

# Genome-wide dissection of changes in maize root system architecture during modern breeding

Received: 6 February 2022

Accepted: 12 October 2022

Published online: 17 November 2022

 Check for updates

Wei Ren<sup>1</sup>, Longfei Zhao<sup>1</sup>, Jiaying Liang<sup>1</sup>, Lifeng Wang<sup>2</sup>, Limei Chen<sup>3</sup>, Pengcheng Li<sup>1,4</sup>, Zhigang Liu<sup>1</sup>, Xiaojie Li<sup>1</sup>, Zhihai Zhang<sup>1</sup>, Jieping Li<sup>1</sup>, Kunhui He<sup>1</sup>, Zheng Zhao<sup>1</sup>, Farhan Ali<sup>5</sup>, Guohua Mi<sup>1</sup>, Jianbing Yan<sup>6</sup>, Fusuo Zhang<sup>1</sup>, Fanjun Chen<sup>1,7</sup>✉, Lixing Yuan<sup>1,3</sup>✉ & Qingchun Pan<sup>1,7</sup>✉

Appropriate root system architecture (RSA) can improve maize yields in densely planted fields, but little is known about its genetic basis in maize. Here we performed root phenotyping of 14,301 field-grown plants from an association mapping panel to study the genetic architecture of maize RSA. A genome-wide association study identified 81 high-confidence RSA-associated candidate genes and revealed that 28 (24.3%) of known root-related genes were selected during maize domestication and improvement. We found that modern maize breeding has selected for a steeply angled root system. Favourable alleles related to steep root system angle have continuously accumulated over the course of modern breeding, and our data pinpoint the root-related genes that have been selected in different breeding eras. We confirm that two auxin-related genes, *ZmRSA3.1* and *ZmRSA3.2*, contribute to the regulation of root angle and depth in maize. Our genome-wide identification of RSA-associated genes provides new strategies and genetic resources for breeding maize suitable for high-density planting.

Maize (*Zea mays* L.) is grown worldwide and serves as a major pillar of global food security<sup>1</sup>. Breeding has contributed to about 50% of the increase in maize yield over the past century<sup>2</sup>, and breeders continue to select new maize varieties that can tolerate high-density planting systems, thereby increasing the yield per unit area<sup>2,3</sup>. Reductions in ear height, leaf angle, tassel branch number, tiller number and anthesis–silking interval have all enabled maize adaptation to high-density planting systems<sup>3,4</sup>. Genetic improvement of root system architecture (RSA) is also thought to contribute to increased crop yields in densely

planted fields<sup>5,6</sup>. High-density planting systems increase the competition between plants for nutrients and water, and aspects of RSA such as deep rooting can improve water and nitrogen use efficiency and reduce root-to-root competition<sup>7–9</sup>. Steeper maize root angles have also been found in high-density planting systems<sup>8,10</sup>. Therefore, genetic modifications of root traits provide an opportunity for crop improvement.

In recent years, researchers have attempted to increase crop yields through genetic improvement of root architecture. Two rice genes, *Deeper rooting1 (DROI)* and *Phosphorus-starvation tolerance 1 (PSTOL1)*,

<sup>1</sup>College of Resources and Environmental Sciences, National Academy of Agriculture Green Development, Key Laboratory of Plant-Soil Interactions, Ministry of Education, China Agricultural University, Beijing, China. <sup>2</sup>Cereal Crops Research Institute, Henan Academy of Agricultural Sciences, Zhengzhou, China. <sup>3</sup>Center for Crop Functional Genomics and Molecular Breeding, China Agricultural University, Beijing, China. <sup>4</sup>Key Laboratory of Plant Functional Genomics of the Ministry of Education, Yangzhou University, Yangzhou, China. <sup>5</sup>Cereal Crops Research Institute, Pirsabak, Nowshera, Pakistan. <sup>6</sup>National Key Laboratory of Crop Genetic Improvement, Huazhong Agricultural University, Wuhan, China. <sup>7</sup>Sanya Institute of China Agricultural University, Sanya, China. ✉e-mail: [caucfj@cau.edu.cn](mailto:caucfj@cau.edu.cn); [yuanlixing@cau.edu.cn](mailto:yuanlixing@cau.edu.cn); [panqingchun@cau.edu.cn](mailto:panqingchun@cau.edu.cn)

regulate root architecture and increase grain yield under stress conditions<sup>11,12</sup>. The *Enhanced Gravitropism2 (EGT2)* gene of barley was cloned and shown to regulate root growth angle in barley and wheat<sup>13</sup>. Many maize genetic studies have been performed on aspects of RSA such as root length, root number and root weight; fewer studies have focused on root angle and root depth, and most have involved only primary quantitative trait locus mapping<sup>14,15</sup>. A genome-wide association study (GWAS) of maize root traits, followed by functional validation, identified two genes, *ZmTIP1* and *ZmCIPK15*, that control root angle<sup>16,17</sup>. However, so far, only two studies related to GWAS analysis of maize root architecture have been conducted under field conditions<sup>18,19</sup>.

Root traits, like aboveground traits, have been selected during the processes of domestication and improvement<sup>20–24</sup>. The extent of root elongation has been shown to increase under N stress in maize, whereas the number of tillers decreases under N stress in teosinte, the ancestor of maize<sup>25,26</sup>. Compared with modern maize, teosinte shows greater variation in root aerenchyma<sup>27</sup>. Maize landraces also show greater variation in many root architectural traits, with the exception of nodal root number and branching<sup>21</sup>. For instance, landraces have larger xylem and stele areas, longer nodal roots, wider nodal systems and more seminal roots, whereas teosintes are smaller but have more nodal roots with greater branching<sup>21</sup>. Newly released maize varieties were found to have deeper root systems than older varieties<sup>28,29</sup>. The introduction of favourable chromosome segments from teosinte can improve maize tolerance to flooding stress<sup>30</sup>. Moreover, some genes under selection during maize domestication have been reported to participate in regulating aspects of root system development<sup>22,31</sup>. For instance, *teosinte branched1 (tb1)* regulates the number of crown roots and lateral roots<sup>22</sup>, and *teosinte glume architecture 1 (tga1)* regulates the number of brace root whorls<sup>31</sup>. The mining of elite root haplotypes during maize breeding can enhance tolerance to high planting densities and increase the absorption of water and nutrients<sup>16,32</sup>. A superior haplotype with variants in the *ZmTIP1* promoter region was associated with increased root hair length and, thus, improved drought resistance<sup>16</sup>. Therefore, the identification of root-related genes selected during maize domestication and improvement can be valuable for the development of new varieties with improved root traits.

The molecular basis of root development has been studied extensively in *Arabidopsis* and rice. Plant hormone (for example, auxin and cytokinin) signal transduction, hormone transport, multiple transcription factors, external nutrient levels and biosynthetic pathway genes all participate in root development<sup>33,34</sup>. Unlike the root system of *Arabidopsis*, the maize root system is complex and contains several parts with structurally and functionally distinct characteristics. Moreover, only a limited number of root developmental genes have been characterized in maize<sup>15</sup>. For instance, *RUM1* encodes an Aux/IAA protein that is a central regulator of auxin signalling, and *rum1* mutant is defective in maize lateral root production from the primary root<sup>35,36</sup>. Hence, more studies are needed to fully characterize the molecular events that regulate maize root development.

In this Article, we used an association mapping panel of 380 maize inbred lines to study the genetic architecture of adult maize root traits under field conditions. GWAS and root transcriptomic analyses revealed multiple high-priority candidate genes. We used transgenic lines to functionally validate the molecular mechanisms of two high-priority candidate genes that regulate maize RSA. We also identified maize root genes that have been under selection during domestication, improvement and modern maize breeding, and we identified favourable alleles. Our results provide valuable insight into the regulation of root architecture in maize and other gramineous crops.

## Results

### RSA traits differ among subpopulations and breeding eras

We phenotyped root architectural traits of 380 maize inbred lines in an association panel across four environments (2 locations × 2 years) in

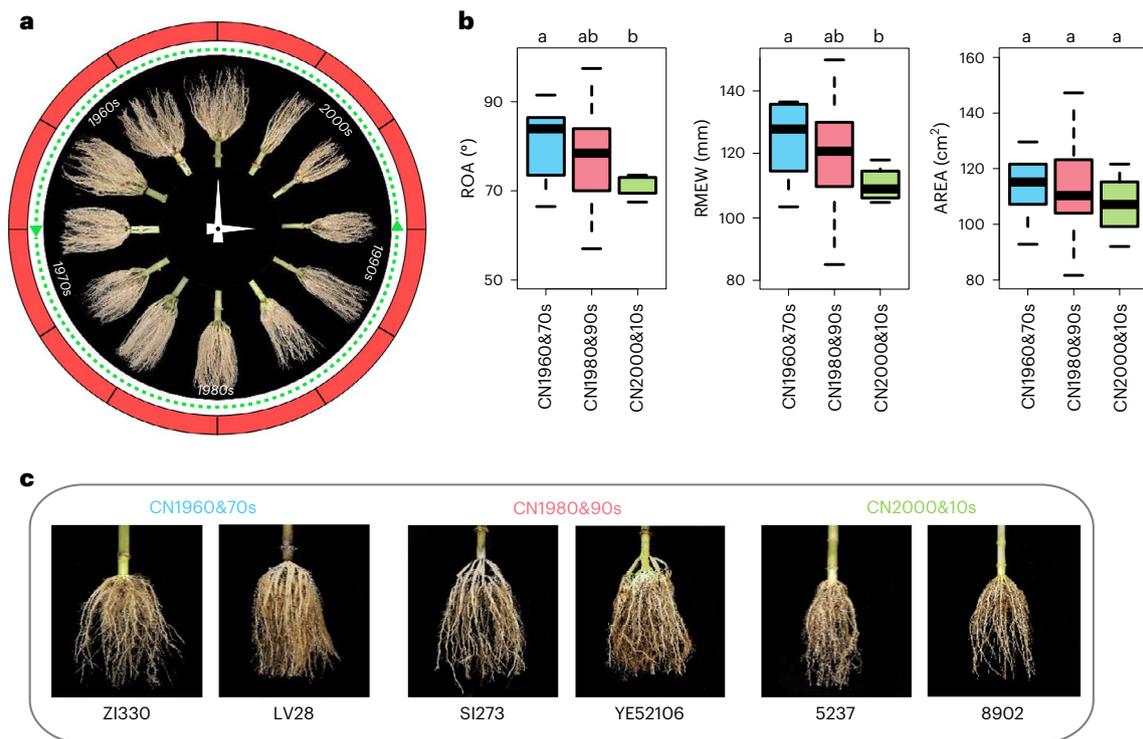
the field. Root system samples with a diameter of 40 cm and a depth of 30 cm were excavated from around the stem of approximately 15,000 individual plants, and eight root traits were measured on each sample: opening root angle (ROA), top root angle (RTA), bottom root angle (RBA), median width of the root system (RMEW), maximum width of the root system (RMAW), root width after correction (ROIW), projected root area (AREA) and area of the root convex hull (ACH) (Fig. 1a and Supplementary Table 1). The best linear unbiased prediction (BLUP) values for each trait across the four environments were estimated to represent the root phenotype of each line. The eight root traits were highly correlated (Supplementary Table 2, Extended Data Fig. 1 and Supplementary Note). They were grouped into three categories by principal component analysis (PCA) of the phenotypes: root angle (ROA, RTA and RBA), root width (RMEW, RMAW and ROIW) and root area (AREA and ACH) (Supplementary Fig. 1a,b, Extended Data Fig. 2 and Supplementary Note). Their broad-sense heritabilities ranged from 0.63 for AREA to 0.85 for RTA and were generally lower than those of aboveground agronomic traits<sup>37</sup> (Supplementary Table 3).

The inbred lines were divided into four subpopulations according to their genetic relationships in the association population<sup>38</sup> (Methods). Our phenotypic data showed that the stiff stalk subpopulation (SS) had a smaller root angle, width, and area than the non-stiff stalk (NSS), tropical and subtropical (TST) and mixed subpopulations ( $P < 0.05$ ; Extended Data Fig. 3 and Supplementary Note). The representative SS lines were B73 and GEMS66, and the representative lines of the other subpopulations were MO17, ZHENG58 and GEMS25. According to the release year of the inbred lines, we divided the Chinese inbred lines of this association panel into three subgroups<sup>4</sup> (Methods). Compared with earlier inbred lines, the newly released inbred lines had steeper root angles and narrower root widths ( $P < 0.05$ ; Fig. 1b,c and Supplementary Table 4). These results suggest that the root architecture of maize may have become steeper and narrower during the modern breeding process in China, and these root architecture features have been proposed to promote water and nitrogen absorption in maize and improve adaptability to high-density planting<sup>6,32</sup>.

### Identification and prioritization of RSA candidate genes

We used GWAS to explore the genetic basis of root architectural variation in the association panel, and we identified 371 single-nucleotide polymorphisms (SNPs) associated with the eight root traits (Fig. 2a, Supplementary Fig. 2, Supplementary Table 5 and Methods). The proportions of phenotypic variance explained by these SNPs ranged from 4.1% to 11.2% with a median value of 5.3%. By searching for candidate genes within 50 kb up- and downstream of the associated SNPs, we identified 795 GWAS candidate genes (Supplementary Table 6). These genes were involved mainly in molecular function pathways such as anatomical structure morphogenesis, transport, response to stimulus, signal transduction and cellular component organization (Supplementary Fig. 3). Several known root-related genes (for example, *ZmIAA4* (GRMZM2G159285), *ZmRTH1* (GRMZM2G099056), *ZmAUX1* (GRMZM2G127949), *ZmIAA10* (*ZmRSA3.1*, GRMZM2G138268) and *ZmEXPA5* (GRMZM2G361064)) were identified in our GWAS results within 50 kb of associated SNPs (Fig. 2a and Supplementary Table 7; Fisher's exact test,  $P < 0.01$ ). For instance, *ZmRTH1* (significant SNP located within the gene) was significantly associated with ROA (root angle category;  $P = 3.85 \times 10^{-5}$ ), RMEW (root width category;  $P = 7.61 \times 10^{-5}$ ) and ROIW (root width category;  $P = 2.39 \times 10^{-5}$ ) (Fig. 2b), and its mutant displays a defective root hair phenotype in maize<sup>39</sup>. *ZmAUX1* (2.85 kb from the closest significant SNP), which was associated with RBA (root angle category;  $P = 5.64 \times 10^{-5}$ ) and RMAW (root width category;  $P = 9.82 \times 10^{-5}$ ), encodes an auxin transporter-like protein and affects the geotropism of the maize root system<sup>40</sup> (Fig. 2c).

To more efficiently identify high-priority candidate genes, a gene co-expression network was constructed using root transcriptome data. Eighty-one of the 795 GWAS candidate genes were selected on the



**Fig. 1 | Changes in RSA over the course of Chinese maize breeding.** **a**, Variations in RSA among association populations in the GWAS panel. **b**, Values of ROA, RMEW and AREA in maize from different eras of Chinese breeding. In each box, the lower and upper boundaries represent the 25th and 75th percentiles, respectively. The middle horizontal line represents the median. The whiskers represent 1.5× the interquartile range. Different letters above the boxes indicate

significant differences at  $P < 0.05$  (one-way analysis of variance followed by Duncan's multiple-comparison test). Nine inbred lines were released during 1960–1979 (CN1960&70s,  $n = 9$ ), 80 during 1980–1999 (CN1980&90s,  $n = 80$ ) and 4 after 2000 (CN2000&10s,  $n = 4$ ). **c**, Representative inbred lines produced during different eras of Chinese maize breeding.

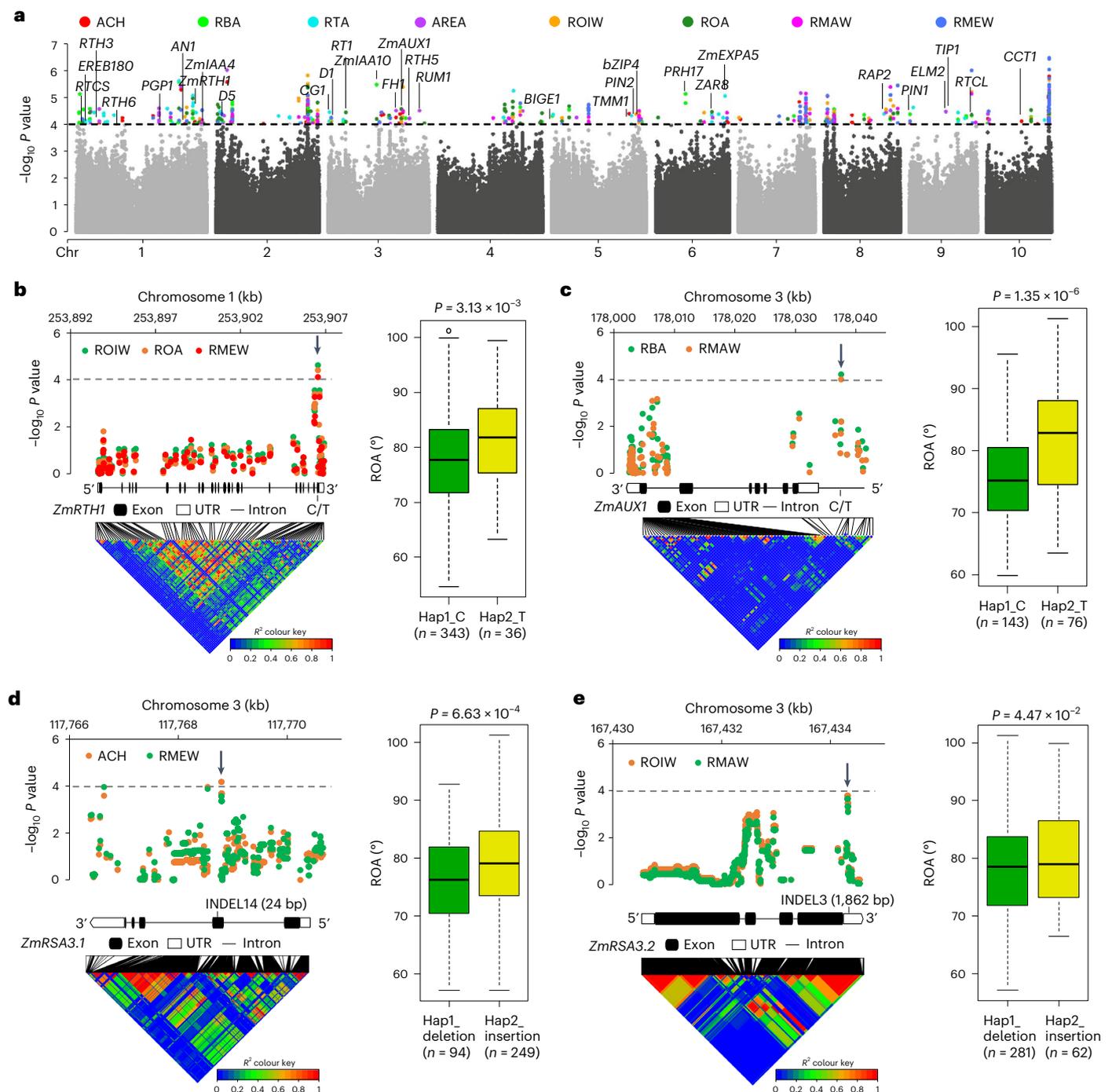
basis of their strong co-expression in the maize root and were defined as high-priority candidate genes (Supplementary Fig. 4 and Supplementary Table 8). The functions of these high-priority candidate genes included anatomical structure morphogenesis, cell differentiation and cell growth (Supplementary Fig. 5). Among the high-priority candidate genes, four (*ZmAUX1*, *ZmIAA4* (2.99 kb from the closest significant SNP), *ZmRSA3.1* (significant SNP located within the gene) and *ZmRSA3.2* (GRMZM2G044055; 3.06 kb from the closest significant SNP)) have previously been reported to be related to root development<sup>40–44</sup> (Fisher's exact test,  $P < 0.01$ ).

### Selective sweeps during modern maize breeding in China

To detect the indirect selection of root-related genes during modern maize breeding, we collected 172 maize improved inbred lines from a previous study<sup>4</sup>. The inbred lines released in 1960–1979, 1980–1999 and after 2000 were named CN1960&70s, CN1980&90s and CN2000&10s, respectively. A cross-population composite likelihood ratio (XP-CLR)<sup>45</sup> was used to detect putatively selected regions during modern Chinese breeding (selection in maize improved lines from three different breeding eras in China). The analysis of genome-wide selective sweeps identified 3,793 genes (Fig. 3a,b and Supplementary Table 9), and we identified 3,322 and 2,540 genes that were selected during domestication and improvement, respectively (Supplementary Fig. 6, Supplementary Tables 10 and 11, and Methods). Eighty-one, 93 and 59 of our GWAS candidate genes were selected during modern Chinese breeding, domestication and improvement, respectively (Fig. 3a,b and Supplementary Fig. 6). Furthermore, 71, 24 and 41 GWAS candidate genes were selected during the three breeding eras of CN1980&90s versus CN1960&70s, CN2000&10s versus CN1960&70s, and CN2000&10s versus CN1980&90s, respectively (Fig. 3a,c).

By analysing the whole-genome selection signals of three modern breeding eras in China, we found that five, three and seven known root-related genes were indirectly selected during CN1980&90s versus CN1960&70s, CN2000&10s versus CN1960&70s, and CN2000&10s versus CN1980&90s, respectively (Fig. 3b). For example, the well-studied maize gene *RTH1* (GRMZM2G099056) was indirectly selected between CN1980&90s and CN1960&70s. *RTH1* encodes a *sec3* homologue and regulates root hair elongation by mediating the polar exocytosis of secretory vesicles<sup>39</sup>. *RUM1* (GRMZM2G037368), which was indirectly selected between CN2000&10s and CN1960&70s, encodes a 269-amino-acid monocot-specific Aux/IAA protein that is required for the initiation of embryonic seminal roots and post-embryonic lateral roots from primary roots of maize<sup>36</sup>.

To detect indirect selection on root-related genes during maize domestication (from teosinte to maize landraces) and improvement (from maize landraces to improved inbred lines)<sup>46</sup>, we carried out selective sweep detection using 68 teosintes, 55 maize landraces and 172 maize improved lines<sup>47</sup> (Methods). Genes related to shoot-borne roots (brace and crown roots), which determine the skeleton of the whole root system (Supplementary Table 12), were mainly indirectly selected during the domestication process from teosinte to maize landraces<sup>46</sup>. During the process of improvement from maize landraces to improved inbred lines<sup>46</sup>, indirectly selected traits involved mainly the seminal roots and lateral roots, and the functions of related genes included response to external stimulus, cell communication, anatomical structure development and biosynthetic process (Supplementary Table 12). Interestingly, the root candidate gene *ZmRSA3.2* identified in this study was also indirectly selected during the improvement process (Supplementary Fig. 6b).



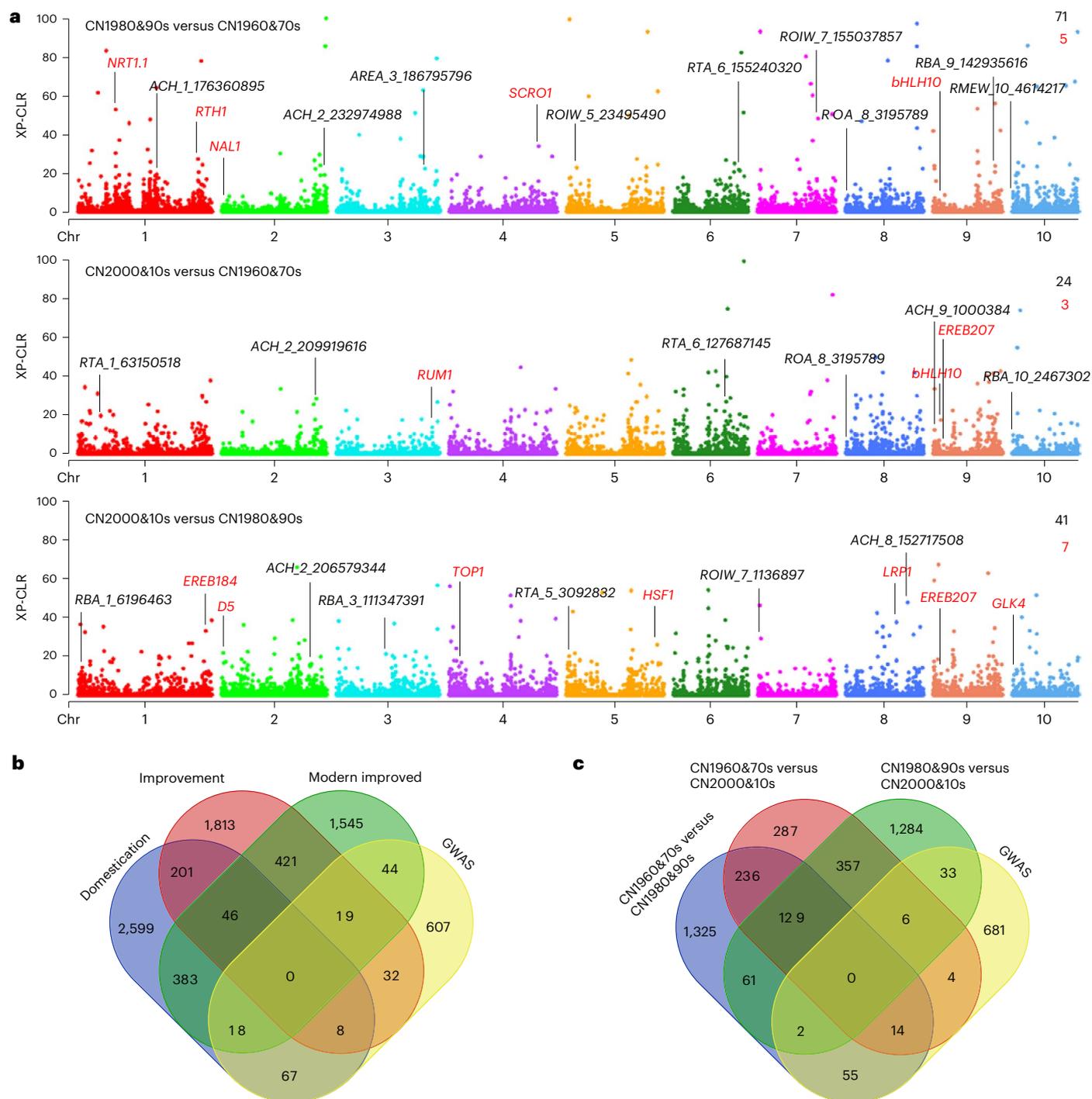
**Fig. 2 | GWAS identification of candidate genes for variation in maize root traits. a**, Manhattan plot of eight root traits analysed by GWAS. Several known root-related genes were identified in our GWAS results within 50 kb of associated SNPs (Fisher's exact test,  $P < 0.01$ ). The  $P$  value was calculated in TASSEL 5.0 software using the CMLM model. **b, c**, GWAS identification of *ZmRTH1* and *ZmAUX1* as candidate genes for RSA variation. Each set of plots includes a partial Manhattan plot (top left), the candidate gene structure, an LD heatmap (bottom left) and the ROA values of different haplotypes (box plot). *ZmRTH1* was co-located by ROA, ROIW and RMEW, and *ZmAUX1* was co-located by RBA and

RMAW. LD, linkage disequilibrium. **d, e**, *ZmRSA3.1* and *ZmRSA3.2* were identified on the basis of GWAS and co-expression analysis. *ZmRSA3.1* was co-located by the ACH and RMEW, and *ZmRSA3.2* was co-located by ROIW and RMAW. The known root-related genes marked in Fig. 2a are included in Supplementary Table 7. In each box, the lower and upper boundaries represent the 25th and 75th percentiles, respectively. The middle horizontal line represents the median. The whiskers represent  $1.5 \times$  the interquartile range. The  $P$  values in the box plots are derived from two-sided Student's  $t$ -tests, and 'n' in **b–e** denotes the number of inbred lines in each haplotype group.

### The role of favourable alleles during maize breeding

To study changes in favourable allele frequencies over the course of modern breeding, we used SNPs that were significantly related to root phenotypes in the GWAS. As modern breeding continued, newly released inbred lines tended to have a greater number of favourable

alleles associated with steep root angle and narrow root width, especially inbred lines released after 2000 (CN1960&70s versus 2000&10s and CN1980&90s versus 2000&10s,  $P < 0.05$ ; Fig. 4a and Supplementary Tables 4 and 13). Furthermore, on the basis of inbred lines released in different breeding eras, we summarized the numbers of favourable alleles



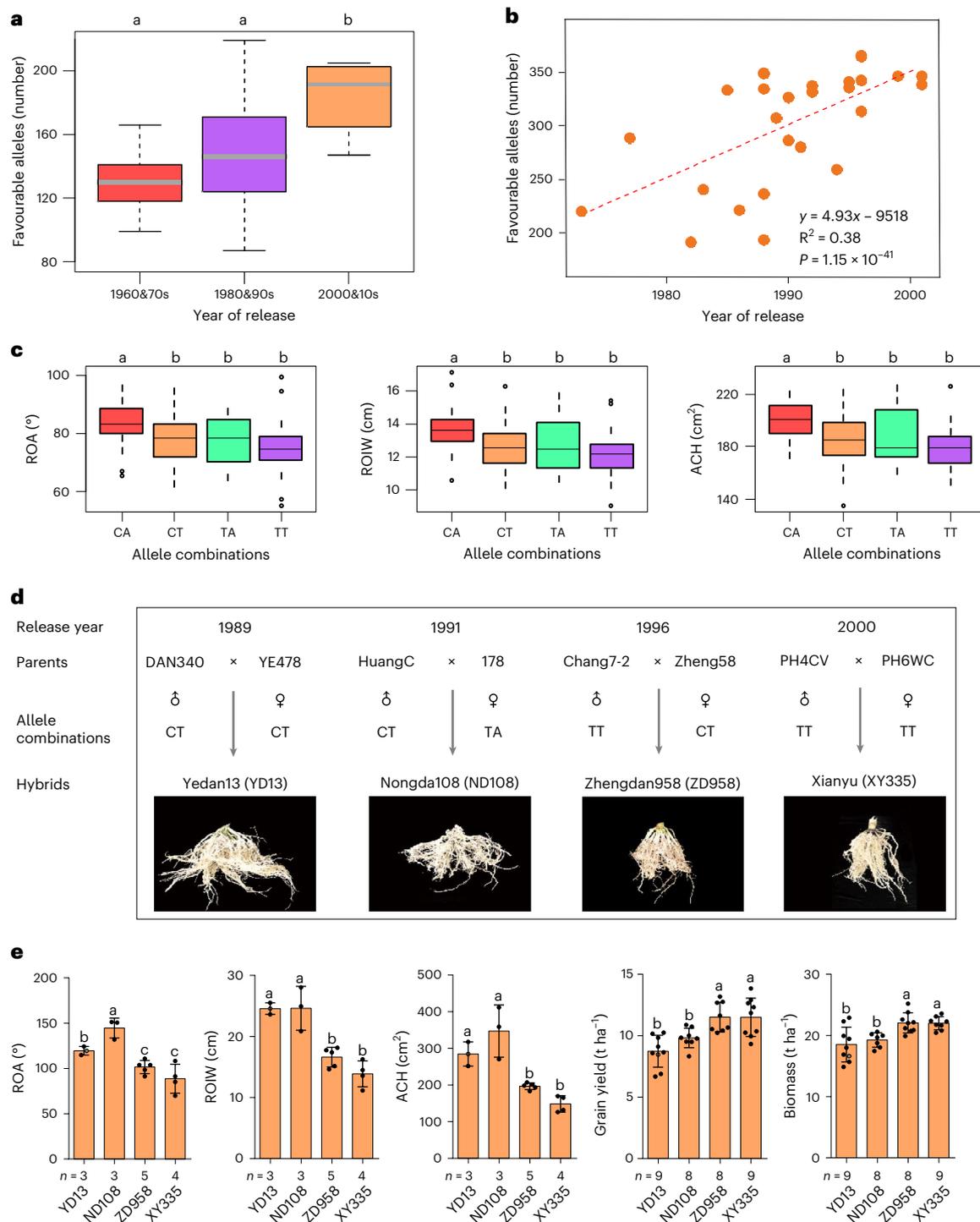
**Fig. 3 | Profiling of selective sweeps during maize domestication, improvement and modern breeding. a**, Genome-wide selection signals (XP-CLR scores) of different maize breeding eras in China. **b**, Numbers of genes shared among significant GWAS loci and selective sweeps associated with domestication, improvement and modern Chinese breeding. **c**, Numbers of genes shared among significant GWAS loci and selective sweeps associated

with different eras of Chinese breeding. The black number at the top left of each Manhattan plot is the number of GWAS candidate genes selected during each breeding era, and the red number is the number of GWAS candidate genes related to root development that have been functionally characterized. Inbred lines released during 1960–1979, 1980–1999 and after 2000 were named CN1960&70s, CN1980&90s and CN2000&10s, respectively.

in matching hybrids identified by GWAS (Supplementary Table 14). We found that the total number of favourable alleles carried by hybrids also increased significantly over the course of modern breeding ( $R^2 = 0.38$ ,  $P = 1.15 \times 10^{-41}$ ; Fig. 4b). To further verify this phenomenon, we used candidate gene association analysis to determine the association between root phenotype and polymorphic sites of known root-related genes and high-priority candidate genes identified in this study (Methods).

As maize breeding advanced, more favourable alleles accumulated in inbred lines and hybrids (Supplementary Fig. 7).

To further examine the above results in detail, we used eight inbred lines and four hybrids with different allele combinations at two newly identified root-related genes (*ZmRSA3.1* and *ZmRSA3.2*) for comparative analysis. On the basis of significant SNPs identified by GWAS, *ZmRSA3.1* (C/T) and *ZmRSA3.2* (A/T) could be divided into four allele



**Fig. 4 | The role of favourable alleles in the maize breeding process. a**, The number of favourable alleles related to steep root architecture in maize inbred lines from different breeding eras: 1960&70s ( $n = 9$ ), 1980&90s ( $n = 80$ ) and 2000&10s ( $n = 4$ ). **b**, The number of favourable alleles corresponding to maize hybrids ( $n = 27$ ) selected from inbred lines in our association panel. The x axis represents the release years of hybrids produced by crosses between inbred lines in the association panel. The y axis is the number of favourable alleles related to steep root architecture identified by GWAS of eight root traits. Linear regression was used for the analysis, and the coefficient of determination ( $R^2$ ) and the  $P$  value of the resulting equation are provided. **c**, The ROA, ROIW and ACH of four allele combinations corresponding to significant GWAS SNPs of *ZmRSA3.1* and *ZmRSA3.2*. The alleles corresponding to these two candidate genes are C/T and A/T, respectively. The T alleles is the favourable genotype at both loci. Four allele

combinations were possible at the two loci: CA ( $n = 46$ ), CT ( $n = 253$ ), TA ( $n = 9$ ) and TT ( $n = 57$ ). **d**, The parents and root architectures of four Chinese hybrids released in different years. The release years, parents, allele combinations and root images of the hybrids are shown. **e**, The ROA, ROIW, ACH, grain yield and biomass of four elite hybrids released in different years. The 'n' in ROA, ROIW and ACH indicates the number of biologically independent samples. The 'n' in grain yield and biomass indicates the number of independent experiments. Data are presented as mean  $\pm$  s.d. In the box plots in **a** and **c**, the lower and upper boundaries represent the 25th and 75th percentiles, respectively. The middle horizontal line represents the median. The whiskers represent 1.5 $\times$  the interquartile range. Different letters in **a**, **c** and **e** indicate significant differences at  $P < 0.05$  (one-way analysis of variance followed by Duncan's multiple-comparison test).

combinations (CA, CT, TA and TT). At both loci, the T allelic variant was associated with smaller root system angle, width and area (Fig. 4c). From the eight parents, the four elite hybrids Yedan13 (YD13), Nongda108 (ND108), Zhengdan958 (ZD958) and Xianyu335 (XY335) carried CT/CT, CT/TA, TT/CT and TT/TT allele combinations, respectively (Fig. 4d). The numbers of favourable alleles (T) carried by the four hybrids were thus 2, 2, 3 and 4 (Fig. 4d). YD13, ND108, ZD958 and XY335 were released in 1989, 1991, 1996 and 2000, respectively. In contrast to the older varieties (YD13 and ND108), the newly released varieties (ZD958 and XY335) had steeper and smaller root systems, which may have made them more tolerant of high planting density (Fig. 4d). Quantitative analysis of their crown roots showed that the ROA, ROIW and ACH of ZD958 and XY335 were significantly reduced by 27.96% ( $P < 0.001$ ), 38.01% ( $P < 0.001$ ) and 45.36% ( $P < 0.01$ ) relative to those of YD13 and ND108 (Fig. 4e). Moreover, the grain yield and aboveground biomass of ZD958 and XY335 were significantly higher by 23.99% ( $P < 0.001$ ) and 16.62% ( $P < 0.001$ ) relative to those of YD13 and ND108 (Fig. 4e and Supplementary Table 15).

### Functional validation of two RSA candidate genes

Two high-priority candidate genes, *ZmRSA3.1* and *ZmRSA3.2*, were selected for initial functional validation through the creation of overexpression lines. *ZmRSA3.1* encodes a member of the AUX/IAA protein family with 271 amino acids<sup>48</sup>. *ZmRSA3.2* is a gene with an unknown function in maize; its *Arabidopsis* homologue (AT3G25500.1) regulates actin filament and microtubule dynamics, which affect cell shape<sup>49</sup>. Three independent transformation lines were generated for each candidate gene, and their root phenotypes were compared with that of the wild type under field conditions. The ROA (root angle category) values of the *ZmRSA3.1* and *ZmRSA3.2* overexpression lines were lower than that of the wild type by 11.5–28.7% and 11.6–25.2% ( $P < 0.001$ ), respectively. The root depths of the *ZmRSA3.1* and *ZmRSA3.2* overexpression lines were greater than those of the wild type by 22.0–22.1% and 25.3–29.2% ( $P < 0.001$ ), respectively (Fig. 5a,b, Supplementary Fig. 9 and Supplementary Table 16). There was no significant difference in shoot biomass between wild-type and multiple overexpression lines (Supplementary Fig. 9i,j). These results were consistent with a role for *ZmRSA3.1* and *ZmRSA3.2* in the control of root architecture.

We re-sequenced alleles of *ZmRSA3.1* and *ZmRSA3.2* across the association panel, including the 5' untranslated region (UTR), coding sequence and 3' UTR (Methods). Sequence alignments revealed a 24 bp insertion/deletion (indel) in the second exon of *ZmRSA3.1* that was highly associated with root angle and depth ( $P = 6.63 \times 10^{-4}$ ) (Fig. 2d). Likewise, a large indel (1,862 bp) in the 3' UTR region of *ZmRSA3.2* was associated with root angle and depth ( $P = 4.47 \times 10^{-2}$ ) (Fig. 2e). We found that inbred lines carrying haplotype 1 (Hap1\_deletion; *ZmRSA3.1* and *ZmRSA3.2* deletion) showed smaller ROA values to those with haplotype 2 (Hap2\_insertion; *ZmRSA3.1* and *ZmRSA3.2* insertion) ( $P = 6.63 \times 10^{-4}$  and  $4.47 \times 10^{-2}$ , respectively; Fig. 2d,e). The expression levels of both *ZmRSA3.1* and *ZmRSA3.2* were negatively correlated with ROA ( $P = 1.57 \times 10^{-5}$  and  $3.27 \times 10^{-3}$ , respectively; Supplementary Fig. 10a,b). Moreover, the favourable haplotype (Hap2) of *ZmRSA3.1* and *ZmRSA3.2* has higher gene expression levels and a smaller ROA compared with Hap1 (Supplementary Fig. 10c,d). Collectively, these results show that the two indels of *ZmRSA3.1* and *ZmRSA3.2* may be promising functional sites for future root genetic improvement of maize.

To explore the physiological and molecular mechanisms by which *ZmRSA3.1* and *ZmRSA3.2* may influence root traits, we quantified the IAA concentrations in crown root tips and performed transcriptomic assays. Compared with the wild type, overexpression of *ZmRSA3.1* and *ZmRSA3.2* resulted in a significant increase in IAA concentration by 36.3% ( $P = 4.34 \times 10^{-4}$ ) and 63.9% ( $P = 1.10 \times 10^{-2}$ ), respectively (Fig. 5d). Subcellular localization of fusion proteins in maize protoplasts demonstrated that 35S::*ZmRSA3.1*-GFP and 35S::*ZmRSA3.2*-GFP localized mainly to the nucleus and plasma membrane, respectively (Fig.

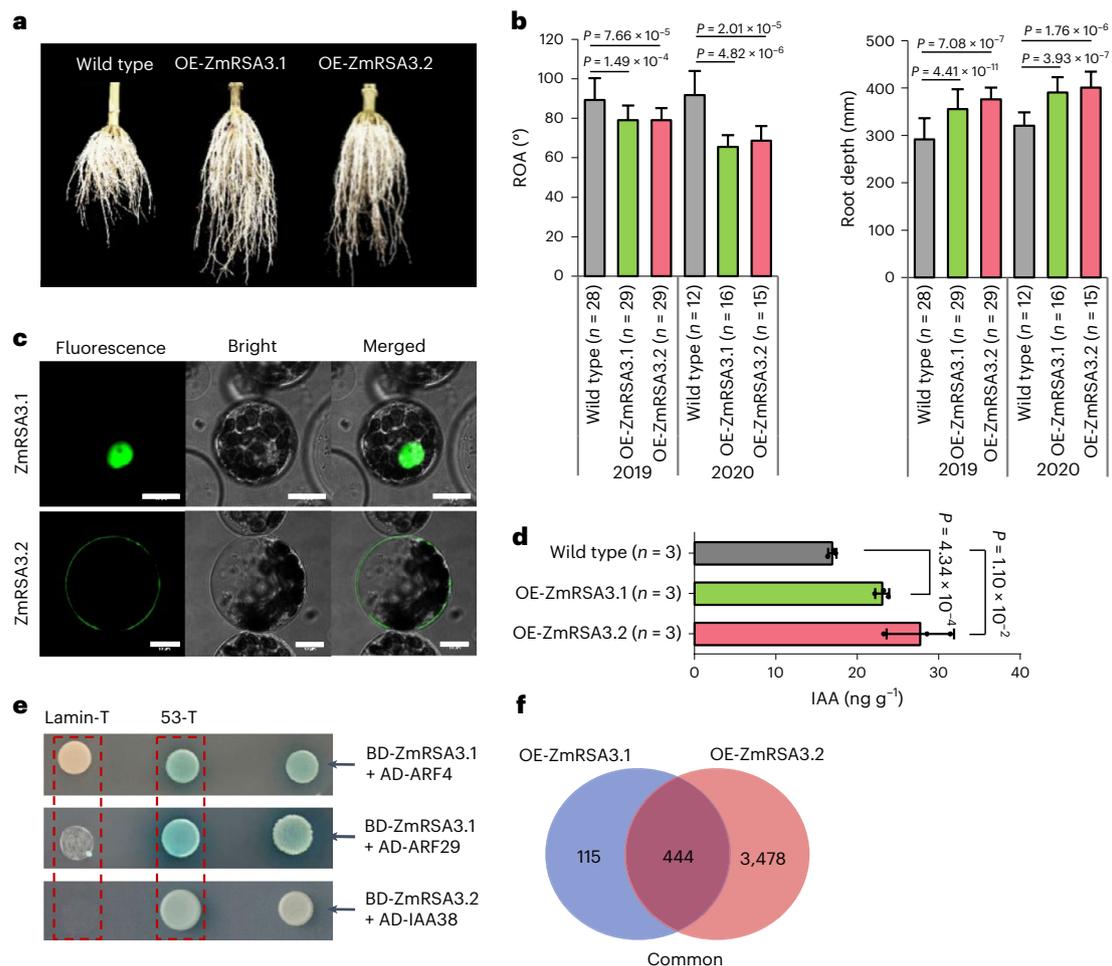
5c and Supplementary Figs. 11 and 12). Protein–protein interaction network analysis indicated that *ZmRSA3.1* may interact with two auxin response factors (*ZmARF4* (GRMZM2G034840) and *ZmARF29* (GRMZM2G086949)) and that *ZmRSA3.2* may interact with *ZmIAA38* (GRMZM2G035465). The predicted interacting proteins were then verified using a yeast two-hybrid assay (Fig. 5e). Transcriptome analysis of differentially expressed genes (DEGs) showed that overexpression of *ZmRSA3.1* and *ZmRSA3.2* caused significant changes in the expression of hormone-related genes involved in response to hormones, auxin stimulation, auxin-mediated pathways and nitrate transport (Supplementary Figs. 13 and 14 and Supplementary Tables 17–22). There were 444 common DEGs between the two comparisons of *OE-ZmRSA3.1* versus wild type and *OE-ZmRSA3.2* versus wild type (Fig. 5f). They included some known root-related genes such as *ZmARGOS8* (GRMZM2G354338), *ZmD3* (GRMZM2G093195) and *ZmD5* (GRMZM2G093603), which have been shown to regulate nodal root number via the auxin and gibberellin pathways<sup>50,51</sup> (Supplementary Tables 23 and 24 and Supplementary Fig. 15). These results suggest that the overexpression of *ZmRSA3.1* and *ZmRSA3.2* affects the maize auxin signalling pathway, which in turn causes changes in auxin distribution in the crown root and ultimately leads to changes in RSA (Supplementary Fig. 16d).

### Discussion

Many studies have highlighted the potential benefits of steeper RSA and have reported that a steeper RSA improves crop stress resistance and tolerance of high planting densities, thus contributing to sustainable crop production<sup>7,8,52,53</sup>. Root architecture has changed over the course of domestication and improvement, especially under modern high-density planting systems<sup>21,28,29</sup>. Nonetheless, there have been few studies on the genetic basis of maize RSA and the genes responsible for changes in RSA during maize domestication and breeding. Here we document changes in the RSA of maize during the modern breeding process and provide genetic evidence (Figs. 1 and 4a,b and Supplementary Fig. 7). GWAS and co-expression analysis were used to identify high-priority candidate genes associated with root phenotypes of maize (Fig. 2a and Supplementary Fig. 4). Moreover, the functions of the two candidate genes in regulating maize RSA were investigated using transgenic materials (Fig. 5a,b). Finally, we summarized and classified the known genes that control maize root development on the basis of their functions (Fig. 6).

Although many genetic studies of maize root traits have been performed, very few genes that influence maize RSA have been cloned<sup>14,15,17,54</sup>. In this study, we sampled 14,301 root systems of adult, field-grown maize plants from an association panel, and 371 SNPs and 795 GWAS candidate genes were identified. These candidates included some known root-related genes, such as *ZmIAA4*, *ZmRTH1*, *ZmAUX1* and *ZmEXPA5*, lending confidence to our GWAS results (Fig. 2a). Co-expression networks provide a powerful basis for prioritizing candidate causal genes from GWAS loci<sup>55</sup>. Here we combined GWAS data with co-expression analysis to identify 81 high-priority candidate genes (Supplementary Fig. 4), providing a wide range of high-quality candidate genes for maize RSA traits. We also collected and summarized all known root-related genes in maize whose functions affect crown roots, seminal roots and lateral roots (Supplementary Table 12). Our work therefore provides a comprehensive summary of root candidate genes in maize. Identifying favourable alleles of these genes is a worthwhile target for future maize root genetic improvement (Supplementary Tables 14 and 15).

In modern maize production, increased planting density has made an important contribution to continuous improvements in grain yield<sup>2</sup>. Compared with older maize hybrids, modern hybrids tend to show higher yield performance under high planting densities (Supplementary Fig. 17). Over the course of maize breeding, RSA has also been altered, and newly released hybrids tend to have longer total root lengths and a greater proportion of root length in the deeper soil



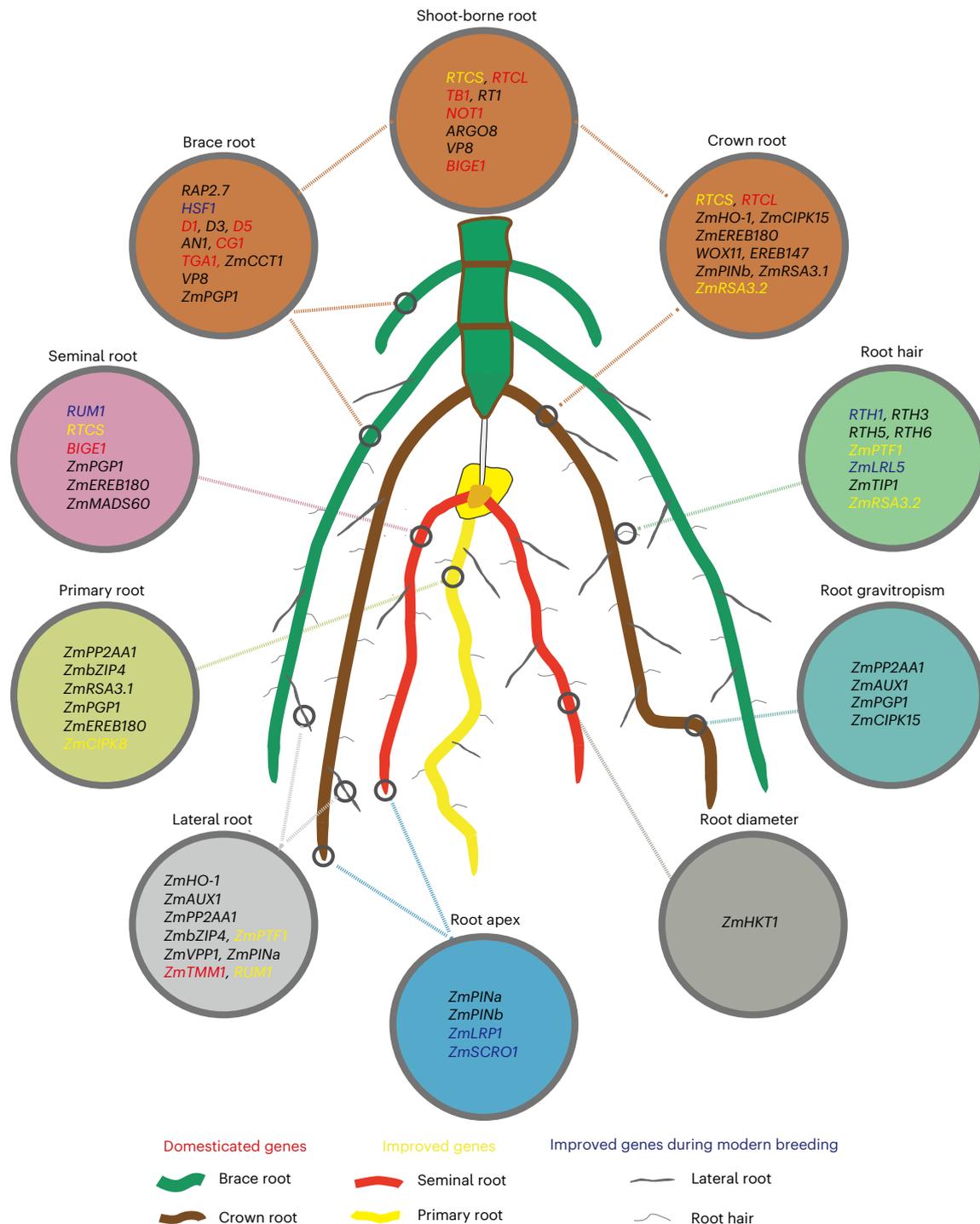
**Fig. 5 | Validation of two candidate genes associated with RSA. a, b**, Root architecture images (**a**) and statistics (**b**), showing the ROA and root depth (DEPTH) of wild-type and *ZmRSA3.1* and *ZmRSA3.2* overexpression plants. To obtain the entire maize root system, root excavation was performed at the silking stage, and the excavation was stopped when the maize root system was no longer visible to the naked eye. The ‘*n*’ indicates the number of biologically independent samples. Bars represent mean  $\pm$  s.d. **c**, Subcellular localization assays of *ZmRSA3.1* and *ZmRSA3.2*. Representative images are shown. Each experiment was repeated four times independently with similar results. Scale bars, 10  $\mu$ m. **d**, The IAA concentrations of wild-type and *ZmRSA3.1* and *ZmRSA3.2* overexpression plants. The ‘*n*’ indicates the number of independent experiments. Data are presented as mean  $\pm$  s.d. **e**, Interactions between *ZmRSA3.1* and *ZmARFs*

(*ZmARF4* and *ZmARF29*) and between *ZmRSA3.2* and *ZmIAA38* are indicated by yeast two-hybrid assays. The coding sequences of *ZmRSA3.1* and *ZmRSA3.2* were ligated into the pGBKT7 vector to generate the pGBKT7-*ZmRSA3.1* and pGBKT7-*ZmRSA3.2* bait vectors (BD-*ZmRSA3.1* and BD-*ZmRSA3.2*, respectively), and the coding sequences of *ZmARF4*, *ZmARF29* and *ZmIAA38* were ligated into the prey vector pGADT7 to generate pGADT7-*ARF4*, pGADT7-*ARF29* and pGADT7-*IAA38* (AD-*ARF4*, AD-*ARF29* and AD-*IAA38*, respectively). The vectors pGBKT7p53/pGADT7-T and pGBKT7-Lam/pGADT7-T were used as the positive and negative controls (Lamin-T and 53-T, respectively). **f**, Venn diagram of the numbers of DEGs derived from root transcriptome analysis of wild-type and *ZmRSA3.1* and *ZmRSA3.2* overexpression plants. *P* values of two-tailed Student’s *t*-tests are indicated in **b** and **d**.

profile<sup>28,29</sup> (Supplementary Fig. 18). However, this change has not previously been documented in released inbred lines, nor has there been genetic evidence to support it. In this study, we found that the root system of inbred lines gradually narrowed over the course of modern maize breeding (Fig. 1b,c and Supplementary Table 4), and this is consistent with previous published literature from China<sup>29</sup>. However, this conclusion is inconsistent with the findings of York et al.<sup>56</sup>. Using fifteen hybrids released by DuPont Pioneer between 1900 and 2000, they found that the most recent material had a shallower root system<sup>56</sup>. The reasons for the inconsistency may be that, on the one hand, the genetic background and breeding age of the materials used in different studies are different, and, on the other hand, the environments are different (for example, soil types and soil organic matter). On the basis of GWAS identification of significant SNPs associated with RSA, we found that favourable alleles (for steeper roots) have accumulated in modern inbred lines (Supplementary Table 14) and their resulting hybrids (Fig. 4). Information on these favourable alleles can provide

important guidance for molecular-marker-assisted breeding (Supplementary Fig. 8 and Supplementary Table 13). This result suggests that, with the accumulation of favourable root alleles, newly released hybrids may have steeper root architecture. Changes in root architecture may enable maize to thrive under high-density planting by enhancing water and nitrogen absorption and improving lodging resistance, thereby increasing grain yield<sup>7–9,52</sup>. Breeders have made improvements in aboveground morphological traits that promote the adaptation of maize to high-density planting systems; these include more upright leaves, lower ear heights and so forth<sup>2–4,57</sup>. These results suggest that, while breeders were improving aboveground agronomic traits, root traits may also have been indirectly selected during the process<sup>20–24</sup>.

Previous research has shown that the RSA of maize has undergone significant changes over the course of domestication and improvement<sup>20–24</sup>. For instance, two known domestication genes (*tb1* and *tga1*) have been reported to regulate maize root system development<sup>22,31</sup>. So far, there has been no systematic identification and comparison of



**Fig. 6 | Genes that have been reported to affect the development of maize roots.** For details on these genes, refer to Supplementary Table 7.

maize root genes selected during domestication and improvement. In this study, whole-genome selection scans revealed that 28 (24.3%) of known root genes were indirectly selected during the domestication and improvement of maize (Supplementary Table 12). By analysing the whole-genome selection signals of three modern breeding eras in China, we found that 13 (11.3%) known root-related genes were indirectly selected during modern breeding eras (Fig. 3b). When we performed Fisher's exact test on the overlaps shown in Fig. 3b,c, we did not obtain a significant result ( $P > 0.05$ ). This may be because these data were not based on the candidate gene identification method of co-expression with GWAS based on intervals. Instead, the method was

based on the XP-CLR values, where the top 20% of values, including selected candidate genes, are precisely located. In the domestication process, genes related to traits of shoot-borne roots (brace and crown roots) were indirectly selected; these roots determine the skeleton of the root system (Fig. 6 and Supplementary Table 12). The functions of these selected genes included signal transduction and responses to endogenous stimulus, external stimulus and abiotic stimulus (Supplementary Table 12). For example, *RTCL* (AC149818.2.FG009) binds to LOB domain (LBD) downstream promoters and acts as a transcription factor to control shoot-borne root elongation early in development<sup>58</sup>, and *BIGE1* (GRMZM2G148937) has been documented to regulate the

number of seminal and shoot-borne roots in maize<sup>59</sup>. In the process of maize improvement and modern maize breeding, genes related to traits of seminal roots and lateral roots were indirectly selected (Supplementary Table 12). For instance, *LRPI* (GRMZM2G077752) acts as a transcriptional activator in auxin signalling downstream of the *Aux/IAA* gene *RUMI*, and our results thus suggest that *LRPI* functions in maize root development via the RUMI-dependent auxin signalling pathway<sup>60</sup>. Root genes indirectly selected during domestication could be used for future genetic improvements such as lodging resistance. The root genes indirectly selected during improvement and modern breeding could be used to increase planting density and improve the absorption of nutrients and water. These findings provide new insight into changes in RSA during maize domestication and improvement.

Several known root-related genes in maize have been cloned using a mutant approach<sup>15</sup>, and their identities indicate that key elements of auxin signal transduction play an important regulatory role in maize root development. For example, Aux/IAA and LBD proteins contribute to the initiation of seed roots, shoot-borne roots and lateral roots<sup>58,61</sup>. *RTCS* and *RTCL* control the initiation and elongation of maize crown roots; they encode plant-specific LBD transcription factors and are both involved in the auxin signal transduction pathway<sup>61</sup>. Increased auxin in maize root tips causes the plant to allocate more carbohydrates to the roots, thereby improving root growth<sup>32,62</sup>. Here *ZmRSA3.1* identified in this study encodes an Aux/IAA protein. Similar to genes identified in previous studies<sup>32,62</sup>, the overexpression of *ZmRSA3.1* increased the accumulation of auxin in root tips, thereby promoting crown root elongation (Fig. 5a,b,d and Supplementary Fig. 9a–d). A yeast two-hybrid assay showed that, as an Aux/IAA protein, *ZmRSA3.1* could interact with the auxin response factors *ZmARF4* and *ZmARF29* (Fig. 5e). The rice homologue of *ZmRSA3.1*, *OsIAA30* (Os12g40890), has been reported to participate in auxin signalling and to control the expression of genes required for lateral root initiation in rice<sup>44</sup>. Hence, we speculate that, at high IAA concentrations, *ZmRSA3.1* may bind to the auxin receptor SCF<sup>TIR</sup> and be degraded through the ubiquitination pathway. The released auxin response factors *ZmARF4* and *ZmARF29* may then promote maize crown root development by regulating the transcription of auxin response genes.

Auxin plays a central role in plant growth and development<sup>63,64</sup>. The transport and signal transduction of auxin depend on actin organization<sup>65</sup>, and the actin cytoskeleton is required for polarized cell growth in plants<sup>66</sup>. For instance, the *Arabidopsis* gene *AtFHI* (AT3G25500), a formin homology protein, has been reported to participate in the signal transduction cascade that leads to re-arrangement of the actin cytoskeleton<sup>67</sup>, and it may therefore regulate cell morphogenesis and macroscopic organ development<sup>49</sup>. The rice gene *OsFHI* (Os01g67240) has been reported to play a significant role in root-hair elongation via the regulation of actin-dependent transport<sup>43</sup>. In this study, we identified the high-priority candidate gene *ZmRSA3.2*, which is a homologue of *AtFHI* and *OsFHI*. Overexpression of *ZmRSA3.2* caused the accumulation of auxin in the root tips and promoted elongation of the crown root (Fig. 5a,b,d and Supplementary Fig. 9e–h). Furthermore, *ZmRSA3.2* interacted with the auxin response factor *ZmARF38* (Fig. 5e), consistent with a previous study<sup>68</sup>. Therefore, we speculate that *ZmRSA3.2* may be regulated by *ZmARF38* and may in turn regulate the transport of auxin in root tips by mediating the re-arrangement of the actin cytoskeleton. The numerous DEGs shared between the *ZmRSA3.1* and *ZmRSA3.2* overexpression lines suggest that *ZmRSA3.1* and *ZmRSA3.2* may regulate maize root development through some common regulatory pathways, such as the regulation of plant hormone levels (Fig. 5f, Supplementary Fig. 14 and Supplementary Tables 23 and 24). In addition, similar to previous reports of genes that affect root angle<sup>12,69,70</sup>, root gravitropism was found to be stronger after overexpression of *ZmRSA3.1* and *ZmRSA3.2* (Supplementary Fig. 16a–c). These findings give new insight into the molecular regulation of maize root system development and provide new molecular targets for maize breeding. Nonetheless, more extensive

research on the detailed molecular mechanism by which *ZmRSA3.2* regulates maize crown root elongation will be required in the future.

## Methods

### Population materials and growth conditions

A maize (*Zea mays* L.) association panel of 380 genetically diverse inbred lines was prepared, including tropical, subtropical and temperate materials representing global maize diversity<sup>71</sup> (Supplementary Table 25). We divided the inbred lines into three groups (CN1960&70s, CN1980&90s and CN2000&10s) on the basis of their release year<sup>4</sup> (Supplementary Table 4). In addition, these inbred lines can be divided into four subgroups (SS, NSS, TST and mixed) according to their genetic relationship<sup>38</sup> (Supplementary Table 25). The 380 maize inbred lines were grown under field conditions in 2015 and 2016 at the Shangzhuang Experimental Station of China Agricultural University (40°06' N, 116°11' E) in Beijing (location 1) and in 2015 and 2017 at the Xinxiang Comprehensive Scientific Research Experimental Base of the Henan Academy of Agricultural Sciences (35°11' N, 113°80' E) in Henan Province (location 2). Each inbred line was planted in a single row at a density of 100,000 plants ha<sup>-1</sup>, and lines were replicated three times in a randomized block design. The distance between rows was 50 cm, and the inter-plant distance was 20 cm. We applied 200 kg ha<sup>-1</sup> nitrogen (urea), 120 kg ha<sup>-1</sup> phosphate (P<sub>2</sub>O<sub>5</sub>) and 75 kg ha<sup>-1</sup> potassium (K<sub>2</sub>SO<sub>4</sub>) each year.

### Phenotyping of RSA

At the silking stage, three to five adjacent plants with similar above-ground growth were selected from each replicate of each inbred line for root excavation, for up to 15 root systems per inbred line. The overall root architecture of maize is mainly determined by thick and hard nodal roots, and this root skeleton is seldom changed by washing the roots. On the basis of a previously described method, a soil cylinder with a diameter of 40 cm and a depth of 30 cm was dug from around the stem of each plant<sup>72</sup>. A standard shovel was used to perform the root excavation. The excavated roots were gently shaken to remove most of the attached soil. The roots were then soaked in soapy water, and the remaining soil was removed using a washing apparatus with adjustable water pressure. The root systems were then transferred to a studio with a stable light-emitting diode light, and two-dimensional images were obtained with a camera (ILCE-5100, Sony). All root images were stored in JPEG file format (Supplementary Fig. 19). Root system architectural parameters were quantified from the images using Root Estimator for Shovelomics Traits (REST, v1.0.1) and Digital Imaging of Root Traits (DIRT)<sup>73,74</sup>. Eight root traits were derived from the images (Supplementary Fig. 1a and Supplementary Table 1).

### Data organization and phenotype analysis

The BLUP for each trait in the four environments (2 locations × 2 years) was estimated with the 'lme4' package in R-3.6.1 (ref. <sup>75</sup>) using the following mixed linear model:

BLUP = (1|rep%in%env) + (1|env) + (1|gen) + (1|env:gen) with genotype (gen), environment (env), genotype × environment (env:gen) and replications (rep) as random effects. The BLUP value of each inbred line was used for basic statistical analysis of the phenotypes, correlation analysis, cluster analysis and PCA. Correlation among the eight root traits was estimated using Pearson correlations between pairs of traits with the 'cor' function in R. The formula used was:

$$\rho_{X,Y} = \frac{\text{cov}(X, Y)}{\sigma_X \times \sigma_Y}$$

where  $\text{cov}(X, Y)$  is the co-variance between two root traits, and  $\sigma_X$  and  $\sigma_Y$  represent the standard deviation (s.d.) of the two traits. The cluster analysis of eight root traits was performed using the 'hclust' function in R with the 'ward' method. The Euclidean distance between all pairs of

root traits was calculated with the ‘dist’ function in R, and the formula used was:

$$d_{ij} = \sqrt{\sum_{p=1}^P (x_{ip} - x_{jp})^2}$$

where  $i$  and  $j$  are traits selected from the eight root traits, and  $p$  is the number of variables. The ‘principal’ functions of the ‘psych’ R package were used to perform PCA (R version R-3.6.1).

The broad-sense heritability ( $h^2$ ) was calculated using the following formula:

$$h^2 = \sigma_G^2 / \left( \sigma_G^2 + \frac{\sigma_{GE}^2}{n} + \frac{\sigma_\varepsilon^2}{rn} \right)$$

where  $\sigma_G^2$  is the genotypic variance,  $\sigma_{GE}^2$  is the interaction variance between genotype and environment,  $\sigma_\varepsilon^2$  is the residual variance,  $r$  is the number of replications and  $n$  is the number of environments in the study.

### Genotyping and population genetic analyses

The genotypes of the 380 inbred lines in this study were obtained from four different genotyping platforms: RNA sequencing<sup>38,76</sup>, a 50K array derived from the maize high-density array (600K) (ref. <sup>77</sup>), the Illumina MaizeSNP50 Bead Chip<sup>78</sup> and genotyping-by-sequencing technology<sup>79</sup>. The genotypes from the four platforms were merged after strict quality controls for each dataset, and 1.25 million SNPs with minor allele frequency (MAF)  $\geq 5\%$  were used for further analysis<sup>80</sup>. The physical locations of the SNPs were based on the B73 RefGen\_v2 genome sequence (<https://www.maizegdb.org/>). The final merged genotype set, population structure and kinship were obtained from the Maizego website (<http://www.maizego.org/Resources.html>).

### GWAS of root traits and candidate gene annotations

Genome-wide association analysis was performed in TASSEL (v5.2.50) software using the compressed mixed linear mixed (CMLM) model with 1,253,814 SNP markers, population structure (Q) and relative kinship matrix (K) files, and BLUP root phenotypic values<sup>81,82</sup>. The threshold 4.15 corresponded to a Bonferroni correction of  $1/n$ , where  $n$  was the number of independent markers determined by PLINK (v1.90) (ref. <sup>83</sup>) (window size 200, step size 5,  $r^2 = 0.05$ ). Given the rigour of the mixed linear model, and to balance false positives and false negatives, we conservatively chose  $-\log_{10}P = 4.0$  as a moderate threshold for calling significant associations. The GWAS results were visualized by Manhattan plots generated with the ‘CMplot’ R package (<https://github.com/YinLiLin/R-CMplot>). On the basis of previous research<sup>38</sup>, GWAS candidate genes were obtained by searching for genes within 50 kb up- and downstream of each unique, significantly associated SNP base on the B73 RefGen\_v2 maize genome sequence. For gene functional annotation, the candidate genes identified from the association mapping analyses were submitted to MaizeGDB (<https://www.maizegdb.org/>), National Center for Biotechnology Information (NCBI) (<https://www.ncbi.nlm.nih.gov/>) and The *Arabidopsis* Information Resource (TAIR; <https://www.arabidopsis.org/>) to search for the best match gene annotations. To evaluate the association results of this study, 115 known genes underlying root-related traits are summarized from the published literature and from functional annotations of their homologous genes (Supplementary Table 7).

### Construction of a gene co-expression network for maize roots

A total of 35 maize inbred lines were randomly selected from the association panel and planted in the Shangzhuang Experimental Station at China Agricultural University in the summer of 2019 (Supplementary Table 26). Two to three plants from each line were selected at the silking stage, and two layers of brace roots below the soil surface were

collected and frozen in liquid nitrogen. Total RNA was extracted using TRIzol reagent (Takara) according to the manufacturer’s instructions. RNA concentration and purity were measured using a NanoDrop 2000 instrument (Thermo Fisher Scientific). RNA integrity was assessed using the RNA Nano 6000 Assay Kit and the Agilent Bioanalyzer 2100 system (Agilent Technologies). Sequencing libraries were generated using the NEBNext Ultra RNA Library Prep Kit for Illumina (NEB) following the manufacturer’s recommendations. Paired-end, stranded libraries were sequenced on an Illumina HiSeq 6000 system to obtain 150 bp reads. Clean reads were deposited into the Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) under project number PRJNA694491. HISAT2 (v2.1.0) (ref. <sup>84</sup>) was used for sequence alignment, and StringTie (v2.0) (ref. <sup>85</sup>) was used to assemble and quantify gene expression levels as fragments per kilobase per million mapped fragments (FPKM) values after the alignment analysis was completed. The raw FPKM table was imported into Camoco (v.0.6.3) (ref. <sup>55</sup>), a computational framework for integrating significant GWAS loci with gene co-expression networks, and passed through the quality control pipeline.

### Co-expression network analysis

Camoco<sup>55</sup> was used to prioritize causal RSA-related genes in maize. The GWAS locus information was derived from the genome-wide association analysis, and density metrics were used to perform the co-expression network analysis. The candidate window size was 50 kb, and the maximum number of flanking genes was 2. The candidate windows were calculated both upstream and downstream of the input SNPs. Finally, a score  $>0.1$  was used as a filtering parameter to generate a high-priority candidate gene set. Cytoscape (v3.9.0) (ref. <sup>86</sup>) was used to visualize the network and the high-priority candidate gene set. The Cytoscape plugin BiNGO (v3.0.3) (ref. <sup>87</sup>) was used to calculate over-represented Gene Ontology (GO) terms in the high-priority candidate genes and display them as a network of significant GO terms.

### Selective sweep detection for domestication and improvement

A composite likelihood approach (XP-CLR) was used to scan for genome-wide selective sweeps<sup>45</sup>. Sixty-eight teosintes<sup>47</sup> were used as a reference, and 55 maize landraces<sup>47</sup> were used as a query to identify potential domestication-related sweeps (teosinte versus maize landraces) (Supplementary Table 27). To detect possible improvement-related sweeps, 55 maize landraces were used as a reference, and 172 maize improved lines<sup>4</sup> were used as a query (maize landraces versus maize improved lines) (Supplementary Table 27). For detection of sweeps associated with the modern maize breeding process, 172 improved maize inbred lines released in different breeding eras were used. Whole-genome, high-throughput sequencing data for teosinte, landraces and maize improved lines were downloaded from the Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/sra>) using BioProject numbers PRJNA616247 and PRJNA609577. The short sequencing reads were mapped to the B73 reference genome using BWA-MEM (v0.7.17-r1188) with default parameters<sup>88</sup>. Format conversion and sorting of the mapping result were performed with SAMtools (v1.7) (ref. <sup>89</sup>). Bcftools (v1.8) (ref. <sup>90</sup>) was used to call the raw variants, merge all samples and filter variants with max-missing = 0.8. Selective sweep analysis was performed with a step of 10 kb and a sliding window of 10 kb, and the top 20% of features were identified as putatively selected features.

### Obtaining aligned sequences of high-priority candidate genes and known root-related genes

Contigs of high-priority candidate genes and known root-related genes were retrieved from a published genome article<sup>91</sup>. Each de novo assembly was mapped to the maize B73 reference genome 4.0 (AGPv4) using minimap2 (v2.17-r941) (ref. <sup>92</sup>) with the parameters ‘-c -x asm5-B5-O4,16-no-long-join -r 85 -N 50 -s 65 -z 200-mask-level 0.9-min-occ

200 -g 2500-score-N 2-cs'. The alignments were then filtered using Quast (v5.0.2) (ref.<sup>93</sup>), which kept the best alignment for each contig to minimize the variants. The filtered alignment results from Quast were converted to AXT format using the custom script 'coord2axt.sh' and converted to MAF format using 'axtToMaf' from the University of California, Santa Cruz bioinformatics utilities (<http://genome.ucsc.edu>). The multiple alignments based on AGPv4 regions and the consensus sequences were retrieved using the custom pipeline 'PANZ\_aln\_extractor'. In general, to obtain multiple alignments based on AGPv4 regions, we subsetted each individual's MAF file according to the query AGPv4 region using `maf_extract_ranges_indexed.py` from `bx-python`. Each individual's sequences were then retrieved from the alignment blocks, and multiple alignment was performed using sequences from all individuals in the association mapping panel as input in MAFFT (v7.427) (ref.<sup>94</sup>) with the options '-op 4-ep 0-retree 1' to prevent gap opening within previously aligned blocks in each MAF file. The consensus sequence for each individual was then acquired by retaining the most frequently counted base for each individual.

### Analysis of favourable alleles by GWAS

Significant SNP sets generated by GWAS, known root-related genes and high-priority candidate genes identified from co-expression analysis were used to investigate changes in the frequency of favourable alleles during Chinese maize breeding. The candidate association analysis of known root-related genes and high-priority candidate genes was performed in TASSEL (v5.2.50) using a general linear mixed model that included population structure (Q) and BLUP root phenotypic value<sup>81,82</sup>. The allele types associated with steeper root growth angle were deemed to be the favourable alleles. The numbers of favourable alleles carried by Chinese hybrids selected from the inbred lines in our association panel were calculated as the sum of the favourable alleles of the two corresponding parents.

### Root phenotypic and yield data of maize hybrids from different breeding eras

To further investigate the changes in favourable alleles over the course of maize breeding, we compared the total number of favourable alleles from the two parents of maize hybrids from different breeding eras. Four maize hybrids (Yedan13, Nongda108, Zhengdan958 and Xianyu335) representing different years were used as an example to further verify the effects of favourable alleles on RSA, aboveground biomass and grain yield. Root images of the four hybrids were obtained from our previous unpublished work. REST (v1.0.1) (ref.<sup>74</sup>) was used for quantitative analysis of RSA to obtain the values of the ROA, ROIW and ACH. The grain yield and the aboveground biomass of the four hybrids were obtained from a previous study<sup>29</sup>.

### Construction of maize transgenic overexpression lines

To create the overexpression constructs, the coding sequences of *ZmRSA3.1* (GRMZM2G138268) and *ZmRSA3.2* (GRMZM2G044055) were amplified from the cDNA library of B73 and inserted into the pBCXUN vector under the control of the maize Ubiquitin 1 promoter (Ubi), forming the *pUbi::ZmRSA3.1* and *pUbi::ZmRSA3.2* vectors. The two constructs were transformed into *Agrobacterium tumefaciens* strain EHA105 by the freeze-thaw method and then into young maize embryos of inbred line B73\_329 by *Agrobacterium*-mediated transformation<sup>95</sup>. All the overexpression lines were cultivated at the Center for Crop Functional Genomics and Molecular Breeding of China Agricultural University.

### RSA phenotyping of transgenic overexpression lines under field conditions

In the summer of 2019 and 2020, transgene-negative plants (B73\_329) and three transgenic overexpression lines of each gene were planted at a density of 80,000 plants ha<sup>-1</sup> with three biological replicates at the

Shangzhuang Experimental Station of China Agricultural University (Supplementary Table 16). The distance between rows was 50 cm, and the inter-plant distance was 25 cm. The fertilization levels and field management were consistent with those used during the root phenotype survey of the GWAS panel. To obtain the entire maize root system, root excavation was performed at the silking stage, and the excavation was stopped when the maize root system was no longer visible to the naked eye. REST (v1.0.1) (ref.<sup>74</sup>) and DIRT<sup>73</sup> were used to quantify ROA and the depth of the root skeleton (DEPTH). Plant height and aboveground biomass were also measured at the silking stage.

### Subcellular localization assay

The coding regions of *ZmRSA3.1* and *ZmRSA3.2* were amplified from the complementary DNA library of inbred line B73 and cloned into the pCAMBIA1302 vector, which contains a green fluorescent protein at the C terminus. The restriction endonuclease NcoI was used to digest the pCAMBIA vector. The In-Fusion HD Cloning Kit (TaKaRa) was used to clone *ZmRSA3.1* and *ZmRSA3.2* into the pCAMBIA1302 vector by the In-Fusion reaction to generate pCAMBIA-*ZmRSA3.1*-mGFP5 and pCAMBIA-*ZmRSA3.2*-mGFP5. The GFP fusion proteins were expressed in maize mesophyll protoplasts under the control of the CaMV 35 S promoter, and images were collected with a Nikon C2-ER confocal microscope (Nikon). Images of the empty vector control are provided in Supplementary Figs. 11 and 12. Sequences of primers are listed in Supplementary Table 28.

### Yeast two-hybrid assay

The yeast two-hybrid assays were performed using the yeast strain Y2H Gold. The coding sequences of GRMZM2G138268\_P01 (*ZmRSA3.1*), GRMZM2G044055\_P01 (*ZmRSA3.2*), GRMZM2G086949\_P01 (*ZmARF29*), GRMZM2G034840\_P02 (*ZmARF4*) and GRMZM2G035465\_P03 (*ZmIAA38*) were PCR amplified from the cDNA library of inbred line B73. The coding sequences of *ZmRSA3.1* and *ZmRSA3.2* were ligated into the pGBKT7 vector to generate pGBKT7-*ZmRSA3.1* and pGBKT7-*ZmRSA3.2* as bait vectors (BD-*ZmRSA3.1* and BD-*ZmRSA3.2*, respectively), and the coding sequences of *ZmARF29*, *ZmARF4* and *ZmIAA38* were ligated into the prey vector pGADT7 to generate pGADT7-ARF4, pGADT7-ARF29 and pGADT7-IAA38 (AD-ARF4, AD-ARF29 and AD-IAA38, respectively). The pGBKT7-p53/pGADT7-T and pGBKT7-Lam/pGADT7-T vectors were transferred to Y2H Gold competent cells and used as the positive and negative controls (53-T and Lamin-T, respectively). Interactions were determined by co-transforming the prey and bait vectors into yeast Y2H Gold and growing the yeast colonies on SD-His-Leu-Trp+X- $\alpha$ -Gal plates. Sequences of primers are listed in Supplementary Table 28.

### Quantification of IAA concentrations

Wild-type maize and the *ZmRSA3.1* and *ZmRSA3.2* overexpression lines were planted under field conditions, and three plants of each genotype were selected at the silking stage. The root tips of the first layer of crown roots below the soil surface were sampled and ground in liquid nitrogen. The root samples were then used for IAA quantification and transcriptome sequencing. For IAA quantification, IAA was extracted and purified as described previously<sup>96</sup>. IAA was quantified on an ultraperformance liquid chromatography system (Waters) with a liquid chromatograph-tandem mass spectrometer coupled to a 6500 Q-Trap system (AB SCIEX) following a previously reported procedure<sup>97</sup>.

### Analysis of DEGs between transgenic lines and wild type

Collection of root samples was carried out during the 2019 field trial to identify the root phenotype of the overexpression lines. The root tips of the first layer of crown roots below the soil surface were sampled for transcriptome sequencing. Three biological replicates were used for the wild-type and each overexpression line, and three plants from each biological replicate were selected for root sampling and pooling.

In total, nine libraries (three maize genotypes (WT, *OE-ZmRSA3.1* and *OE-ZmRSA3.2*) with three biological replicates per genotype) were sequenced. The methods used for root transcriptome sequencing, sequence alignment and transcript assembly were the same as those described for the construction of the co-expression network above. DEGs were identified with DESeq2 (ref. <sup>98</sup>) using  $|\log_2(\text{fold change})| \geq 1$  and a false discovery rate (Benjamini–Hochberg adjusted *P*-value)  $\leq 0.001$ . GO enrichment analysis of the DEGs was performed using the R package topGO (<https://bioconductor.org/packages/release/bioc/html/topGO.html>). The GO enrichment network was visualized using Cytoscape (v3.9.0) (ref. <sup>86</sup>). Clean reads were deposited into the Sequence Read Archive under BioProject number [PRJNA693427](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA693427).

### Reporting summary

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

### Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information. The genotype set, population structure and kinship data can be downloaded from the Maizego website (<http://www.maizego.org/Resources.html>). All root phenotype data for the 380 inbred maize lines are included in Supplementary Table 25. The RNA-sequencing reads used to construct the co-expression network and the root transcriptome sequencing reads were deposited in the NCBI Sequence Read Archive (<https://www.ncbi.nlm.nih.gov/>) under accession codes [PRJNA694491](https://www.ncbi.nlm.nih.gov/PRJNA694491) and [PRJNA693427](https://www.ncbi.nlm.nih.gov/PRJNA693427), respectively. Source data are provided with this paper.

### Code availability

All scripts for GWAS, co-expression network analysis, selective sweep detection for domestication and improvement, and obtaining aligned sequences of high-priority candidate genes and known root-related genes (<https://doi.org/10.5281/zenodo.7112683>) are available on Zenodo.

### References

- Tilman, D., Balzer, C., Hill, J. & Befort, B. L. Global food demand and the sustainable intensification of agriculture. *Proc. Natl Acad. Sci. USA* **108**, 20260–20264 (2011).
- Duvick, D. N. The contribution of breeding to yield advances in maize (*Zea mays* L.). *Adv. Agron.* **86**, 83–145 (2005).
- Tian, J. et al. Teosinte ligule allele narrows plant architecture and enhances high-density maize yields. *Science* **365**, 658–664 (2019).
- Wang, B. et al. Genome-wide selection and genetic improvement during modern maize breeding. *Nat. Genet.* **52**, 565–571 (2020).
- Hochholdinger, F. Untapping root system architecture for crop improvement. *J. Exp. Bot.* **67**, 4431–4433 (2016).
- Lynch, J. P. Root phenotypes for improved nutrient capture: an underexploited opportunity for global agriculture. *N. Phytol.* **223**, 548–564 (2019).
- Lynch, J. P. Steep, cheap and deep: an ideotype to optimize water and N acquisition by maize root systems. *Ann. Bot.* **112**, 347–357 (2013).
- Mi, G., Chen, F., Yuan, L. & Zhang, F. Ideotype root system architecture for maize to achieve high yield and resource use efficiency in intensive cropping systems. *Adv. Agron.* **139**, 73–97 (2016).
- Thorup-Kristensen, K. et al. Digging deeper for agricultural resources, the value of deep rooting. *Trends Plant Sci.* **25**, 406–417 (2020).
- Shao, H. et al. Genotypic difference in the plasticity of root system architecture of field-grown maize in response to plant density. *Plant Soil* **439**, 201–217 (2019).
- Gamuyao, R. et al. The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. *Nature* **488**, 535–539 (2012).
- Uga, Y. et al. Control of root system architecture by *DEEPER ROOTING 1* increases rice yield under drought conditions. *Nat. Genet.* **45**, 1097–1102 (2013).
- Kirschner, G. K. et al. *ENHANCED GRAVITROPISM 2* encodes a STERILE ALPHA MOTIF-containing protein that controls root growth angle in barley and wheat. *Proc. Natl Acad. Sci. USA* **118**, e2101526118 (2021).
- Bray, A. L. & Topp, C. N. The quantitative genetic control of root architecture in maize. *Plant Cell Physiol.* **59**, 1919–1930 (2018).
- Hochholdinger, F., Yu, P. & Marcon, C. Genetic control of root system development in maize. *Trends Plant Sci.* **23**, 79–88 (2018).
- Zhang, X. M. et al. Genetic variation in *ZmTIP1* contributes to root hair elongation and drought tolerance in maize. *Plant Biotechnol. J.* **18**, 1271–1283 (2020).
- Schneider, H. M. et al. Root angle in maize influences nitrogen capture and is regulated by calcineurin B-like protein (CBL)-interacting serine/threonine-protein kinase 15 (*ZmCIPK15*). *Plant Cell Environ.* **45**, 837–853 (2022).
- Schneider, H. M. et al. Genetic control of root architectural plasticity in maize. *J. Exp. Bot.* **71**, 3185–3197 (2020).
- Zheng, Z. et al. Shared genetic control of root system architecture between *Zea mays* and *Sorghum bicolor*. *Plant Physiol.* **182**, 977–991 (2020).
- Chen, Z. et al. Plasticity of root anatomy during domestication of a maize-teosinte derived population. *J. Exp. Bot.* **73**, 139–153 (2022).
- Burton, A. L., Brown, K. M. & Lynch, J. P. Phenotypic diversity of root anatomical and architectural traits in *Zea* species. *Crop Sci.* **53**, 1042–1055 (2013).
- Gaudin, A. C. M., McClymont, S. A., Soliman, S. S. M. & Raizada, M. N. The effect of altered dosage of a mutant allele of *Teosinte branched 1 (tb1-ref)* on the root system of modern maize. *BMC Genet.* **15**, 23 (2014).
- Perkins, A. C. & Lynch, J. P. Increased seminal root number associated with domestication improves nitrogen and phosphorus acquisition in maize seedlings. *Ann. Bot.* **128**, 453–468 (2021).
- Wang, H. et al. Natural variation and domestication selection of *ZmCKX5* with root morphological traits at the seedling stage in maize. *Plants* **10**, 1 (2021).
- Gaudin, A. C. M., McClymont, S. A. & Raizada, M. N. The nitrogen adaptation strategy of the wild teosinte ancestor of modern maize, *Zea mays* subsp. *parviglumis*. *Crop Sci.* **51**, 2780–2795 (2011).
- Gao, K., Chen, F., Yuan, L., Zhang, F. & Mi, G. A comprehensive analysis of root morphological changes and nitrogen allocation in maize in response to low nitrogen stress. *Plant Cell Environ.* **38**, 740–750 (2015).
- Mano, Y. et al. Variation for root aerenchyma formation in flooded and non-flooded maize and teosinte seedlings. *Plant Soil* **281**, 269–279 (2006).
- Ning, P., Li, S., Li, X. & Li, C. New maize hybrids had larger and deeper post-silking root than old ones. *Field Crop Res.* **166**, 66–71 (2014).
- Chen, X. et al. Changes in root size and distribution in relation to nitrogen accumulation during maize breeding in China. *Plant Soil* **374**, 121–130 (2014).
- Mano, Y. & Omori, F. Flooding tolerance in interspecific introgression lines containing chromosome segments from teosinte (*Zea nicaraguensis*) in maize (*Zea mays* subsp. *mays*). *Ann. Bot.* **112**, 1125–1139 (2013).
- Wang, H., Studer, A. J., Zhao, Q., Meeley, R. & Doebley, J. F. Evidence that the origin of naked kernels during maize domestication was caused by a single amino acid substitution in *tga1*. *Genetics* **200**, 965–974 (2015).

32. Li, Z. et al. Enhancing auxin accumulation in maize root tips improves root growth and dwarfs plant height. *Plant Biotechnol. J.* **16**, 86–99 (2018).
33. Zhao, Y. et al. The interaction between rice ERF3 and WOX11 promotes crown root development by regulating gene expression involved in cytokinin signaling. *Plant Cell* **27**, 2469–2483 (2015).
34. Asim, M., Ullah, Z., Oluwaseun, A., Wang, Q. & Liu, H. B. Signalling overlaps between nitrate and auxin in regulation of the root system architecture: insights from the *Arabidopsis thaliana*. *Int. J. Mol. Sci.* **21**, 2880 (2020).
35. Woll, K. et al. Isolation, characterization, and pericycle-specific transcriptome analyses of the novel maize lateral and seminal root initiation mutant *rum1*. *Plant Physiol.* **139**, 1255–1267 (2005).
36. von Behrens, I. et al. *Rootless with undetectable meristem 1* encodes a monocot-specific AUX/IAA protein that controls embryonic seminal and post-embryonic lateral root initiation in maize. *Plant J.* **66**, 341–353 (2011).
37. Tang, J. et al. Genetic dissection of plant height by molecular markers using a population of recombinant inbred lines in maize. *Euphytica* **155**, 117–124 (2007).
38. Li, H. et al. Genome-wide association study dissects the genetic architecture of oil biosynthesis in maize kernels. *Nat. Genet.* **45**, 43–50 (2013).
39. Wen, T. J., Hochholding, F., Sauer, M., Bruce, W. & Schnable, P. S. The *rootless1* gene of maize encodes a homolog of *sec3*, which is involved in polar exocytosis. *Plant Physiol.* **138**, 1637–1643 (2005).
40. Huang, P. et al. *Sparse panicle1* is required for inflorescence development in *Setaria viridis* and maize. *Nat. Plants* **3**, 17054 (2017).
41. Rademacher, E. H. et al. Different auxin response machineries control distinct cell fates in the early plant embryo. *Dev. Cell* **22**, 211–222 (2012).
42. Rinaldi, M. A., Liu, J., Enders, T. A., Bartel, B. & Strader, L. C. A gain-of-function mutation in *IAA16* confers reduced responses to auxin and abscisic acid and impedes plant growth and fertility. *Plant Mol. Biol.* **79**, 359–373 (2012).
43. Huang, J. et al. Formin homology 1 (OsFH1) regulates root-hair elongation in rice (*Oryza sativa*). *Planta* **237**, 1227–1239 (2013).
44. Kitomi, Y., Inahashi, H., Takehisa, H., Sato, Y. & Inukai, Y. OsIAA13-mediated auxin signaling is involved in lateral root initiation in rice. *Plant Sci.* **190**, 116–122 (2012).
45. Chen, H., Patterson, N. & Reich, D. Population differentiation as a test for selective sweeps. *Genome Res.* **20**, 393–402 (2010).
46. Hufford, M. B. et al. Comparative population genomics of maize domestication and improvement. *Nat. Genet.* **44**, 808–811 (2012).
47. Chen, Q. et al. The genetic architecture of the maize progenitor, teosinte, and how it was altered during maize domestication. *PLoS Genet.* **16**, e1008791 (2020).
48. Wang, Y., Deng, D., Bian, Y., Lv, Y. & Xie, Q. Genome-wide analysis of primary auxin-responsive *Aux/IAA* gene family in maize (*Zea mays* L.). *Mol. Biol. Rep.* **37**, 3991–4001 (2010).
49. Rosero, A. et al. *Arabidopsis* FH1 formin affects cotyledon pavement cell shape by modulating cytoskeleton dynamics. *Plant Cell Physiol.* **57**, 488–504 (2016).
50. Shi, J. et al. Ectopic expression of ARGOS8 reveals a role for ethylene in root-lodging resistance in maize. *Plant J.* **97**, 378–390 (2019).
51. Evans, M. M. & Poethig, R. S. Gibberellins promote vegetative phase change and reproductive maturity in maize. *Plant Physiol.* **108**, 475–487 (1995).
52. Trachsel, S., Kaepler, S. M., Brown, K. M. & Lynch, J. P. Maize root growth angles become steeper under low N conditions. *Field Crop Res* **140**, 18–31 (2013).
53. Kell, D. B. Breeding crop plants with deep roots: their role in sustainable carbon, nutrient and water sequestration. *Ann. Bot.* **108**, 407–418 (2011).
54. Hund, A., Reime, R. & Messmer, R. A consensus map of QTLs controlling the root length of maize. *Plant Soil* **344**, 143–158 (2011).
55. Schaefer, R. J. et al. Integrating coexpression networks with GWAS to prioritize causal genes in maize. *Plant Cell* **30**, 2922–2942 (2018).
56. York, L. M., Galindo-Castaneda, T., Schussler, J. R. & Lynch, J. P. Evolution of US maize (*Zea mays* L.) root architectural and anatomical phenes over the past 100 years corresponds to increased tolerance of nitrogen stress. *J. Exp. Bot.* **66**, 2347–2358 (2015).
57. Chen, F. et al. Breeding for high-yield and nitrogen use efficiency in maize: lessons from comparison between Chinese and US cultivars. *Adv. Agron.* **166**, 251–275 (2021).
58. Xu, C. et al. Cooperative action of the paralogous maize lateral organ boundaries (LOB) domain proteins RTCS and RTCL in shoot-borne root formation. *N. Phytol.* **207**, 1123–1133 (2015).
59. Suzuki, M., Sato, Y., Wu, S., Kang, B. H. & McCarty, D. R. Conserved functions of the MATE transporter BIG EMBRYO1 in regulation of lateral organ size and initiation rate. *Plant Cell* **27**, 2288–2300 (2015).
60. Zhang, Y. et al. LATERAL ROOT PRIMORDIA 1 of maize acts as a transcriptional activator in auxin signalling downstream of the *Aux/IAA* gene *rootless with undetectable meristem 1*. *J. Exp. Bot.* **66**, 3855–3863 (2015).
61. Taramino, G. et al. The maize (*Zea mays* L.) RTCS gene encodes a LOB domain protein that is a key regulator of embryonic seminal and post-embryonic shoot-borne root initiation. *Plant J.* **50**, 649–659 (2007).
62. Zhang, M. et al. Auxin efflux carrier ZmPGP1 mediates root growth inhibition under aluminum stress. *Plant Physiol.* **177**, 819–832 (2018).
63. Benjamins, R. & Scheres, B. Auxin: the looping star in plant development. *Annu Rev. Plant Biol.* **59**, 443–465 (2008).
64. Ikeda, Y. et al. Local auxin biosynthesis modulates gradient-directed planar polarity in *Arabidopsis*. *Nat. Cell Biol.* **11**, 731–738 (2009).
65. Lanza, M. et al. Role of actin cytoskeleton in brassinosteroid signaling and in its integration with the auxin response in plants. *Dev. Cell* **22**, 1275–1285 (2012).
66. Banno, H. & Chua, N. H. Characterization of the *arabidopsis* formin-like protein AFH1 and its interacting protein. *Plant Cell Physiol.* **41**, 617–626 (2000).
67. Martiniere, A., Gayral, P., Hawes, C. & Runions, J. Building bridges: formin1 of *Arabidopsis* forms a connection between the cell wall and the actin cytoskeleton. *Plant J.* **66**, 354–365 (2011).
68. Li, G. et al. Rice actin-binding protein RMD is a key link in the auxin-actin regulatory loop that controls cell growth. *Proc. Natl Acad. Sci. USA* **111**, 10377–10382 (2014).
69. Ogura, T. et al. Root system depth in *Arabidopsis* is shaped by *EXOCYST70A3* via the dynamic modulation of auxin transport. *Cell* **178**, 400–412 (2019).
70. Yang, P. et al. Light modulates the gravitropic responses through organ-specific PIFs and HY5 regulation of *LAZY4* expression in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **117**, 18840–18848 (2020).
71. Yang, X. et al. Characterization of a global germplasm collection and its potential utilization for analysis of complex quantitative traits in maize. *Mol. Breed.* **28**, 511–526 (2011).
72. Trachsel, S., Kaepler, S. M., Brown, K. M. & Lynch, J. P. Shovelomics: high throughput phenotyping of maize (*Zea mays* L.) root architecture in the field. *Plant Soil* **341**, 75–87 (2010).

73. Das, A. et al. Digital imaging of root traits (DIRT): a high-throughput computing and collaboration platform for field-based root phenomics. *Plant Methods* **11**, 51 (2015).
74. Colombi, T. et al. Next generation shovelomics: set up a tent and REST. *Plant Soil* **388**, 1–20 (2015).
75. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
76. Fu, J. et al. RNA sequencing reveals the complex regulatory network in the maize kernel. *Nat. Commun.* **4**, 2832 (2013).
77. Unterseer, S. et al. A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k SNP genotyping array. *BMC Genomics* **15**, 823 (2014).
78. Ganai, M. W. et al. A large maize (*Zea mays* L.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PLoS ONE* **6**, e28334 (2011).
79. Elshire, R. J. et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE* **6**, e19379 (2011).
80. Liu, H. J. et al. MODEM: multi-omics data envelopment and mining in maize. *Database* **2016**, baw117 (2016).
81. Yu, J. et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38**, 203–208 (2006).
82. Zhang, Z. et al. Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* **42**, 355–360 (2010).
83. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
84. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
85. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
86. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
87. Maere, S., Heymans, K. & Kuiper, M. BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics* **21**, 3448–3449 (2005).
88. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
89. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
90. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
91. Gui, S. et al. ZEAMAP, a comprehensive database adapted to the maize multi-omics era. *iScience* **23**, 101241 (2020).
92. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
93. Mikheenko, A., Prijbelski, A., Saveliev, V., Antipov, D. & Gurevich, A. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* **34**, i142–i150 (2018).
94. Nakamura, T., Yamada, K. D., Tomii, K. & Katoh, K. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics* **34**, 2490–2492 (2018).
95. Ishida, Y., Hiei, Y. & Komari, T. Agrobacterium-mediated transformation of maize. *Nat. Protoc.* **2**, 1614–1621 (2007).
96. Wang, B. et al. Tryptophan-independent auxin biosynthesis contributes to early embryogenesis in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **112**, 4821–4826 (2015).
97. Tan, H. et al. A crucial role of GA-regulated flavonol biosynthesis in root growth of *Arabidopsis*. *Mol. Plant* **12**, 521–537 (2019).
98. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

## Acknowledgements

This study was financially supported by the National Key Research and Development Program of China (grant nos. 2021YFF1000500 (Q.P.), 2021YFD1200700 (F.C.) and 2016YFD0100700 (L.Y.)), the National Natural Science Foundation of China (grant nos. 31972485 (F.C.) and 31971948 (Q.P.)), the Hainan Natural Science Foundation Innovation Research Team Project (grant no. 321CXTD443 (F.C.)), the Hainan Provincial Science and Technology Plan Sanya Yazhou Bay Science and Technology City Joint Project (grant no. 320LH011 (Q.P.)) and the China Postdoctoral Science Foundation (grant no. 2021M693431 (W.R.)). The transgenic maize seeds were produced by the Center for Crop Functional Genomics and Molecular Breeding of China Agricultural University.

## Author contributions

Q.P., L.Y. and F.C. conceived and designed the research. W.R., L.Z., J. Liang, L.W., P.L., Z.L., X.L., Z. Zhang and J. Li performed phenotypic measurements. W.R. and Q.P. performed the data analyses. L.C. performed plasmid construction and genetic transformation. W.R., K.H. and Z. Zhao characterized the transgenic overexpression lines. J.Y. provided the maize inbred lines and genotype set. W.R. and Q.P. wrote the manuscript. F.A., G.M., J.Y., F.Z., F.C., L.Y. and Q.P. revised the manuscript. All authors contributed to the final version of the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s41477-022-01274-z>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41477-022-01274-z>.

**Correspondence and requests for materials** should be addressed to Fanjun Chen, Lixing Yuan or Qingchun Pan.

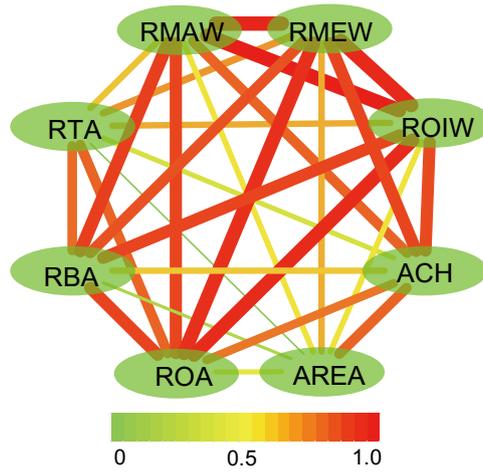
**Peer review information** *Nature Plants* thanks Yusaku Uga, Ana Letycia Basso Garcia, Ana Caño-Delgado and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

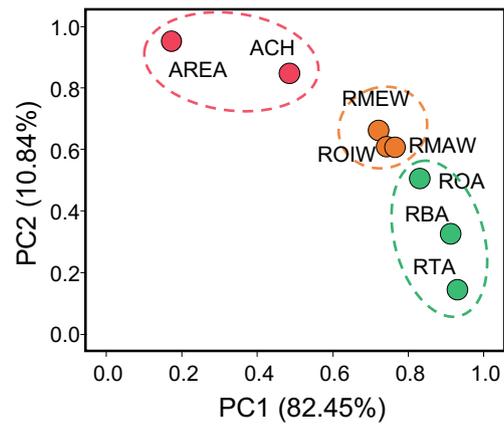
**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

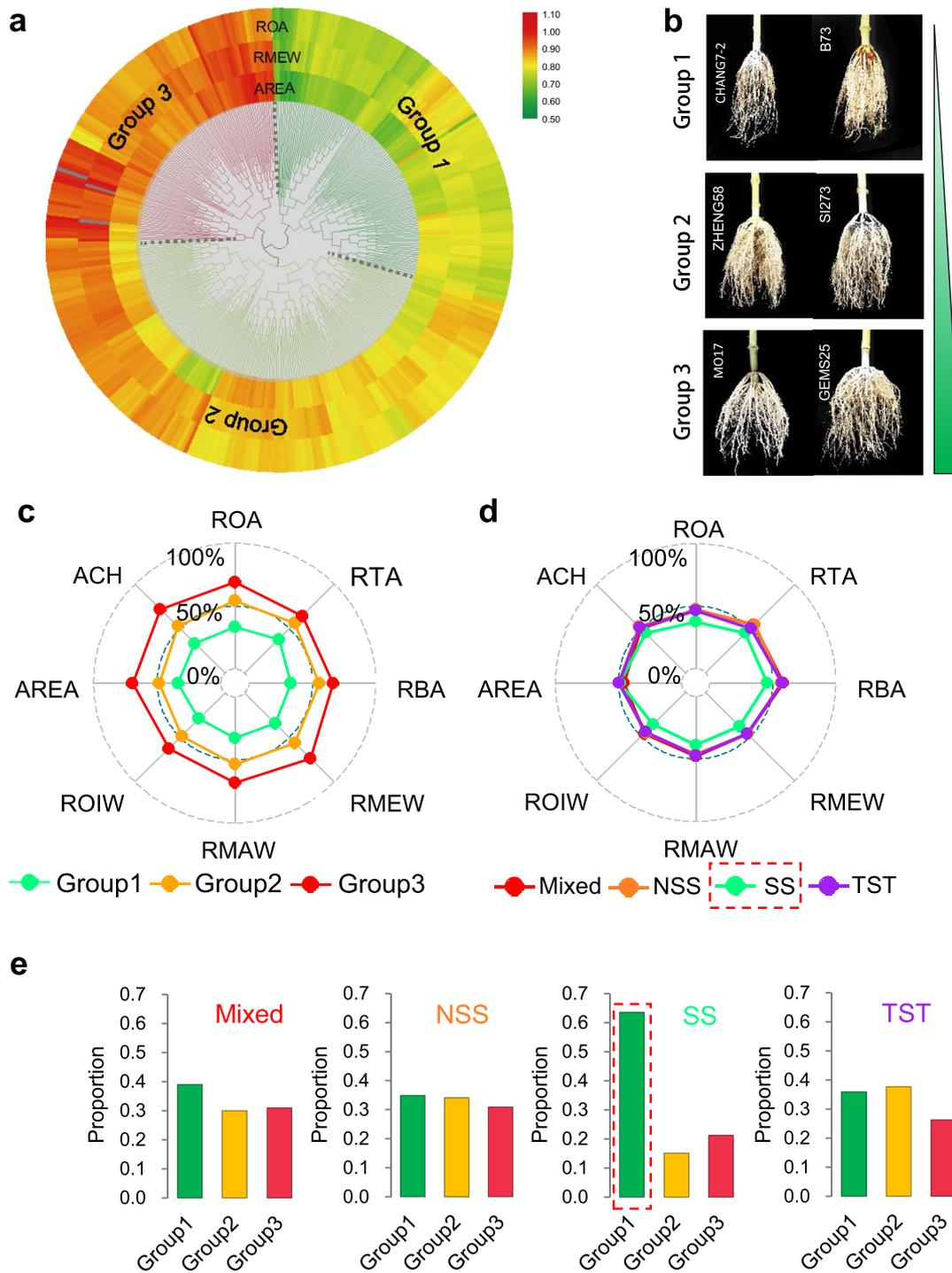
© The Author(s), under exclusive licence to Springer Nature Limited 2022



**Extended Data Fig. 1 | Pearson correlations among eight root traits.** The red lines represent positive correlations, and the green lines represent negative correlations. The line width represents the strength of the correlation. Yellow lines indicate that the correlation coefficient was close to zero.



**Extended Data Fig. 2 | Principal component analysis of eight root traits.** The red ellipse indicates the area-related traits; the yellow ellipse indicates the width-related traits; and the green ellipse indicates the angle-related traits.



**Extended Data Fig. 3 | Cluster analysis of 380 maize inbred lines based on root traits. (a)** Cluster analysis of 380 inbred lines based on ROA, RMEW, and AREA. **(b)** Representative inbred lines from the three clusters (groups 1–3). **(c)** Comparison of eight root traits among groups 1–3. **(d)** Comparison of eight root traits among

four subpopulations. **(e)** Proportion of lines from each of four subgroups in the three cluster groups. The four subgroups (Mixed, SS, NSS, and TST) are based on genetic relationships among the different inbred lines. Mixed, mixed group; SS, stiff stalk group; NSS, non-stiff stalk group; TST, tropical and subtropical group.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a                                 | Confirmed  |
|-------------------------------------|--|
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated   |

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection We used open source software and codes for data collection. All the data collection methods in this manuscript have been described in detail in the Methods section.

Data analysis All softwares and codes used in the present study are publicly available and the corresponding versions are described in detail in the Online Methods, Data availability, and Code availability.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data supporting the findings of this work are available within the paper and its Supplementary Information. The genotype set, population structure, and kinship data can be downloaded from the Maizego website (<http://www.maizego.org/Resources.html>). All root phenotype data for the 380 inbred maize lines are included in

Supplementary Table 25. The RNA-sequencing reads used to construct the co-expression network and the root transcriptome sequencing reads were deposited in the NCBI Sequence Read Archive under accession codes PRJNA694491 and PRJNA693427, respectively. Source data are provided with this paper.

## Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender	<input type="text" value="not applicable"/>
Population characteristics	<input type="text" value="not applicable"/>
Recruitment	<input type="text" value="not applicable"/>
Ethics oversight	<input type="text" value="not applicable"/>

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences     Behavioural & social sciences     Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	<input type="text" value="A maize association panel of 380 genetically diverse inbred lines was prepared, including tropical/subtropical (n=253), and temperate (n=127) materials. The 380 inbred lines used in this study were collected from all over the world and represent global maize diversity, with high genetic diversity sufficient for genome-wide association study. Three transgenic lines of each gene (ZmRSA3.1 and ZmRSA3.2) were used to verify the root phenotype. No sample size calculation was performed in this study. This was also described in the manuscript."/>
Data exclusions	<input type="text" value="No data in this study were excluded."/>
Replication	<input type="text" value="The 8 root traits for 380 inbred lines were repeatedly measured across four environments. Three sampling replicates were used to verify the root phenotypes of ZmRSA3.1 and ZmRSA3.2 across two environments. Each replicate consists of root phenotypes from 4-10 biologically independent samples."/>
Randomization	<input type="text" value="All phenotype identification trials used a randomized complete block design."/>
Blinding	<input type="text" value="The investigators were blinded to group allocation during data collection and analysis."/>

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Included in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging