}), ψ_0 is the soil air entry value (hPa), and τ is a fitting parameter (-). From matching the PDI model[95]and Brooks and Corey model[94] for the experimentally-measured series of potentials we obtained the parameters for Eqn. S5. According to de Jong van Lier *et al.* (2008)[96], and assuming a steady-rate water flow in the soil, Eqn. S1 was reformulated using the matric flux potential (Φ, $cm^2\ s^{-1}$):

$$\Phi\ (\psi_{soil}) = \int_{-\infty}^{\psi_{soil}} K_s(x)dx. \qquad \text{(Eqn. S6)}$$

Combining Eqn. 5, Eqn. 6, and the radial Richards equation, we obtain the flux boundary condition at the root-soil interface, $\psi_{root\text{-}soil}$, according to Schröder *et. al*., (2009)[97]:

$$\Phi_{root\_soil} = -\frac{E}{2\pi L}\left(\frac{1}{2} - r_b^2\ \frac{\ln(r_b/r_0)}{r_b^2 - r_0^2}\right) + \Phi_{soil}, \qquad \text{(Eqn. S7)}$$

where Φ_{root_soil} is the matric flux potential at the root-soil interface ($cm^2\ s^{-1}$), and Φ_{soil} is the matric flux potential in the bulk soil ($cm^2\ s^{-1}$). Φ_{root_soil} is obtained calculating Eqn. S5 and S6.

The water flow in the root system is given by

$$E = -K_{root}\left(\psi_{root\_soil} - \psi_{root\_xylem}\right), \qquad \text{(Eqn. S8)}$$

where ψ_{root_xylem} and ψ_{root_soil} are the water potential at the root-soil interface and at the xylem collar and K_{root} ($cm^3\ s^{-1}\ MPa^{-1}$) is the root conductance (assumed to be constant).

The water flow in the aboveground xylem is defined as:

$$E = K_{xylem}\ (\psi_{xylem\_root} - \psi_{leaf}) \qquad \text{(Eqn. 9)}$$

where K_{xylem} is the aboveground xylem conductance ($cm^3\ MPa^{-1}\ s^{-1}$), and $\psi_{xylem\_root} - \psi_{leaf}$ is the water potential difference (driving force) between root and leaf (MPa). K_{xylem} is derived from K_{root} and declines as leaf water potential falls below the xylem embolism threshold:

$$K_{xylem}(\psi) = K_{root}\left(\frac{\psi_{leaf}}{\psi_{xylem\_0}}\right)^{-\tau_x} \qquad \text{(Eqn. 10)}$$

where ψ_{xylem_0} is the water potential (MPa) at which emboli arise in the xylem, resulting in a decrease in K_{xylem}. τ_x is a fitted parameter. The water flow in the plant is defined as:

$$E = K_{plant}\ (\psi_{soil\_root} - \psi_{leaf}) \qquad \text{(Eqn. 11)}$$

Similarly, the water flow in the soil-plant system is represented by a function of soil-plant hydraulic conductance and water potential difference between soil and leaf:

$$E = K_{soil\_plant} (\psi_{soil} - \psi_{leaf}) \quad \text{(Eqn. 12)}$$

The onset of hydraulic limitation is defined as the point at which the slope of $E(\psi_{leaf})$ reaches 70% of its maximum.

Lignin analysis and salt stress treatments in hydroponics

To understand how root physiology is associated with the availability of water, we applied osmotic/salt stress to simulate a limited water availability in hydroponic solution. A total of 27 representative traditional varieties with contrasting numbers of seminal roots from this 218 traditional varieties panel collected from the US were investigated for stress resistance. 10 kernels per accession were sterilized and pregerminated in a paper roll system[45] with deionized water for three days. Germinated seeds were then transferred to half-strength Hoagland solution for another four days in hydroponics. The conditioned seedlings were then separately grown under control conditions with the same Hoagland solution and salt stress treatment using 160 mM NaCl, equivalent to a water potential of -0.8 MPa. The plants were harvested after six days of growth under these two treatments. The primary roots of those accessions were hand dissected from the root tip to the first emerged lateral root, and stored in 70% ethanol. The length and diameter of the dissected root segment were recorded for normalization of the data. Each biological replicate contains three independent primary roots from three seedlings each.

Root samples were enzymatically digested using 2% cellulase (Novozymes, Sigma-Aldrich, USA) and 2% pectinase (Novozymes, Sigma-Aldrich, USA) dissolved in 0.01 M citric buffer (Carl Roth, Karlsruhe, Germany) with the pH adjusted to 3.0. To inhibit microbial growth, 1 mM of NaN_3 (Fluka, Sigma-Aldrich, USA) was added to the enzyme solution. Once every two days, the solution was changed until all lignified cell walls were free of non-lignified tissues. Isolated lignified cell walls were cleaned with 10 mM borate buffer, subsequently rinsed with deionized water (Carl Roth, Karlsruhe, Germany) and dried. The samples were weighed and about 2 mg of the sample were ground in a steel ball mill (Retsch MM400; Retsch GmbH, Haan, Germany) for 8 min at a 30 Hz frequency in 2 ml Eppendorf tubes. The powdered samples were vortexed thoroughly in 0.75 ml of acetone. This solution was transferred to pre-labeled, pre-weighed glass vials. Vials were placed on a heat block (MHR 23, HLC BioTech, Germany) to evaporate all the liquid in the vial, and the final dry matter was used for lignin analysis.

For chemical analysis of lignin, the samples were depolymerized by thioacidolyisis[98]. This method was optimized for microanalysis following published work[99]. After thioacidolysis, samples were extracted with chloroform and spiked with dotriacontane (20 µg of dotriacontane in 100 µl of solution; Fulka, Sigma-Aldrich, USA) as an internal standard. GC-FID (CG-Hewlett Packard 5890 series H, Agilent) analysis was used to quantify lignin samples, and GC-MS (quadrupole mass selective detector HP 5971, Hewlett Packard, Agilent) analysis was used to identify individual lignin monomers. 1 µl of each sample was analyzed on 30 m GC columns (DB-1 Columns, Agilent, USA) using split/splitless injection. The compounds were then identified using reference tables with molecular ion fragments from published studies[100,101].

Construction and sequencing of maize MAGIC population

A Multiparent Advanced Generation Intercross (MAGIC) population was generated from eight Mexican native maize varieties sourced from across the Mexican environmental range, available from the CIMMYT germplasm bank. The eight founder parents and associated environmental data are provided further in Supplementary Table 15.

To make the MAGIC population, outbred accessions OAXA521, NAYA15 and JALI43 were substituted by the related inbred derivatives MR23 (Ames 30536), MR18 (Ames 30532) and MR21 (Ames 32914), respectively. In the first generation, four F1 stocks were generated by making the crosses A×B, C×D, E×F and G×H. In the second generation, a single individual per F1 was used as both male and female to make the four-way hybrids (A×B)×(C×D), (C×D)×(G×H), (G×H)×(E×F) and (E×F)×(A×B). The use of single F1 individuals ensured the population captured only 8 founding haplotypes despite the use of outbred founders. 120 seeds were advanced per four-way hybrid. In the third generation, eight-way

pools were made by mating individuals of (A×B)×(C×D) with those of (G×H)×(E×F), and separately those of (C×D)×(G×H) with (E×F)×(A×B). 80 ears (families) were advanced from each of the two eight-way pools. In a first cycle of intermating individuals from the 80 families in the first eight-way pool were crossed with this in the second, with no more than one cross between any two given families. 500 ears were recovered from the first intermating and a single seed per ear advanced to generate 500 seed bulks. In a second cycle of intermating, individuals from five 500 seed bulks were randomly crossed within bulks. Progeny were pooled across bulks to give 474 ears, again advanced as single seed. A third cycle of intermating used three 474 seed bulks, again individuals randomly mated within bulks and progeny pooled across bulks, recovering 390 ears. Three 390 bulks were self-pollinated and the recovered ears advanced ear-to-row through two more generations of self-pollination to generate the final Mexico MAGIC (MEMA) families. A pilot set of 129 families was genotyped using an Illumina 3k SNP array (https://www.illumina.com/products/by-type/microarray-kits/maize-ld.html) and evaluated for SRN. Founder parents and F1 individuals were also genotyped to allow founder haplotypes to be inferred using a standard triplet approach. SNP association with SRN was performed with R/qtl2 (doi: 10.1534/genetics.118.301595) using the R/qtl2::scan1snps function controlling for residual population structure with a kinship calculated by R/qtl2::calc_kinship under the loco option. SRN was estimated for "pseudoparents" (i.e. individuals homozygous for each of the eight founder haplotypes) using R/rrBLUP (doi:10.3835/plantgenome2011) to estimate marker effects from the 129 MEMA families. Ten estimates were calculated for each parental haplotype based on setting marker effects for each of the ten chromosomes in turn to zero. Effects for the 8 parental alleles were estimated using R/qtl2::scan1 and parental genotype probabilities calculated with R/qtl2::calc_genoprob.

Estimation of different germplasm for seminal root number

We first compiled the two published datasets from US[43] and European[44] maize inbred panels. Liu et al. (2003)[43] calculated the maximum-likelihood estimates of the allelic constitution of an inbred group or specific inbred line, given different proportions of ancestry from the four historical germplasm pools (Southern Dent, Northern Flint, Tropical Highland maize, and Tropical Lowland maize). Gouesnard et al. (2017)[44] performed the AMIXTURE analysis and identified seven different genetic ancestries (Italy, Argentina, Popcorn, Lacaune, Dent, Northern flint and Pyrenees-Galicia) based on the European germplasm. Since the classification of the germplasm groups differed between US and European, we performed the Pearson correlation analysis separately.

Identification of the transposon-induced mutants

To verify candidate genes from GWAS, we next identified potential loss-of-function maize mutations by exploring our in-house sequence indexed collection *BonnMu*[29]. Mutants of the *BonnMu* resource were derived from Mutator-tagged F_2-families in various genetic backgrounds, such as B73 and F7. All 160 candidate genes generated from the GWAS analysis were screened through our mutant stocks. In brief, the F_2 segregating seeds were phenotyped in paper-roll culture and the seedling plants were scanned using the scanner Expression 12000XL (Epson, Suwa, Japan). Seminal root number, primary root length and lateral root density were recorded. Statistical analyses were performed by pair-wise Student´s *t* tests with *F* statistics.

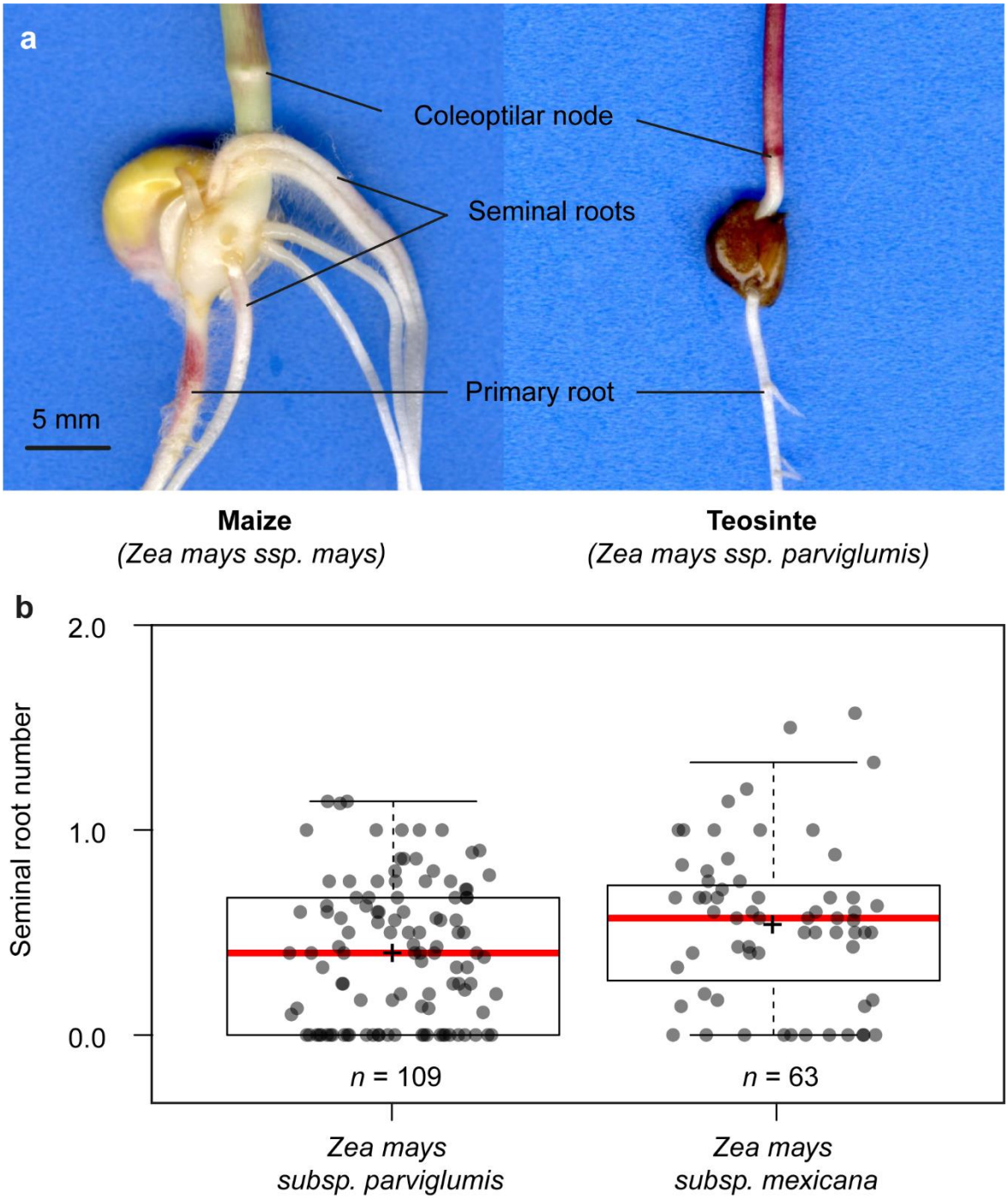
Drought tolerance trials using traditional varieties and the *rtcs* mutant

To understand whether different traditional maize varieties have variable capacity in tolerance to drought, we grew 82 georeferenced varieties with different C (n = 41) and A (n = 41) allele variations of the gene *ZmHb77*. Note that in this drought experiment we used the small pot size (5 cm × 5 cm × 17 cm, length × width × depth) with one plant grown per pot using 500 g of the same soil mixture. We used the small pots because we have only a limited amount of seeds of the traditional varieties. The drought treatment was maintained with the similar strategy as described above with 22% soil water capacity. The shoot dry biomass was determined for both drought and well-watered conditions after 20-d of cultivation.

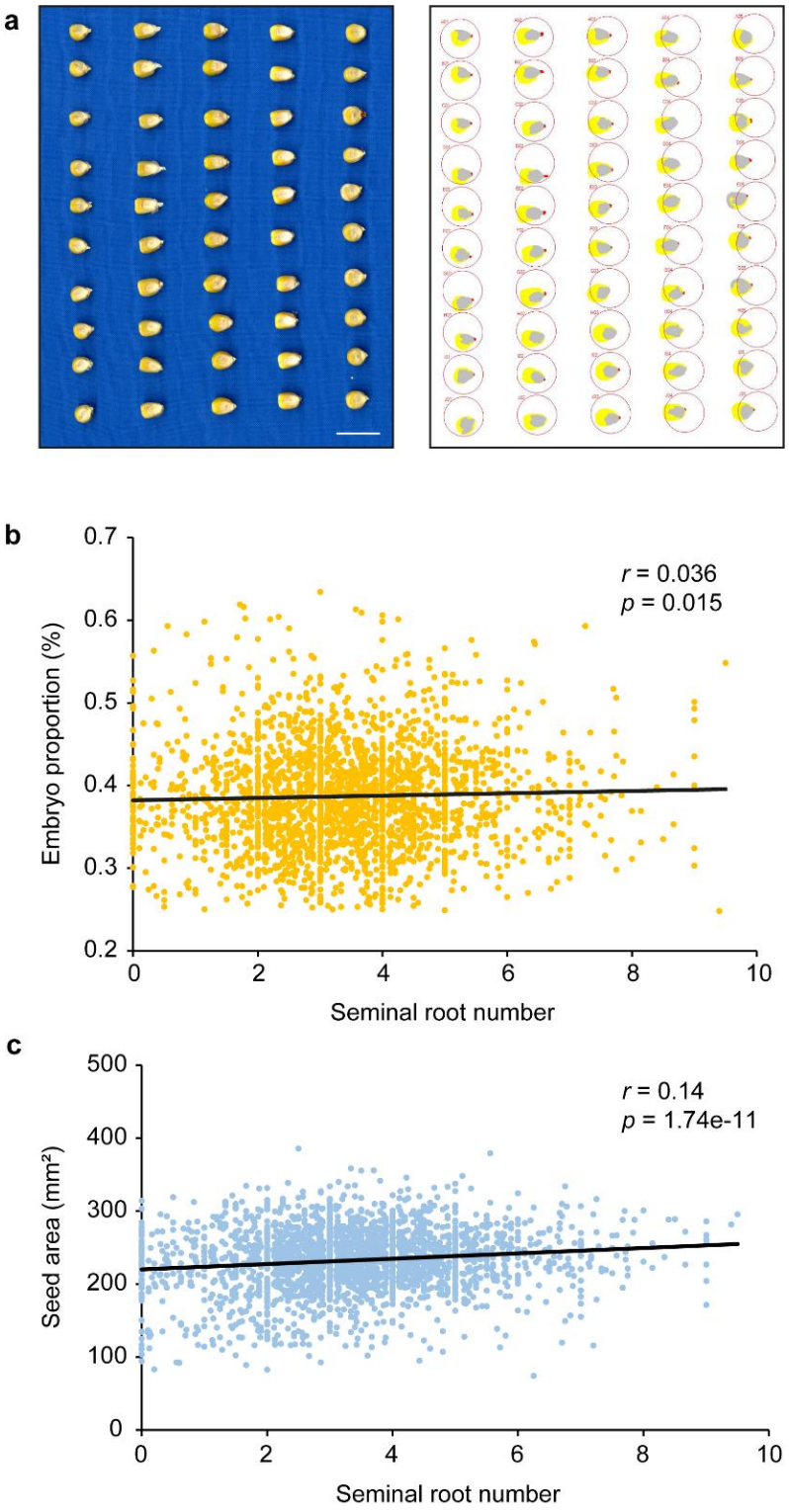
Moreover, we tested the drought tolerance and resilience of the *rtcs* mutant and its wild type that differed in seminal root number and lateral root branching. We performed the drought and recovery trials in the same small pot experiment. The reason why we used individual pots is that we had to genotype every plant separately to distinguish between mutant and wild type plants. Therefore, the cultivation box was not suitable for this experiment. We applied well-watered, drought and drought followed by re-watering

248 treatments to testify the drought tolerance and resilience as described above. The shoot dry biomass
249 and stomatal conductance were recorded.

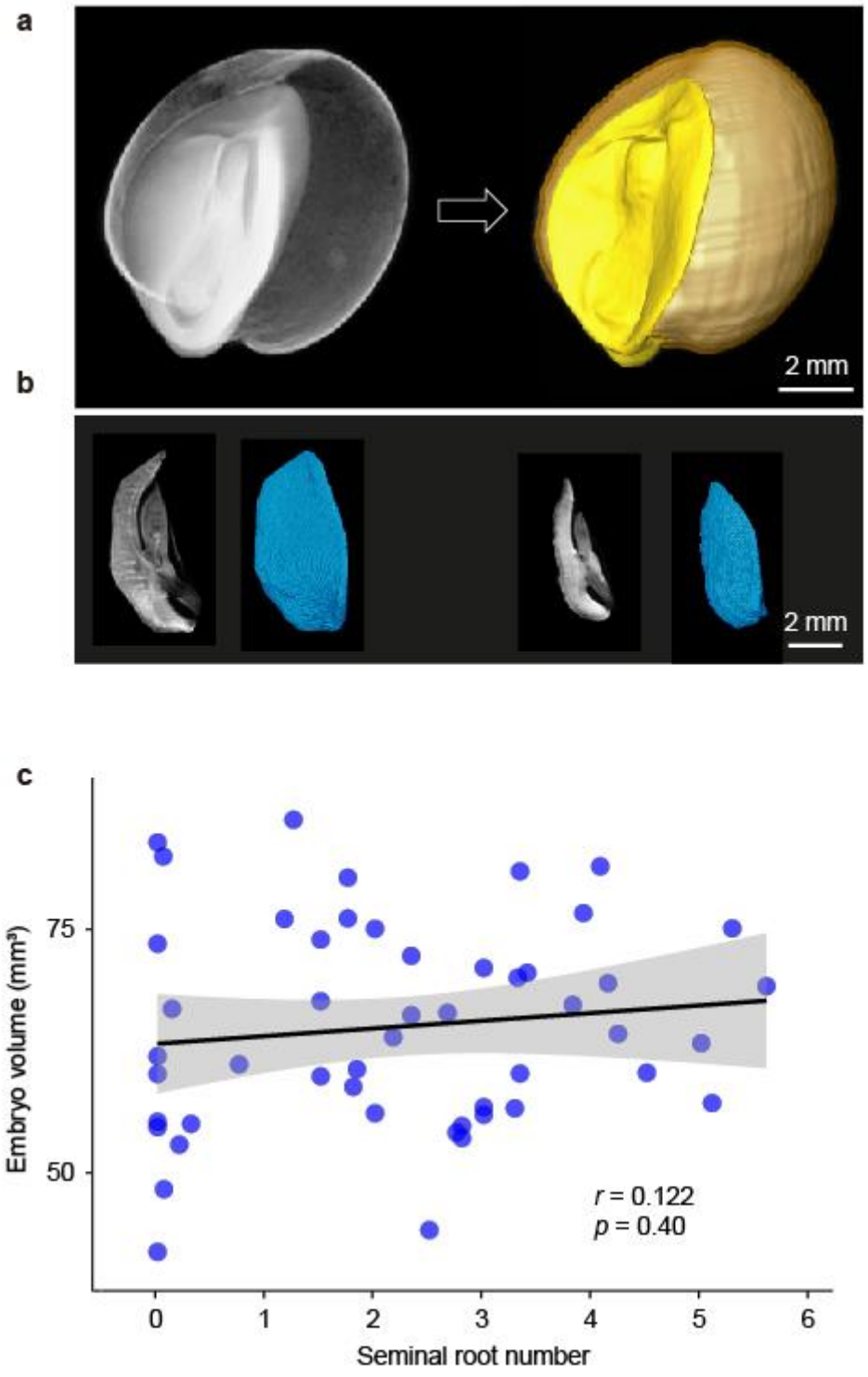
250



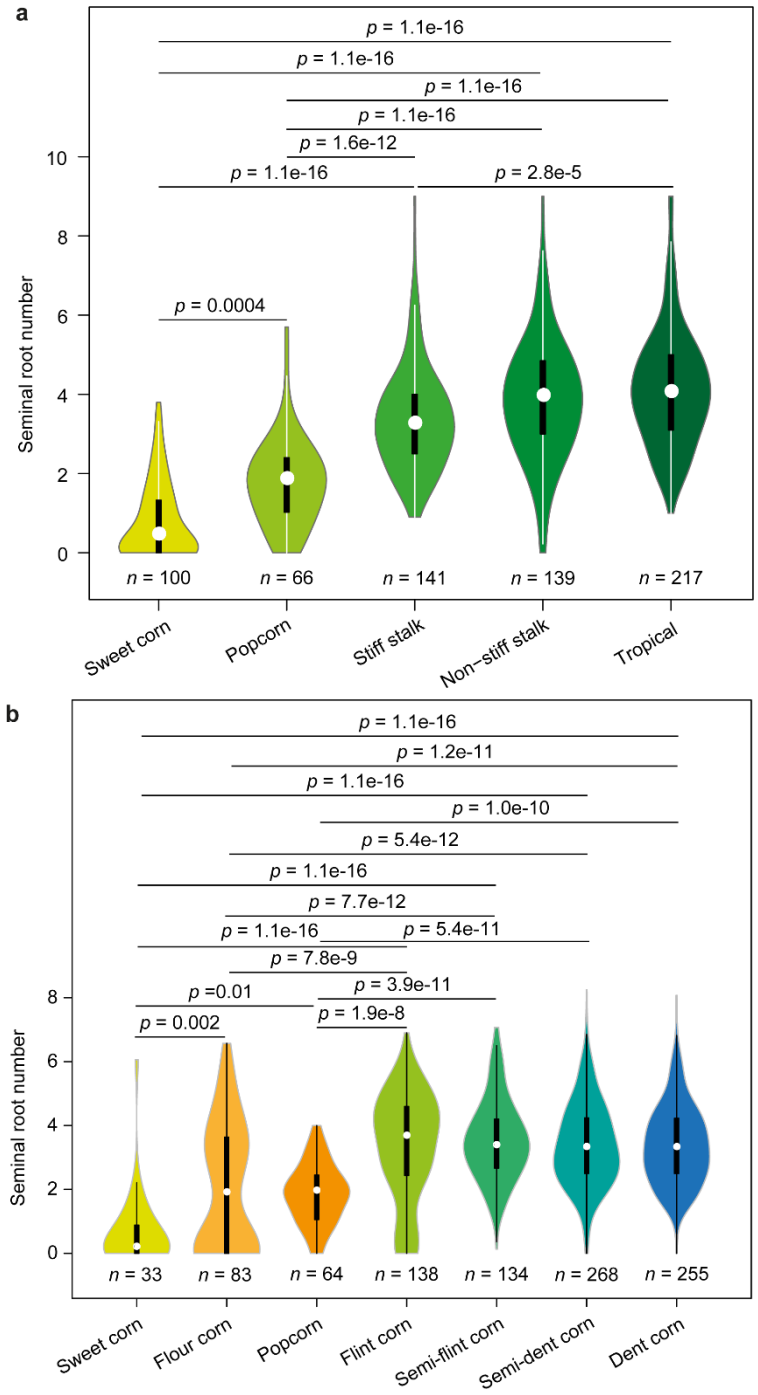
Supplementary Fig. 1. Morphological differences in maize and teosinte root systems. a, Visual appearance of the maize and teosinte seedling root system. Seeds were germinated for 10 days in wet paper-rolls before seedlings were scanned. **b**, Seminal root number of the wild lowland teosinte *Zea mays* ssp. *parviglumis* and highland teosinte ssp. *mexicana*. Boxes span from the first to the third quartiles, red lines represent the median and whiskers extend to 1.5× the interquartile range of the lower and upper quartiles. Data points outside of whiskers represent outliers. According to two-sided Wilcoxon rank sum test, SRN was significantly different at $p = 0.0145$.



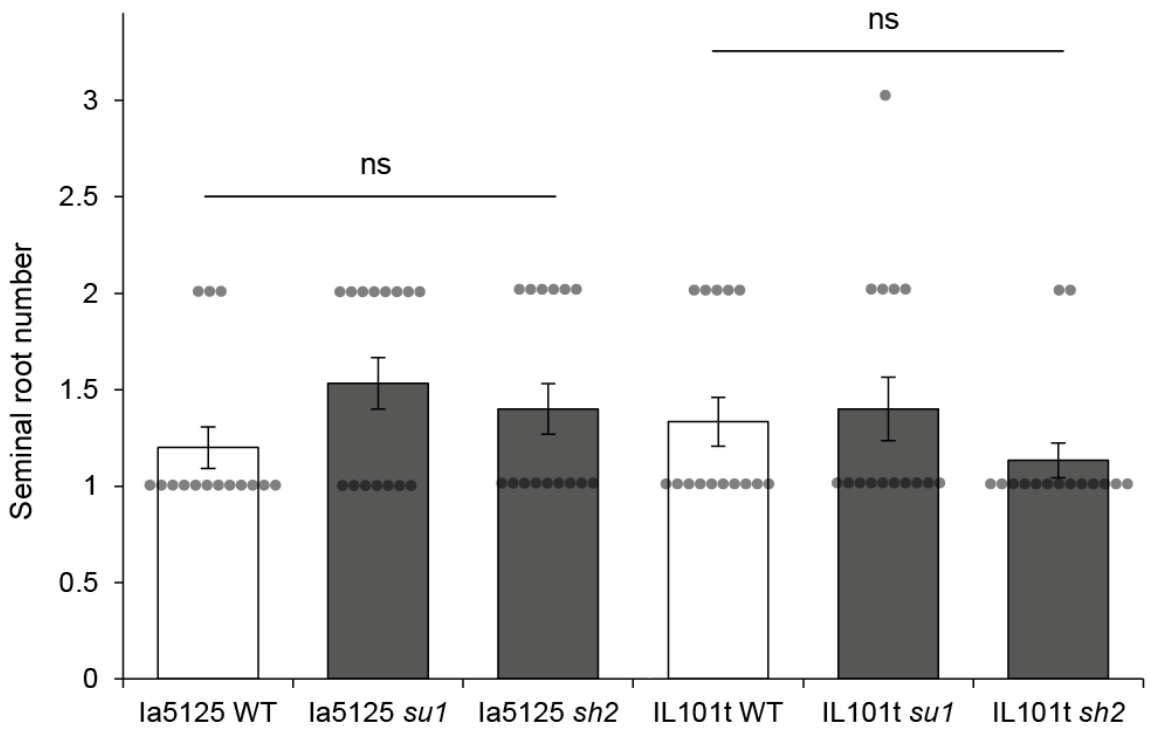
Supplementary Fig. 2. Relationship between seminal root number and embryo size in maize inbred lines. Dry maize seeds were scanned at high-resolution and the resulting images were processed in LemnaGrid (LemnaTech). **a**, The areas of embryo and endosperm were calculated separately. Scale bar = 2 cm. The size of seed (**b**) and ratio of the embryo to the whole seed size (**c**) were correlated with the number of seminal roots for the whole Ames panel. Scatter plots show best fit (solid line) and 95% confidence interval (colour shading) for linear regression. For both **b** and **c**, two-sided tests of significance with Bonferroni correction were performed to adjust the *P* value for multiple independent tests. *n* = 2429.



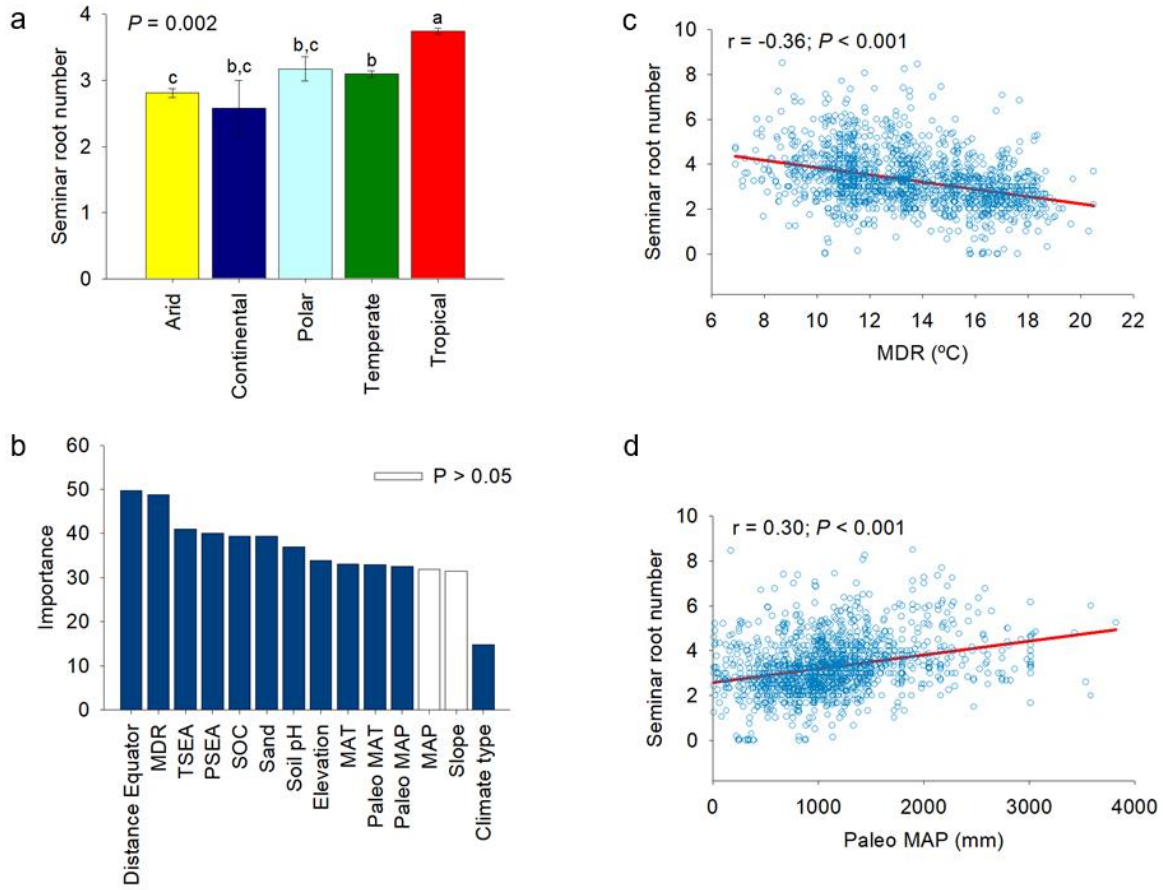
Supplementary Fig. 3. Relationship between seminal root number and embryo size in maize traditional varieties. 3-D visualization of a dry maize seed (**a**) and embryo (**b**) by nuclear magnetic resonance (NMR) imaging. (**c**) Scatter plot indicating the relationship between seminal root number and the volume of dry embryos in maize traditional varieties collected from diverse geographical regions in the US. Pearson correlation was performed and the *p* value indicates the probability at which the correlation coefficient (solid line) is zero (null hypothesis) at 95% confidence interval (shaded area). Two-sided test of significance with Bonferroni correction was performed to adjust the *P* value for the multiple independent tests. *n* = 50.



Supplementary Fig. 4. Variation in seminal root number across different maize germplasm of inbred lines (a) and traditional varieties (b). The inbred lines are from the USDA-ARS North Central Regional Plant Introduction Station (NCRPIS) in Ames, Iowa, and classified according to breeding program of origin, with most of the US programs in the two major germplasm groups, recognized by temperate maize breeders (referred to as stiff stalk and non-stiff stalk) and additional tropical origin. The traditional variety accessions are contributed from NCRPIS and CIMMYT. The germplasm groups of traditional varieties are derived from the narrative information of the US National Plant Germplasm System (https://npgsweb.ars-grin.gov/gringlobal). The exact number of inbred lines or traditional varieties are highlighted in the figure. *n* = 10 biologically independent seedlings per inbred line and *n* = 20 biologically independent seedlings per traditional variety. Violin plots show reaction norms and phenotypic variation within indicated sample sizes of different germplasm groups. White dots represent the median values and whiskers extend to 1.5× the interquartile range of the lower and upper quartiles. Significant differences among means are indicated by exact values (one-way ANOVA, Tukey's HSD).

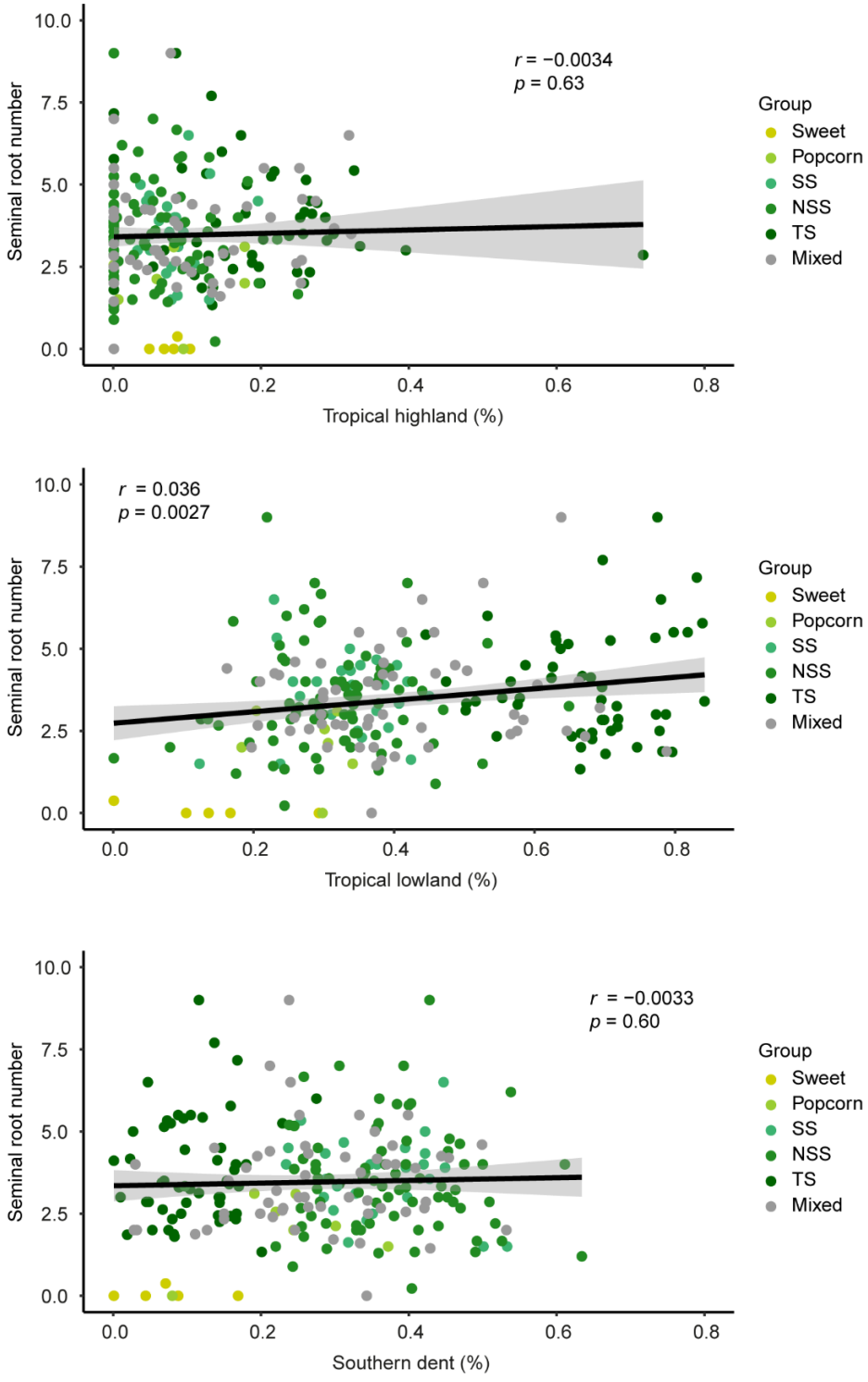


Supplementary Fig. 5. Seminal root number of starch-related mutants and their wild types in maize. *sugary1* (*su1*) and *shrunken2* (*sh2*) are mutant lines in endosperm starch biosynthesis. Mutant alleles of *su1* and *sh2* in two different genetic backgrounds la5125 and IL101t were germinated and the roots were scanned and counted. Significant differences were determined by one-way ANOVA with Tukey HSD at $p = 0.05$; ns, not significant. Data are presented as mean values +/- SEM and $n = 10$ biologically independent seedlings per genotype.

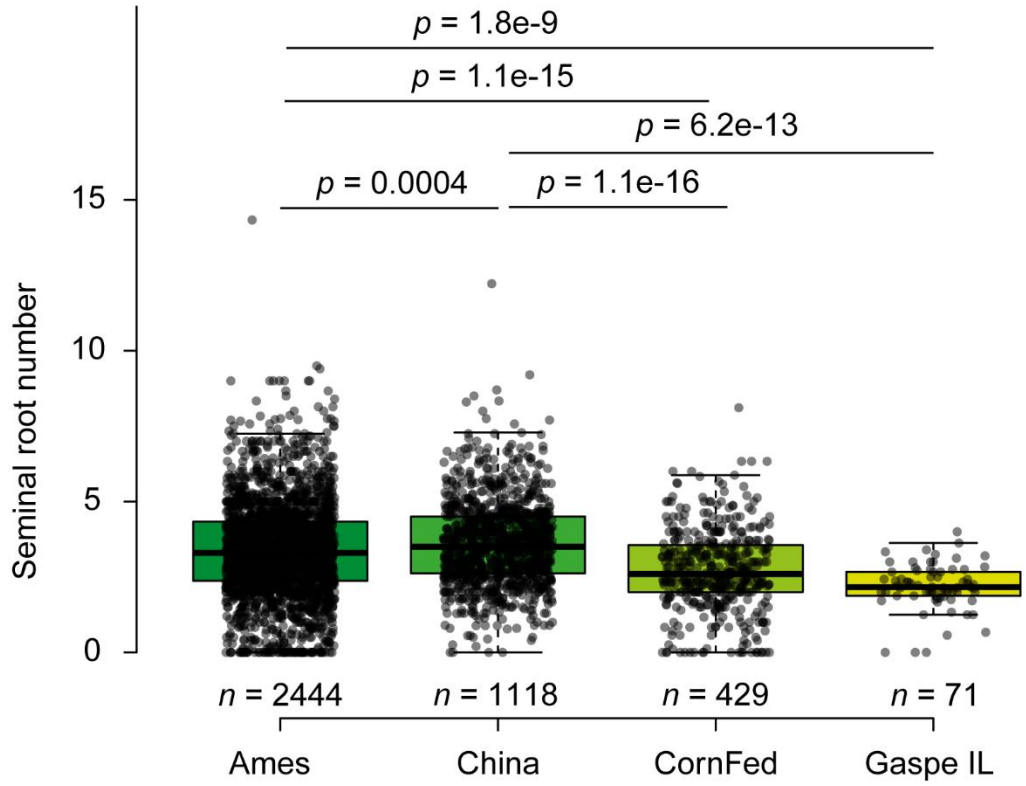


297

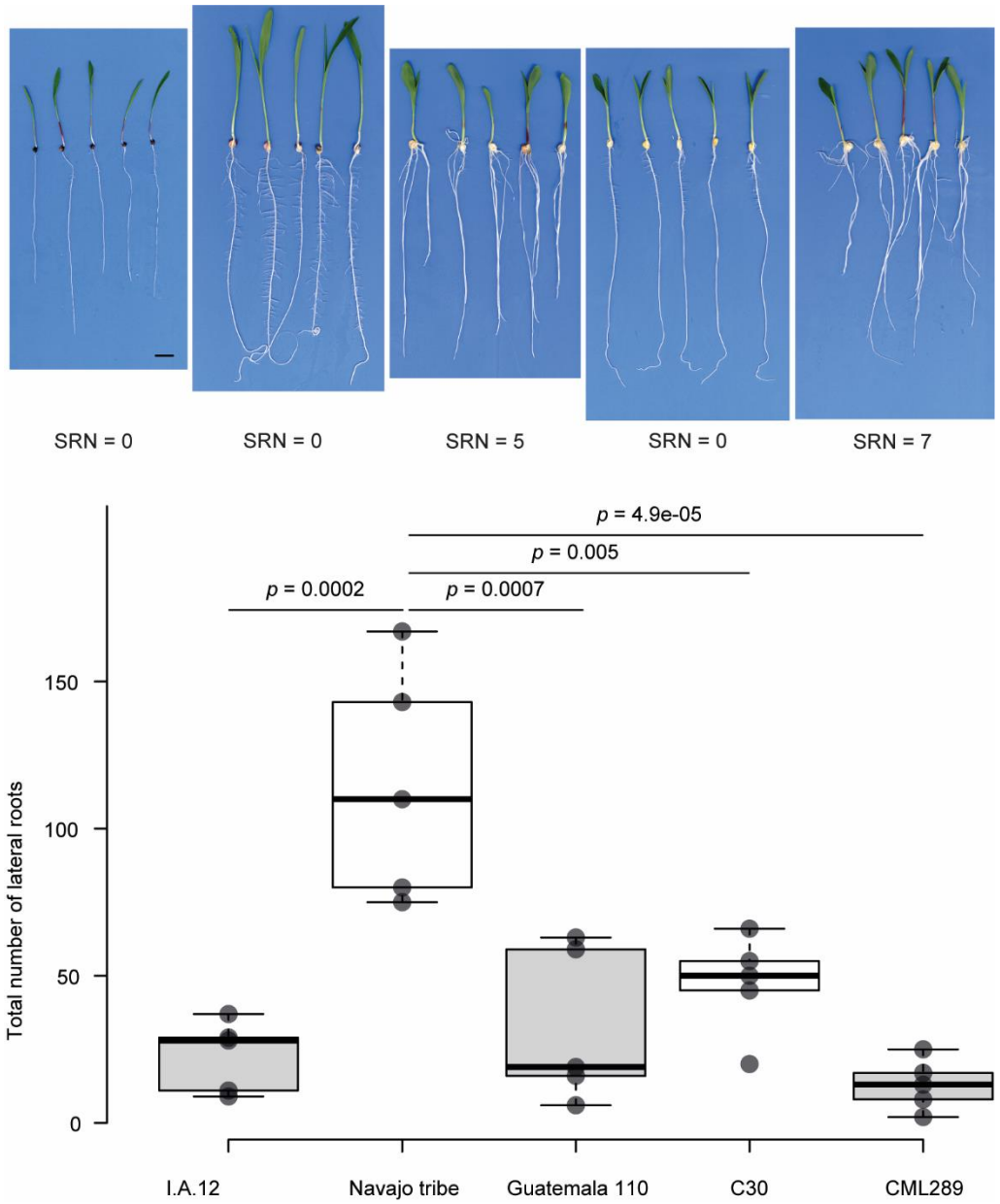
298 **Supplementary Fig. 6. Impact of environmental factors on seminal root number in maize**
299 traditional varieties**. a**, Comparison of seminal root numbers in maize lines originating from different
300 climate groups according to the Köppen-Geiger climate classification maps (v1; http://koeppen-
301 geiger.vu-wien.ac.at/present.htm). Significant differences were tested by one-way PERMANOVA
302 (Permutational Multivariate Analysis of Variance) with post-hoc test at $p < 0.05$. Arid ($n = 317$),
303 continental ($n = 16$), polar ($n = 45$), temperate ($n = 531$), tropical ($n = 575$). Data are presented as mean
304 values +/- SEM. **b**, Relative importance of soil and climate factors on seminal root number as obtained
305 by a Random Forest machine learning approach. Pearson correlation between MDR (**c**) and Paleo MAP
306 (**d**) with seminal root number. The *p* value indicates the probability at which the correlation coefficient is
307 zero (null hypothesis). For both **c** and **d**, two-sided tests of significance with Bonferroni correction were
308 performed to adjust the *P* value for the multiple independent tests. MDR, mean diurnal temperature
309 range; PSEA, precipitation seasonality; TSEA, temperature seasonality; MAT, mean annual
310 temperature; MAP, mean annual precipitation; Paleo MAT and MAP are mid-Holocene MAT and MAP
311 (about 6000 years ago) as described[102]. $n = 1484$ in panels **c** and **d**. Slope (terrain slope) in degrees.



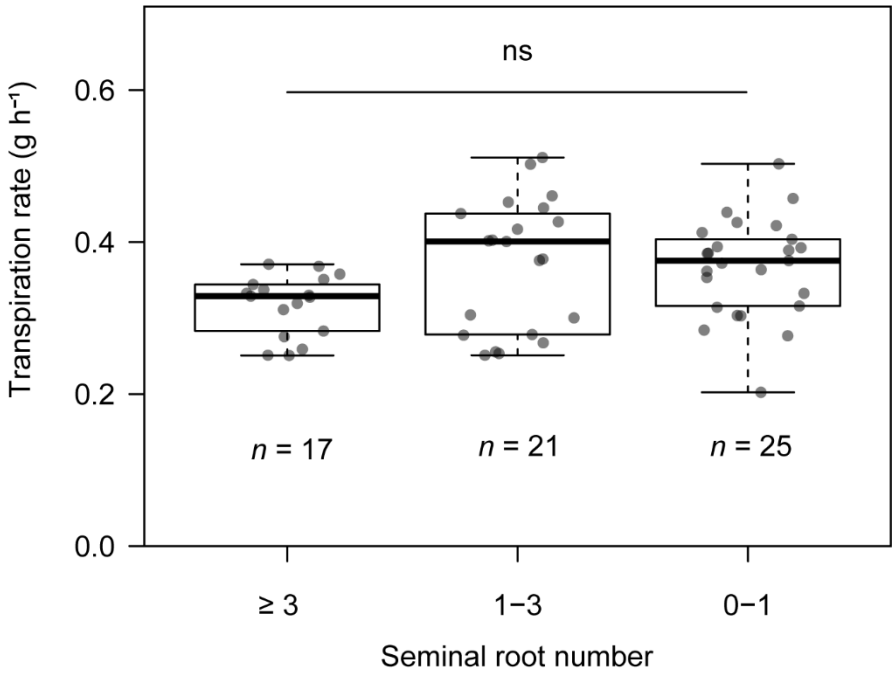
Supplementary Fig. 7. Pearson correlation between the proportion of Tropical highland, Tropical lowland, Southern dent germplasm and seminal root numbers in the US Ames panel (*n* = 225). Estimates of historical sources for individual inbred lines were extracted from published articles[43,44]. Here, the proportion of alleles from different germplasm pools (Tropical highland, Tropical lowland, Southern dent) was correlated with the number of seminal roots in the corresponding maize inbred lines. SS, stiff-stalk; NSS, non-stiff stalk; TS, tropical/sub-tropical. The *p* value denotes the probability at which the correlation coefficient (solid line) is zero (null hypothesis) at 95% confidence interval (shaded area). From **a** to **c**, two-sided tests of significance with Bonferroni correction were performed to adjust the *P* value for the multiple independent tests.



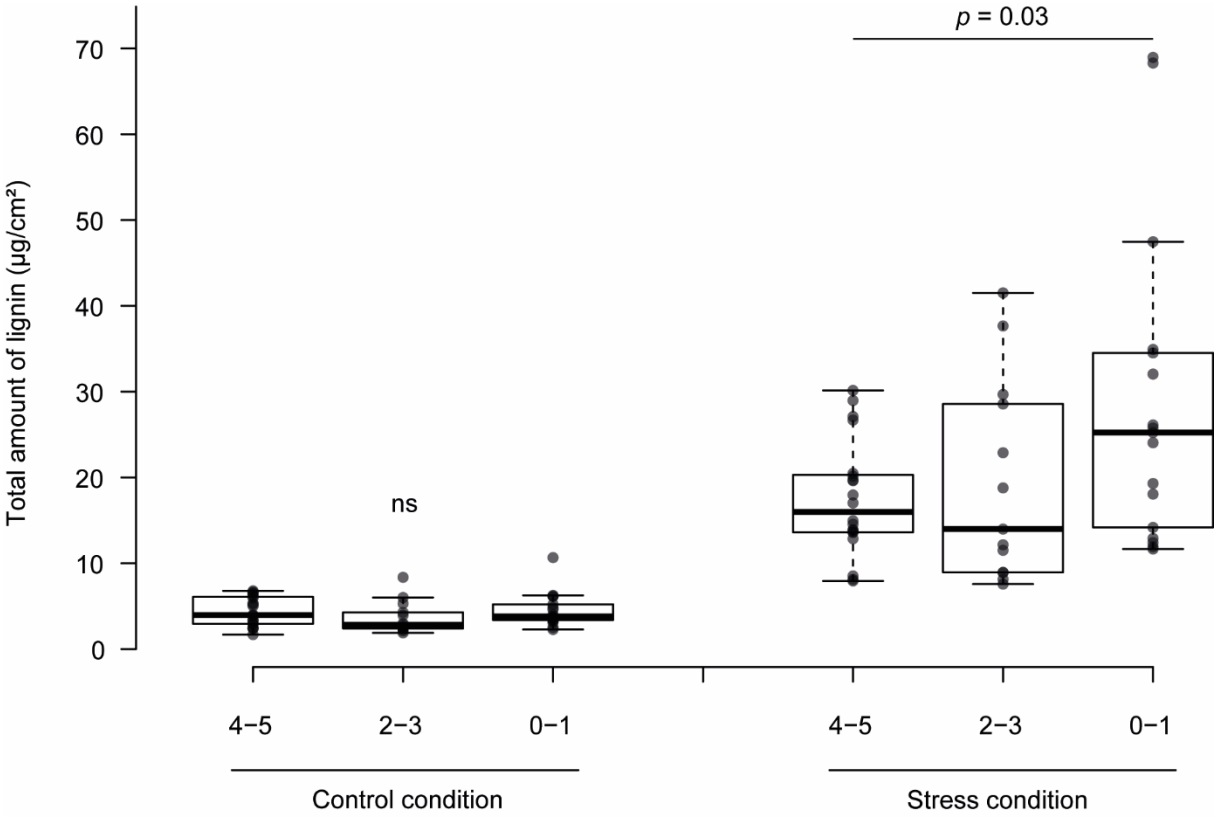
Supplementary Fig. 8. Comparison of seminal root numbers among lines from different breeding pools. Examined maize lines derived from the U.S. Ames panel, inbred lines from China, the European collection and the Gaspé flint introgression panel. The exact number of inbred lines are highlighted in the figure. *n* = 10 biologically independent seedlings per inbred line. Boxes span from the first to the third quartiles, center lines represent the median values and whiskers extend to 1.5× the interquartile range of the lower and upper quartiles. Data points outside of whiskers represent outliers. Significant differences are indicated by exact values according to one-way ANOVA with Tukey HSD.



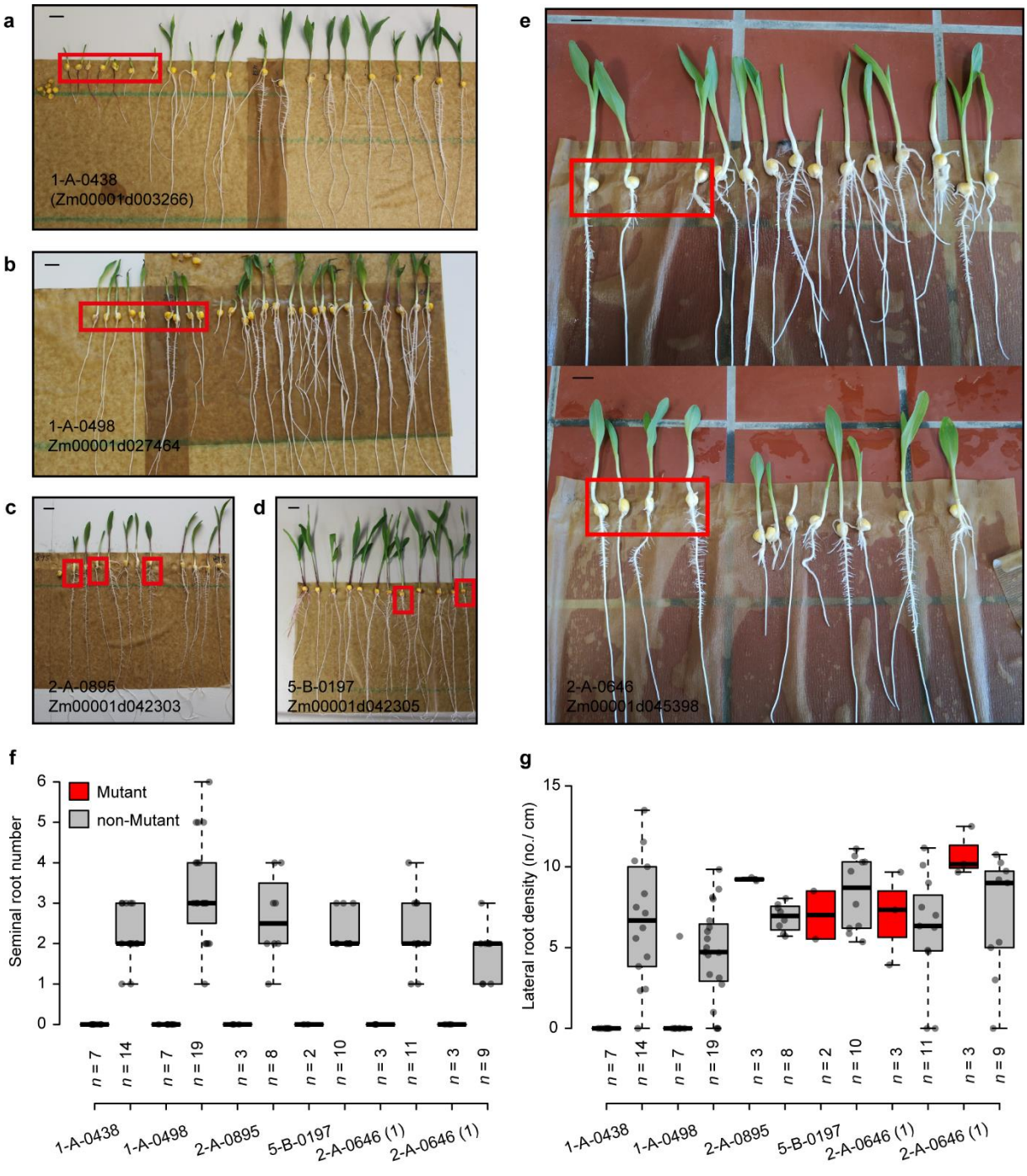
Supplementary Fig. 9. Lateral root phenotype of representative maize varieties. The seeds were germinated for 10 days and scanned by an Epson 14000 scanner. Teosinte: I.A.12 (Ames 21793); Traditional varieties: Navajo tribe (PI 311229); Guatemala 110 (PI 490825); Modern inbred lines: C30 (Ames 26815); CML289 (Ames 32336). *n* = 5 biologically independent seedlings per genotype. Statistical tests were performed by one-way ANOVA with Tukey HSD and labeled with exact values. Boxes span from the first to the third quartiles, center lines represent the median values and whiskers show data lying within 1.5× interquartile range of the lower and upper quartiles. Data points at the ends of whiskers represent outliers. For all figure panels, scale bar = 2cm.



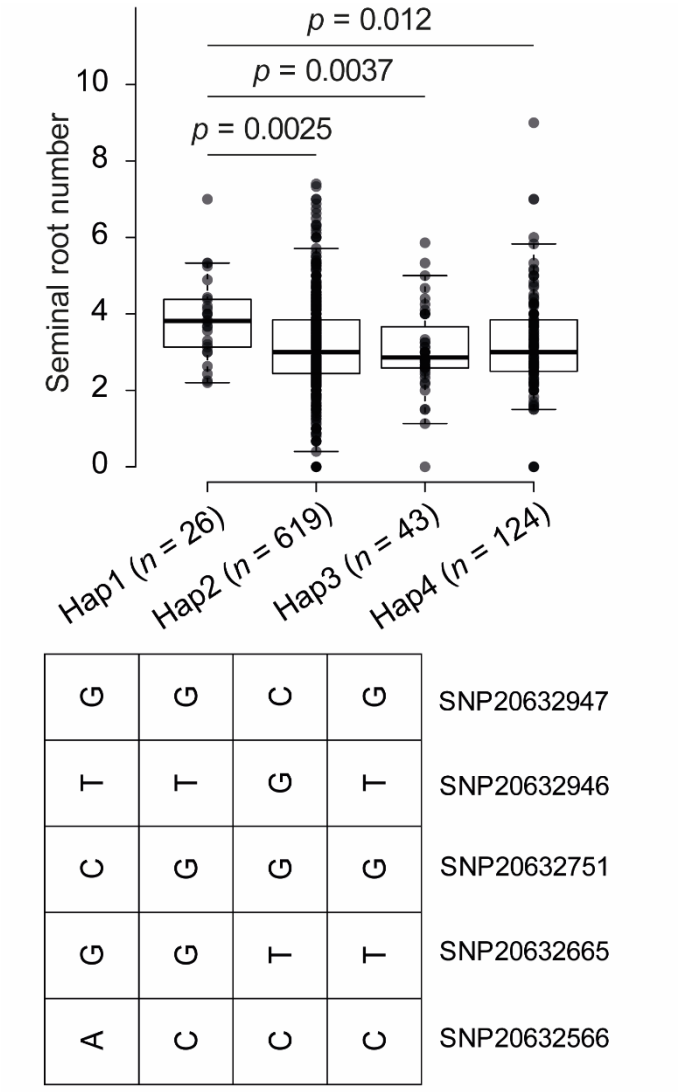
Supplementary Fig. 10. Transpiration rate of maize traditional varieties with different seminal root numbers. The exact numbers of traditional varieties are highlighted in the figure. Transpiration rates were determined gravimetrically by weighing the pots at three-hour intervals in three-week old maize plants. At the end of the measurement the leaf area of each plant was measured by scanning each leaf individually. Boxes span from the first to the third quartiles, center lines represent the median values and whiskers expand to 1.5× the interquartile range of the lower and upper quartiles. Significant differences were tested by one-way ANOVA with Tukey HSD at $p = 0.05$. ns, not significant.



Supplementary Fig. 11. Lignin density in roots of traditional varieties with different seminal root numbers. Roots were exposed to 160 mM NaCl solution to provoke osmotic stress in hydroponic plant culture. Root tips were harvested and enzymatically digested using cellulase and pectinase, followed by chemical analysis of lignin by thioacidolyisis. The whole experiment was done for 50 traditional varieties for each seminal root groups with 4 to 5 seminal roots (*n* = 20), 2 to 3 seminal roots (*n* = 13) and 0 to 1 seminal root (*n* = 17), respectively. Boxes span from the first to the third quartiles, center lines represent the median and whiskers extend to1.5× the interquartile range of the lower and upper quartiles. Data points outside of whiskers represent outliers. Statistical tests were performed by one-way ANOVA with Tukey HSD and labeled with exact values. ns, not significant.

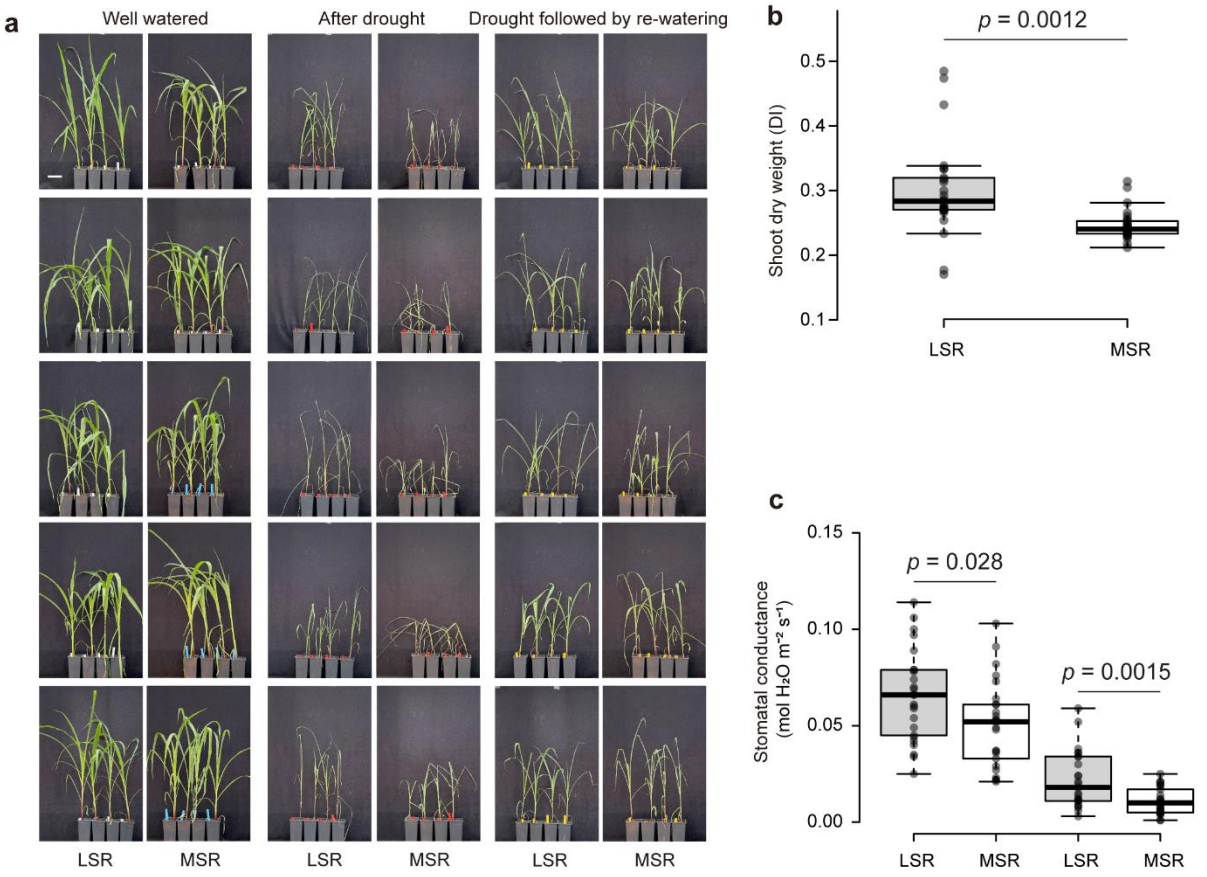


Supplementary Fig. 12. Root phenotyping of candidate mutants in the *BonnMu* stock collection. **a-e,** Root phenotypes of five candidate genes. Mutant lines are labeled with red frames. Seminal root number (**f**) and lateral root density (**g**) were illustrated by box plots. The exact number (n) of mutant or non-mutant seedlings are highlighted in the figure. Boxes span from the first to the third quartiles, center lines represent median values and whiskers show data lying within 1.5× interquartile range of the lower and upper quartiles. Data points at the ends of whiskers represent outliers. 160 candidate genes identified by GWAS were found in the *BonnMu* collection and germinated in paper rolls for 10 days before photographs were taken. For all figure panels, scale bar = 2 cm.

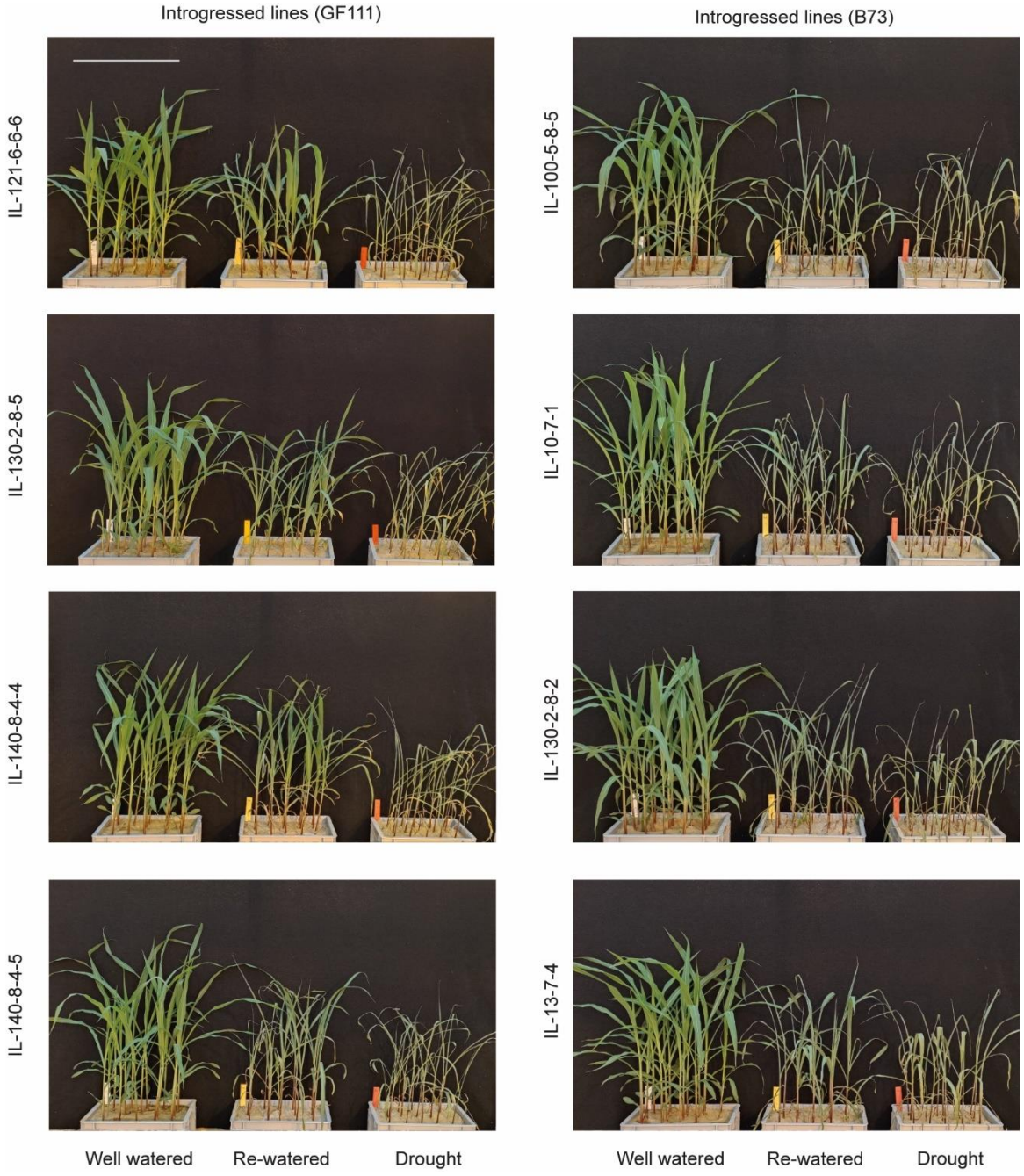


366 **Supplementary Fig. 13. Haplotype analysis of maize inbred lines.** The exact number (n) of inbred
367 lines are highlighted in the figure. Significant differences among haplotype groups are indicated by exact
368 values (one-way ANOVA, Tukey's HSD). Boxes span from the first to the third quartiles, center lines
369 represent median values and whiskers show data lying within 1.5× interquartile range of the lower and
370 upper quartiles. Data points at the ends of whiskers represent outliers.

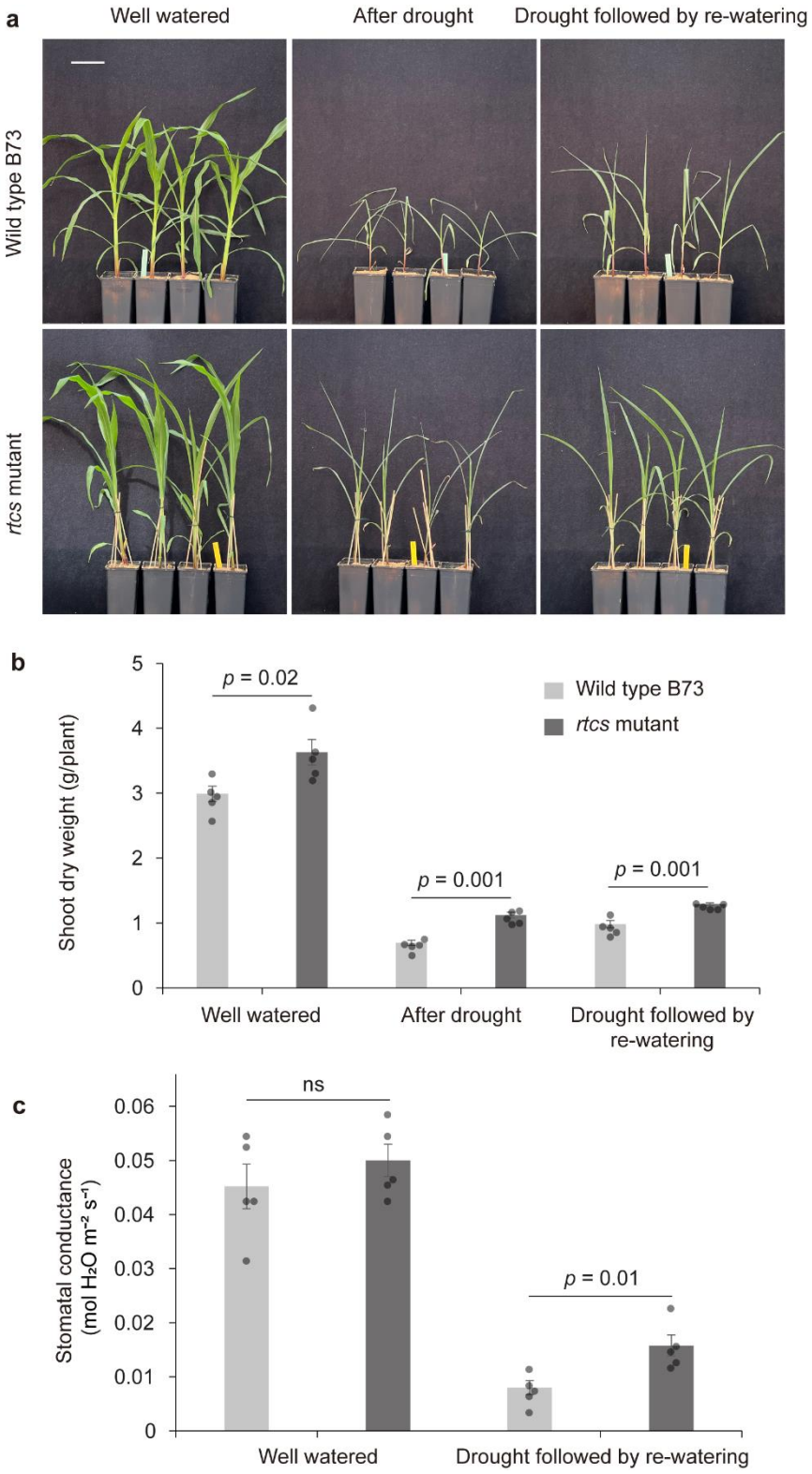
371



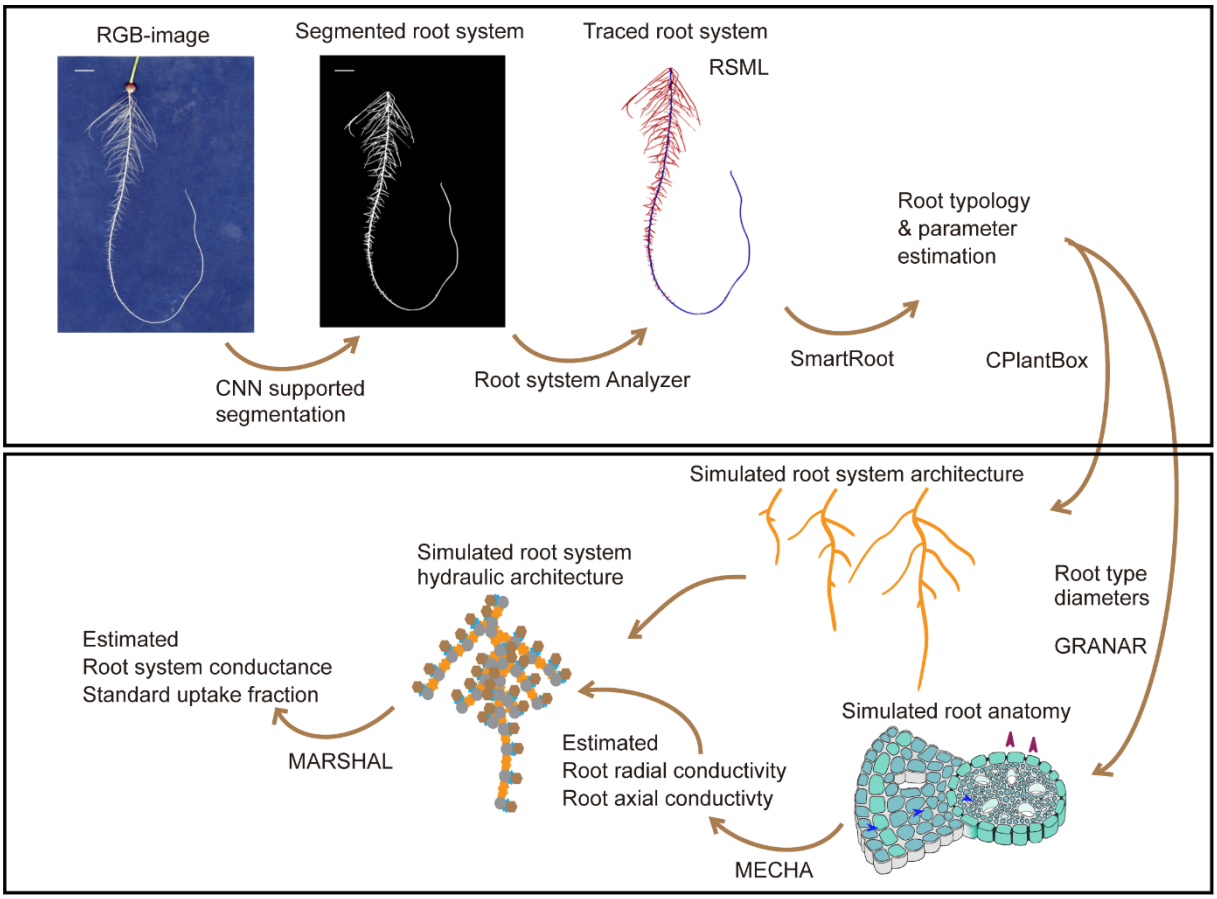
Supplementary Fig. 14. Drought experiment using different traditional maize varieties. (**a**) Representative seedlings of Northern Flint sourced varieties (n = 5) with less seminal roots (LSR) and other traditional varieties (n = 5) with more than 5 seminal roots (MSR) grown under well-watered, drought and drought followed by re-watering conditions. Shoot dry biomass (**b**) and stomatal conductance (**c**) of LSR and MSR varieties after re-watering. The Northern flint sourced varieties with LSR are PI 213730, PI 213733, PI 213735, PI 213748 and PI 217410 are displayed from top to bottom. The other dent or tropical sourced varieties with MSR are PI 213697, PI 221880, PI 222310, PI 222471 and PI 311235 accordingly. Boxes span from the first to the third quartiles, center lines represent median values and whiskers show data lying within 1.5× interquartile range of the lower and upper quartiles. Data points at the ends of whiskers represent outliers. The exact p values were determined by two-sided unpaired Student's t test. DI, drought index. For all figure panels, scale bar = 11 cm.



Supplementary Fig. 15. Representative seedlings of different introgression lines grown under well-watered, drought and drought followed by re-watering conditions. Four GF111^{ZmHb77} introgression lines and another four B73^{ZmHb77} introgression lines were displayed on the left and right, respectively. GF111 is an inbred line, which was developed by repeated selfing (eight cycles) and selection starting from the Northern Flint variety Gaspé Flint[31]. For all figure panels, scale bar = 30 cm.



Supplementary Fig. 16. Drought experiment using the *rtcs* mutant and its wild type siblings. (**a**) Representative seedlings of the *rtcs* mutant without seminal roots and wild type seedlings with three seminal roots grown under well-watered, drought and drought followed by re-watering conditions. Shoot dry biomass (**b**) and stomatal conductance (**c**) of these two genotypes grown under well-watered and drought followed by re-watering conditions. The exact *p* values were determined by two-sided unpaired Student's *t* test. Data are presented as mean values +/- SEM and *n* = 5 biologically independent replicates per treatment. ns, not significant. For all figure panels, scale bar = 11 cm.



Supplementary Fig. 17. A four-step structural-functional modelling pipeline to estimate hydraulic conductance in root systems. The scanned RGB images were automatically segmented by a convolutional neural network (CNN) model with *RootPainter*. Scale bars = 2 cm. Each root system was traced with *Root System Analyzer* (*RSA*) and the start and end point of the primary and each seminal root was manually defined. A RSML-file was exported from *RSA* and migrated to *SmartRoot* for manually labelling of seminal roots. The parameter set was used to generate five realizations of virtual root systems with the stochastic *CPlantBox* model. The anatomical traits were referred to the published articles and created with the *GRANAR* model. The radial hydraulic conductivity and axial hydraulic conductance (k_r and k_x) of these root anatomies were estimated with the model *MECHA*. From the *CPlantBox* root system architectures and their respective root hydraulic conductance, the whole root system conductance (K_{rs}) and the standard uptake fractions (SUF) were determined with the model *MARSHAL* for each virtual root system.

Supplementary references

84. Hetz, W., Hochholdinger, F., Schwall, M. & Feix, G. Isolation and characterization of rtcs, a maize mutant deficient in the formation of nodal roots. Plant J. 10, 845–857 (1996).
85. Ruiz Corral, J. A. et al. Climatic adaptation and ecological descriptors of 42 Mexican maize races. *Crop Sci.* **48**, 1502–1512 (2008).
86. Perales, H. & Golicher, D. Mapping the diversity of maize races in Mexico. *PLoS ONE* **9**, e114657 (2014).
87. Hufford, M. B., Martínez-Meyer, E., Gaut, B. S., Eguiarte, L. E. & Tenaillon, M. I. Inferences from the historical distribution of wild and domesticated maize provide ecological and evolutionary insight. *PLoS ONE* **7**, e47659 (2012).
88. Dell'Acqua, M. et al. Genetic properties of the MAGIC maize population: a new platform for high definition QTL mapping in *Zea mays*. *Genome Biol.* **16**, 167 (2015).
89. van Dusschoten D. et al. Quantitative 3D analysis of plant roots growing in soil using magnetic resonance imaging. *Plant Physiol.* **170**, 1176–1188 (2016).
90. Jung, B. A. & Weigel, M. Spin echo magnetic resonance imaging. *J. Magn. Reson. Imaging* **37**, 805–817 (2013).
91. Carminati, A. & Javaux, M. Soil rather than xylem vulnerability controls stomatal response to drought. *Trends Plant Sci.* **25**, 868–880 (2020).
92. Abdalla, M., Carminati, A., Cai, G., Javaux, M. & Ahmed, M. A. Stomatal closure of tomato under drought is driven by an increase in soil–root hydraulic resistance. *Plant Cell Environ.* **44**, 425–431 (2021).
93. Cai, G., Ahmed, M. A., Abdalla, M. & Carminati, A. Root hydraulic phenotypes impacting water uptake in drying soils. *Plant Cell Environ.* **45**, 650–663 (2022).
94. Brooks, R. H. & Corey, A. T. Properties of porous media affecting fluid flow. *J. Irrig. Drain. Eng.* **92**, 61–88 (1966).
95. Peters, A., Iden, S. C. & Durner, W. Revisiting the simplified evaporation method: Identification of hydraulic functions considering vapor, film and corner flow. *J. Hydrol.* **527**, 531–542 (2015).
96. De Jong Van Lier, Q., Van Dam, J. C., Metselaar, K., De Jong, R. & Duijnisveld, W. H. M. Macroscopic root water uptake distribution using a matric flux potential approach. *Vadose Zone J.* **7**, 1065 (2008).
97. Schröder, T., Javaux, M., Vanderborght, J., Körfgen, B. & Vereecken, H. Implementation of a microscopic soil–root hydraulic conductivity drop function in a three-dimensional soil–root architecture water transfer model. *Vadose Zone J.* **8**, 783–792 (2009).
98. Foster, C. E., Martin, T. M. & Pauly, M. Comprehensive compositional analysis of plant cell walls (Lignocellulosic biomass) part I: lignin *JoVE* 37, p.e1745 (2010).
99. Qi, Q. et al. Rapid, simplified microscale quantitative analysis of lignin H/G/S composition with GC–MS in glass ampules and glass capillaries. *MethodsX* **6**, 2592–2600 (2019).
100. Rolando, C., Monties, B. & Lapierre, C. Thioacidolysis. In *Methods in lignin chemistry*. Springer Berlin Heidelberg (1992).

101. Zeier, J. & Schreiber, L. Chemical composition of hypodermal and endodermal cell walls and xylem vessels isolated from *Clivia miniata* (identification of the biopolymers lignin and suberin). *Plant Physiol.* **113**, 1223–1231 (1997).
102. Delgado-Baquerizo, M. et al. Palaeoclimate explains a unique proportion of the global variation in soil bacterial communities. *Nat. Eco. Evo.* **1**, 1339–1347 (2017).