1   **Evidence that variation in root anatomy contributes to local adaptation in Mexican**

2   **native maize**

3   Chloee M. McLaughlin[1], Meng Li[2], Melanie Perryman[2], Adrien Heymans[3,4], Hannah

4   Schneider[5], Jesse R. Lasky[6] and Ruairidh J. H. Sawers[2]

5   [1]Intercollege Graduate Degree Program in Plant Biology, The Pennsylvania State University,

6   University Park, PA 16802, USA.

7   [2]Department of Plant Science, The Pennsylvania State University, University Park, PA

8   16802, USA.

9   [3]Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural

10   Sciences, Kempe JCK-2130, Sweden.

11   [4]Earth and life Institute, UC Louvain, Belgium

12   [5]Department of Physiology and Cell Biology, Leibniz Institute for Plant Genetics and Crop

13   Plant Research (IPK), OT Gatersleben, 06466, Seeland Germany

14   [6]Department of Biology, The Pennsylvania State University, University Park, PA 16802,

15   USA.

16   **Running head:** Maize root diversity

17   CMM cmc7333@psu.edu

18   ML mul826@psu.edu

19   MP mpm5587@psu.edu

20   AH adrien.heymans@slu.se

21   HS schneiderh@ipk-gatersleben.de

22   JL jrl35@psu.edu

23   RJHS rjs6686@psu.edu

24

25

26

27

28   **ABSTRACT**

29   Mexican native maize (*Zea mays* ssp. *mays*) is adapted to a wide range of climatic and

30   edaphic conditions. Here, we focus specifically on the potential role of root anatomical

31   variation in this adaptation. In light of the investment required to characterize root anatomy,

32   we present a machine learning approach using environmental descriptors to project trait

33   variation from a relatively small training panel onto a larger panel of genotyped and

34   georeferenced Mexican maize accessions. The resulting models defined potential biologically

35   relevant clines across a complex environment and were used subsequently in genotype-

36   environment association. We found evidence of systematic variation in maize root anatomy

37   across Mexico, notably a prevalence of trait combinations favoring a reduction in axial

38   conductance in cooler, drier highland areas. We discuss our results in the context of

39   previously described water-banking strategies and present candidate genes that are associated

40   with both root anatomical and environmental variation. Our strategy is a refinement of

41   standard environmental genome wide association analysis that is applicable whenever a

42   training set of georeferenced phenotypic data is available.

43   **Keywords:** GEA, Local adaptation, Root anatomy, Maize

44

45

46

47

48

INTRODUCTION

49

50 Abiotic stress is a major driver of plant phenotypic diversity (Lowry 2012), acting to select

51 locally adapted varieties with specific morphological, physiological, and phenological traits

52 (Fumagalli et al. 2011; Stebbins 1952; Hereford 2009https://paperpile.com/c/jHvuxa/s6Ua).

53 Differences in such selective pressures over a continuously varying environment produce

54 clines of genetic and phenotypic variation, reflecting the shifting costs and benefits of diverse

55 biological strategies (Joswig et al. 2022). Although plants are plastic in the face of

56 environmental challenges (Des Marais et al. 2013; Lasky et al. 2014), locally adapted

57 specialists can constitutively express adaptive strategies, anticipating the need for the

58 induction of stress responses (Levins 1968; von Heckel et al. 2016; Aguilar-Rangel et al.

59 2017). As a consequence, genotypes sourced from diverse locations will typically still display

60 trait variation indicative of adaptation to their home environments when grown in a benign

61 common garden (Stinchcombe et al. 2004; Janzen et al. 2022; Shimono et al. 2009).

62 Mexican native maize (*Zea mays* ssp. *mays*) varieties ("landraces") represent an

63 attractive system for the study of local adaptation. Mexico is the center of origin of maize,

64 and today hosts 59 described native varieties cultivated from sea level to an elevation of

65 3,400 m, in environments ranging from semi-arid to hot and humid (Ruiz Corral et al. 2008;

66 Arteaga et al. 2016; Perales and Golicher 2014). This diversity has been extensively sampled,

67 and large collections of georeferenced and genetically characterized material are available

68 (Arteaga et al. 2016; Romero Navarro et al. 2017; Mercer and Perales 2019; Janzen et al.

69 2022). Throughout maize domestication and diversification, farmers have consciously

70 selected for agronomically and culturally desirable traits, principally targeting the female

71 inflorescence (the ear) to produce a rich variety of form (Louette and Smale 2000; Bellon et

72 al. 2018). In parallel, unconscious selection has likely acted to adapt varieties to local

73 conditions (Romero Navarro et al. 2017; Mercer and Perales 2019; Janzen et al. 2022) and

74   enhance tolerance to different environmental stressors after dispersal to new environments

75   (Magalhaes et al. 2007;Eagles and Lothrop 1994; Bayuelo-Jiménez et al. 2011).

76   In this study, we focus specifically on the potential role of root trait variation in the

77   adaptation of maize to different climatic and edaphic environments in Mexico. Although less

78   easily visible than aboveground traits, the maize root system has been substantially impacted

79   by domestication, diversification and modern breeding (Gaudin et al. 2014; Chen et al. 2022;

80   Lopez-Valdivia et al. 2022; Burton et al. 2013; Ren et al. 2022). Roots are fundamental to

81   plant water and nutrient acquisition, and play a key role in both wild and domesticated plants

82   in determining performance under resource limitation (Wahl and Ryser 2000; Markesteijn

83   and Poorter 2009; Ma et al. 2018). Within maize specifically, root trait variation among

84   inbred breeding lines has been linked to performance differences under both water (Jaramillo

85   et al. 2013; Bomfim et al. 2011; Schneider et al. 2020) and nutrient (Schneider, Postma, et al.

86   2017; Galindo-Castañeda et al. 2018) limitation. Extensive root trait variation has also been

87   reported among native maize varieties (Burton et al. 2013), although the associated functional

88   impact and possible adaptive roles remain to be fully characterized. Variation in the plant

89   root system can be considered from the anatomy of individual roots to overall root system

90   architecture (Jung and McCouch 2013; Lynch 2019), traits at all levels interact to determine

91   overall root system function in the context of a given environment (Klein et al. 2020). Here,

92   we limit ourselves to consideration of variation in root anatomy.

93   The maize root system consists of a variety of root classes that vary in function and

94   importance during development (Atkinson et al. 2014; Viana et al. 2022; Hochholdinger

95   2009). Within this range, root anatomy develops on a basic pattern of radially organized

96   tissue types: an external epidermis, the ground tissue, and an inner stele containing the

97   pericycle and vasculature (Lynch et al. 2021). The epidermis protects the inner layers from

98   physical damage and is in direct contact with the rhizosphere, playing a key role in water and

99   nutrient exchange. The ground tissue is further differentiated into cortex and the endodermis

100   that encloses the stele. In the model plant *Arabidopsis,* the cortex is composed of only two

101   layers, a single layer of cortical parenchyma and the endodermis. In maize, however, the

102   cortex divides to form multiple cell layers, impacting the physical (Chimungu et al. 2015),

103   hydraulic (Heymans et al. 2020; Sidhu et al. 2023) and radial nutrient transport (Hu et al.

104   2014; Schneider et al. 2017) properties of the root, as well as accommodating beneficial

105   endomycorrhizal fungi (Sawers et al. 2008; Bennett and Groten 2022). Developmental and

106   environmental cues can trigger cells in the cortex to undergo programmed cell death and form

107   aerenchyma. The resulting cortical air-filled lacunae help maintain gas exchange and mitigate

108   hypoxia under flooding (Colmer 2003; Mano and Nakazono 2021). In addition, the reduction

109   in root metabolic cost following aerenchyma formation can be beneficial in resource limiting

110   conditions including drought and low availability of nitrogen or phosphorus (Jaramillo et al.

111   2013; Galindo-Castañeda et al. 2018). The cell walls of the endodermis are impregnated with

112   suberin to form the Casparian strip and, in later development, further reinforced with lignin to

113   act as a barrier that restricts apoplastic transport into the stele from the surrounding cortical

114   root tissue. The central stele contains the xylem and phloem vessels that axially transport

115   water and nutrients. The xylem is composed of small protoxylem vessels and larger

116   metaxylem vessels, the latter providing the majority of the transport capacity in mature root

117   tissues (Doussan et al. 1998). The size and number of metaxylem vessels and living cortical

118   area influence root radial and axial hydraulic properties (conductivity and conductance;

119   (Frensch and Steudle 1989; Schneider, Wojciechowski, et al. 2017), impacting water capture

120   and plant performance (Richards and Passioura 1989; Couvreur et al. 2012).

121   In this study, we aim to characterize heritable variation in root anatomical traits in

122   Mexican native maize and to associate patterns of phenotypic and genetic variation with the

123   source environment. The genetic basis of local adaptation can be characterized through

124   associations between genotype and phenotypes involved in local adaptation or between

125   genotype and environment (Fournier-Level et al. 2011; Hoban et al. 2016). Approaches using

126   genotype-environment association (GEA) have the advantage of not requiring resource-costly

127   phenotypic characterization, although they must assume local adaptation to have occurred

128   over the tested environments and can be complicated by the confounding effects of

129   population structure (Lasky et al. 2023). One challenge of GEA is that it can be unclear

130   which aspect(s) of the environment might drive selection, such that it may be unclear what

131   associations to test. Here, we present an approach that first fits a trait relationship across the

132   multivariate environment using a training set of phenotypically characterized, georeferenced

133   varieties. This fitted model is then used to project trait values onto a larger set of available

134   genotyped and georeferenced varieties, and the predicted trait values then used in genome-

135   wide association (GWA) to find putative adaptive loci. We compare the results of our

136   approach with standard GEA and a previously published phenotypic GWA using modern

137   maize breeding lines. We present evidence for systematic variation in root anatomy driven by

138   differences across the Mexican environmental landscape, identifying candidate genetic

139   variants and linked genes associated with both phenotypic and genetic clines.

140

141   **RESULTS**

142   *Root anatomy varies among Mexican native maize varieties*

143   To characterize the relationship between root anatomy and environment in Mexican native

144   maize, we assigned georeference data to 39 Mexican accessions phenotypically characterized

145   in a previous root anatomy study (Burton et al., 2013; hereafter, the Burton panel) and

146   extracted associated climate and soil descriptors from publicly available databases (see

147   Materials and Methods). We supplemented published trait data with a re-analysis of the

148   original cross-sectional images to generate a final phenotypic dataset of 16 anatomical traits

149   (Table 1). We performed a principal component (PC) analysis on the phenotypic data: PC1

150   was negatively correlated with root cross-sectional area (Fig. S1; Fig. 1B). PC2 and PC3

151   captured allometric relationships among traits (Fig. 1A, B): PC2 was associated with

152   variation in the relative contribution of the stele to the total area, with *total metaxylem vessel*

153   *area* and associated metaxylem traits loading antagonistically to *total cortical area* and *root*

154   *cross-section area*; PC3 captured an apparent trade-off between *cortical cell file number* and

155   *cortical cell size* along with variation in cortical *aerenchyma area*. In addition to grouping

156   traits by PC analysis, we used a modeling pipeline linking the GRANAR and MECHA

157   packages to simulate cross-sectional anatomy (Heymans et al. 2020) and predict radial

158   conductivity (k_r) and radial and axial conductance (K_r, K_x; Couvreur et al. 2018; Fig. S2).

159   Axial and radial hydraulic conductance were negatively correlated with PC1, indicating the

160   greater capacity of larger diameter roots for water transport (Fig. 1C). PC1 and PC2 were

161   both positively correlated with radial conductivity (Fig. 1D), reflecting the greater ease of

162   water transport across roots with less cortex (Heymans et al. 2020). Further associations

163   between anatomical trait PCs and derived hydraulic properties were not easily captured by

164   simple correlations (Fig. S2).

165   As a first attempt to identify clinal relationships between root anatomy and local

166   environment, we examined the correlation of root anatomy PCs to four basic environmental

167   descriptors (elevation, annual precipitation, mean temperature, and soil pH). After adjusting

168   p-values for multiple comparison testing, we did not find evidence for variation in root

169   anatomy (Fig. S3; S4) or derived hydraulic properties (Fig. S5) to have significant

170   associations to environmental descriptors of accessions' point of origin. We did find mild

171   evidence for root anatomical variation summarized by PC1 to be related to elevation (Fig.

172   S3). When dividing PC1 axes into terciles, individuals with the most positive PC1 loadings

173   were sourced from the highest elevations (Fig. 1E).

174

175   *Combined environmental descriptors predict variation in root anatomy*

176   Local adaptation is driven by varied aspects of the environment and their interactions. To

177   capture more complex trait-environment relationships, we used a feature-reduction method

178   (Boruta algorithm; Kursa and Rudnicki 2010) to select the most informative of a full set of

179   157 available environmental descriptors for each anatomical trait, and subsequently

180   combined the chosen descriptors into random forest (RF) models to relate environment and

181   trait (Fig. 2A; S6). Nine of the 16 tested root anatomical traits were associated with

182   environmental descriptors by the Boruta algorithm (Table S1). 39 different environmental

183   descriptors were used as input for RF models across the 9 modeled anatomical traits, with

184   individual models using from two (*percent of cortex as aerenchyma*) to 16 (*total metaxylem*

185   *vessel area*) environmental descriptors. We observed varying goodness-of-fit from RF

186   models and the R-squared for predicted vs observed trait values ranged from 0.32

187   (*aerenchyma area*) to 0.01 (*total metaxylem vessel area*) (Fig. S7).

188   We used the GRANAR-MECHA pipeline to combine predicted trait values and

189   compared observed and predicted anatomies, both graphically (Fig. 2C) and with respect to

190   hydraulic properties (Fig. 2D). Modeled conductivity and conductance values for predicted

191   anatomies correlated well with values from the observed data (radial conductivity, $r = 0.69$, p

192   < 0.01; radial conductance $r = 0.28$, $p = 0.07$; axial conductance $r = 0.59$, $p < 0.01$; Fig. S8),

193   indicating that our RF models successfully captured differences in anatomical traits that

194   impact root hydraulic properties.

195

Random forest prediction of root anatomy across Mexican native maize

196

197   To estimate root anatomical diversity across a broader sampling of native Mexican maize, we

198   applied our Burton-trained RF models to a larger collection of 1791 genotyped and

199   georeferenced Mexican accessions (hereafter, the CIMMyT panel; Romero Navarro et al.

200   2017; Fig. S9). We used the georeference data to link environmental descriptors to each

201   accession and passed these to the RF models, generating a complete phenotypic set of 9

202   estimated root anatomical traits for the 1791 accessions (Supplementary Information). To

203   summarize patterns among the predicted trait values, we used partition-against-medians

204   (PAM) clustering (Klein et al. 2020; Maechler et al. 2021) to group the accessions into seven

205   phenotypic clusters (Fig. S10, S11). The clusters 1 through 7 were composed of 308, 366,

206   370, 231, 158, 277 and 131 accessions, respectively. The structure defined by the clustering

207   was not strong (mean silhouette value = 0.25), reflecting the continuous nature of the

208   environmental descriptors driving the RF models, but did provide a context for subsequent

209   analyses. We also passed the median trait values of each cluster to the GRANAR-MECHA

210   pipeline to obtain average anatomies and hydraulic properties.

211   Clusters were distinguished by the relative elaboration of cortex and stele and

212   associated hydraulic properties (Fig. 3). In Clusters 1 and 3, the stele (*total stele area: root*

213   *cross-section area*; *total stele area: total cortical area*) was relatively small, although

214   individual metaxylem vessels were large (*individual metaxylem area*, *individual metaxylem*

215   *diameter*) and, consequently, the *total metaxylem vessel area* and axial conductance were

216   relatively high. In contrast, Clusters 5, 2, and 6 were distinguished by a small stele and small

217   metaxylem vessels, associated with low axial conductance relative to other clusters. In

218   Cluster 5, the small size of the metaxylem vessels was further associated with a low *number*

219   *of metaxylem vessels* resulting in the lowest *total metaxylem vessel area* and the lowest axial

220   conductance of the clusters.

221   We examined the clusters with respect to previous morphological-isozymatic

222   (Sanchez G. et al. 2000) and environmental (Ruiz Corral et al. 2008) classifications of

223   Mexican native maize (Table S2, S3; Fig. S11). Following trends reported for overall

224   Mexican maize diversity (Sanchez G. et al. 2000), our clusters were structured with respect to

225   elevation (Fig. 3). Clusters 2, 5 and 6 were enriched (Fisher test for variety count in or out of

226   the cluster, adj.p < 0.05) for varieties belonging to the previously defined highland group

227   (Sanchez G. et al. 2000; Fig. 3; Table S3). Cluster 5, containing the highest elevation

228   varieties, was centered on Mexico City, although it also contained accessions from the

229   highlands of Chihuahua in northern Mexico; Cluster 6 extended from north to south along the

230   Sierra Madre Occidental; Cluster 2 was again centered on Mexico City, although with greater

231   representation further west along the trans Mexican volcanic belt than Cluster 6 and included

232   several accessions from the Chiapas highlands on the southern border of Mexico. The mid-

233   elevation Clusters 4 and 7 were loosely sourced from the center-to-south and center-to-north

234   of Mexico, respectively. The Clusters 1 and 3 were enriched for varieties in the lowland

235   short-to-medium maturity and tropical dent groups (Sanchez G. et al. 2000; Table S2, S3),

236   with Cluster 1 from the Gulf Coast, the Yucatan and lowland Guatemala and Cluster 3 from

237   the Pacific Coast. Prior environmental classification was in line with the observed elevational

238   cline: Clusters 2, 5 and 6 were enriched for varieties previously assigned to "temperate to

239   semi-hot" environments; Clusters 1 and 3 were enriched for varieties assigned to the "very

240   hot" niche (Ruiz Corral et al. 2008; Fig. S12; Table S3).

241   Considering the average anatomies and mean values of environmental descriptors

242   associated with each cluster, we could discern a broad trend of a reduction in axial

243   conductance with increasing elevation (Fig. 3D). Our models associated the colder, drier

244   highland niche (>2,500 masl; Eagles and Lothrop 1994; Ruiz Corral et al. 2008) with both

245   fewer and smaller metaxylem vessel elements (Cluster 5). Conversely, the hot, wet lowlands

246   (Eagles and Lothrop 1994; Ruiz Corral et al. 2008) were associated with a relatively larger

247   stele accommodating a greater number of larger metaxylem vessel elements (Clusters 1 and

248   3).

249

250   *Novel phenotypic evaluation supports random forest predicted variation in axial*

251   *conductance*

252   To empirically evaluate our RF models, we characterized eight maize accessions from across

253   Mexico that had not been used previously in the Burton study (Fig. 4A, B). We measured

254   root anatomical traits following the Burton protocol, and for each trait compared the observed

255   best linear unbiased predictor (BLUP) with the results of our environmental RF predictions

256   (Fig. S13). The correlation between observed BLUPs and RF predictions ranged from

257   relatively high for cortical traits (*aerenchyma area*, $r = 0.71$; *percent of cortex as*

258   *aerenchyma*, $r = 0.65$; *percent of cortex as cortical cells*, $r = 0.52$) to lower for metaxylem

259   vessel traits (*number of metaxylem vessels*, $r = 0.30$; *total metaxylem vessel area, r = 0.27*;

260   *individual metaxylem vessel area*, $r = 0.11$) and allometric traits (*total stele area:root cross-*

261   *section area, r = 0.18; total stele area:total cortical area r = 0.06*). Predictions of *individual*

262   *metaxylem vessel diameter* were not well supported by observed values ($r = -0.24$).

263   We assessed overall concordance between observed and predicted anatomy by using

264   the Procrustes transformation (Schönemann 1966) to minimize the distance between each set

265   of observed and predicted trait values across the eight accessions. Observations and

266   predictions were well matched for six of the eight accessions, with the Jala and Nal Tel

267   accessions being a poorer fit (Fig. 4D; Fig. S14). The difference in overall root anatomy in

268   material sourced from the highlands and lowlands was well supported by both observed and

269   predicted trait values (Fig. 4D, S15) and modeled hydraulic properties (Fig. 4E, F). Overall,

270   although individual traits were not always well predicted for any given accession, our

271   methodology and training data were sufficient to capture broad stratification of anatomical

272   traits and hydraulic properties across the environment.

273

274   **Genome wide association analysis using predicted trait values identifies novel candidate**

275   **genes**

276   To look for genetic evidence linking root anatomy to the local environment in Mexican

277   maize, we ran a GWA analysis on the CIMMyT panel using the RF predicted trait values.

278   Given the nature of the RF models, this prediction GWA is, in effect, a development of a

279   standard environmental GWA analysis with modeled trait values capturing complex

280   combinations of the individual environmental descriptors used in RF model construction. For

281   comparison, we ran separate environmental GWA analyses for each of the 39 environmental

282   descriptors used in RF modeling, and also re-analyzed published phenotypic data for a panel

283   of 175 maize inbred lines (hereafter, the WIDP panel, Schneider et al. 2020). We extracted

284   phenotypic data for 8 of our 9 RF modeled traits (not including *percent of living cortical*

285   *area*), combining values obtained for well-watered and water-limited treatments into a single

286   GWA model (Runcie and Crawford 2019), estimating variant main (G) and variant x

287   treatment (GxE) effects. To facilitate comparison across panels genotyped using different

288   platforms, we used the MAGMA pipeline (de Leeuw et al. 2015) to combine signals across

289   single nucleotide polymorphisms (SNPs) to a single gene level value. Here, we assigned any

290   SNP +/- 2.5 kb from an annotated gene model to that gene. In the following discussion of

291   overlap between our different GWA analyses, we consider only genes captured in both

292   CIMMyT and WIDP markersets. In later identification of the genes of greatest interest from

293   the predicted GWA analysis, we do not take the WIDP markeset into account.

294   To compare CIMMyT "predicted", CIMMyT "environment", WIDP "G" and WIDP

295   "GxE" GWA analyses, we selected the top 100 genes (determined by p-value) per trait for

296   each analysis and combined these into candidate gene lists, obtaining sets of 636 unique

297   "prediction" genes, 1,282 unique "environment" genes, 542 unique "G" genes and 576

298   unique "GxE" genes (Fig. 5; Supplementary Information). Only 19% of the prediction genes

299   were also present in the environment set (Fig. 5B), indicating that the two were not redundant

300   and that the prediction set was capturing patterns not revealed by separate analyses of the

301   individual environmental descriptors. For example, a region of the short arm of chromosome

302   10 was linked to *mean metaxylem vessel diameter* in the prediction analysis (Fig. S16). In this

303   case, the $-\log_{10}(P)$ value of the most significant SNP for the predicted trait is approximately

304   double that of the best supported environmental descriptor (precipitation in October). The

305   best supported SNP in this region fell within the gene *Trichome birefringence-like 10* (*Tbl10*;

306   Zm00001d023378; Fig. S16). Natural variation in *Tbl10* has previously been linked to

307   variation in flowering time (Chen et al. 2012; Kusmec et al. 2017), height (Wang et al. 2022),

308   and root diameter (Pace et al. 2015). As such, *Tbl10* illustrates a compelling candidate for

309   further follow-up that would not have been identified by standard environmental GWA.

310

311   **The gene *Vq29* is linked to variation in both metaxylem traits and source elevation**

312   There was no evidence that the prediction set was enriched for WIDP root anatomy candidate

313   genes with respect to the environment set - both contained 2-3% WIDP G and GxE genes

314   (Fig. 5B). Nonetheless, we do consider the 29 genes identified in both prediction and WIDP

315   (G and/or GxE) GWA to be high confidence candidates for further characterization (Fig. 5B,

316   S17; Table S4). For example, the gene *Vq29* (Zm00001d015397) on the short arm of

317   chromosome 5 was associated with *number of metaxylem vessels* and *total metaxylem vessel*

318   *area* in the prediction GWA and with *individual metaxylem vessel area* and *individual*

319   *metaxylem vessel diameter* in the WIDP G analysis (Fig. 6A). The *Vq29* gene is predicted to

320   encode a VQ domain transcription factor, part of a large family of proteins that interact with

321   members of the WRKY family under stress (Song et al. 2015), including in response to

322   hypoxia, ozone or nitric oxide (León et al. 2020). The minor allele of the highest scoring SNP

323   (S5:88306863) was associated with greater *number of metaxylem vessels* and *total metaxylem*

324   *vessel area* and declined in frequency within our clusters with increasing predicted values of

325   these same traits (Fig. 6B). Based on gene expression atlas data (Walley et al. 2016), *Vq29* is

326   most highly expressed in the roots, consistent with a role in metaxylem development (Fig.

327   6C). Geographically, the minor allele of S5:88306863 was most prevalent in the central

328   Mexican highlands (Fig. 6D), and the MAF increased with mean elevation across our

329   previously defined clusters (Fig. 6E). In summary, *Vq29* nicely illustrates an example of a

330   candidate gene associated with phenotypic variation in root anatomy in the inbred WIDP

331   panel that also shows clinal genetic variation across the Mexican environment.

332

DISCUSSION

334   We have presented evidence that variation in root anatomy contributes to local adaptation in

335   Mexican native maize. We used predictive models to define biologically relevant clines over

336   which we identified both genotypic and phenotypic variation. Shared GWA candidates

337   between Mexican native maize and modern inbred lines indicated an element of common

338   genetic architecture, although we also identified novel candidates specific to the native

339   Mexican material. Phenotypic patterns suggested that local differences in precipitation and

340   temperature are associated with heritable variation in maize root anatomy. Root anatomical

341   variation broadly followed the established grouping of Mexican maize varieties, themselves

342 strongly stratified by environment. Our observations are consistent with a role for root

343 anatomical variation in local adaptation. Our combination of environment-based models and

344 GWA allowed us to leverage a relatively small sample of phenotypically characterized

345 locally adapted varieties to identify novel associations between phenotype, genotype and

346 environment in the context of broader maize diversity.

347   Our analyses highlight a predominance of anatomies predicted to reduce axial

348 conductance in material sourced from arid subtropical or temperate environments. In both

349 observed and predicted phenotypic data, varieties from the cooler, drier highland regions

350 were associated with fewer and/or narrower metaxylem vessels and a reduction in the area of

351 the stele with respect to cortex. Comprehensive revision of data across taxa has previously

352 suggested that the capacity for axial water transport is typically greater in plants from wet

353 environments and reduced in plants adapted to xeric conditions (Feng et al. 2016; Lynch et al.

354 2021). Although somewhat counterintuitive, reducing water uptake under dry conditions may

355 benefit plants by reducing root tip desiccation (Richards and Passioura 1989), preventing

356 cavitation (Nardini et al. 2013) and enabling the conservation of soil water resources across

357 the growing season (Richards and Passioura 1989; Leitner et al. 2014). Studies of

358 interspecific variation in crops support these hypotheses with both narrower metaxylem

359 vessels (Priatama et al. 2022; Allah et al. 2010; Purushothaman et al. 2013; Peña-Valdivia et

360 al. 2005) and fewer metaxylem vessels (Strock et al. 2021) being associated with enhanced

361 drought tolerance. Similarly, selection for reduced xylem vessel diameter in Australian wheat

362 has been reported to successfully increase yield under water limitation (Richards and

363 Passioura 1989). In the Mexican highlands, farmers traditionally plant prior to the beginning

364 of the annual rains to maximize the length of the growing season and ensure crops reach

365 maturity prior to the first frosts (Eagles and Lothrop 1994). As a consequence, seed is deep

366 planted to better access residual soil moisture, as well as to offer protection from low

367   temperatures, a practice also employed in the southwestern US (Collins 1914). For wheat

368   varieties reliant on residual soil moisture during early growth, reduced root conductance has

369   been correlated with increased yield (Passioura 1972). Mexican highland maize may similarly

370   benefit from rationing water use early in the season (Fischer et al. 1983; Hayano-Kanashiro et

371   al. 2009) with reduced axial conductance contributing to this adaptive water saving strategy.

372   We observed variation in the formation of root cortical aerenchyma, with varieties

373   sourced from regions of higher precipitation generally being associated with greater

374   aerenchyma formation. Root cortical aerenchyma forms constitutively in wetland crops such

375   as rice and in maize wild relatives endemic to regions of high precipitation (Mano et al. 2007;

376   Mano and Nakazono 2021). Many cultivated maize genotypes lack constitutive aerenchyma;

377   however, aerenchyma formation can be induced by environmental stresses, such as hypoxia

378   (Yamauchi et al. 2016), drought (Zhu et al. 2010), heat (Hu et al. 2014) or nutrient starvation

379   (Saengwilai et al. 2014; Galindo-Castañeda et al. 2018). Although greenhouse evaluation was

380   conducted in benign condition, substantial aerenchyma production was observed (11% Jala,

381   this study; 16% PI586644 in Burton et al. 2013) of the total cortical area in individual

382   sections. In the field, aerenchyma plays a role in oxygenation of the root tissue under hypoxia

383   (Jackson et al. 1985; Colmer 2003), while, in resource-limited conditions, the reduction in

384   root metabolic cost resulting from aerenchyma formation may enhance the efficiency of

385   foraging in terms of carbon invested (Klein et al. 2020; Lynch et al. 2021). On the other hand,

386   with fewer living cortical cells a plant may be less able to accommodate mutualistic

387   arbuscular mycorrhizal fungi, although the relationships between root anatomy, microbial

388   interactions, environment and cortical burden remain to be fully understood (Saengwilai et al.

389   2014; Galindo-Castañeda et al. 2018; Strock et al. 2019). In our predictive analysis high

390   *aerenchyma area* was associated with the varieties from the Gulf coast, the Yucatan

391   (exemplified by Nal Tel in our greenhouse evaluation) and the region around the southern

392   Mexican border extending into Guatemala. This last region covers the native range of the

393   flooding tolerant teosinte *Zea mays* ssp. *huehuetenangensis* (Mano et al. 2005) and suggests

394   flooding may have also exerted a selective pressure on the endemic maize.

395   The functional impact of root anatomical variation is contingent on root system

396   architecture and, indeed, overall plant phenology (Lynch 2019). Differences in growth angle

397   and branching determine the deployment of roots across the soil profile and will interact,

398   synergistically or antagonistically, with root anatomy to impact overall root function. For

399   example, the water-banking effect of reduced axial conductance discussed above has been

400   shown to be enhanced in the context of a shallow root system architecture, enhancing the

401   performance of inbred maize under drought (Strock et al. 2021). While there is a scarcity of

402   information concerning root system architecture in Mexican maize, the limited data reveal

403   remarkable structural diversity, indicating strong spatio-temporal variation in soil exploration

404   (Heymans 2022). It has been noted that native maize root systems tend to be generally

405   shallower compared to those of inbred lines (Burton et al. 2013; Ren et al. 2022).

406   Interestingly, the highland varieties we found associated with reduced axial conductance have

407   previously been described to have a high tendency to lodge (fall over) due to "poorly

408   developed" root systems (Wellhausen et al. 1952). In practice, traditional management

409   involves pilling of earth around the growing plant, freeing the root system from the need to

410   provide mechanical support and perhaps allowing an overall reduction in root system

411   development that contributes to water-banking.

412   In summary, our analyses indicate that reported variation in Mexican native maize

413   root anatomy is distributed systematically over the environment, consistent with a role in

414   local adaptation. We propose that predictive models based on a set of "signpost" accessions

415   can define biologically relevant clines though complex environments, providing the

416   appropriate axes against which to identify both phenotypic and genetic trends. Significantly,

417 we obtained candidate genes from our predicted trait GWA that were not identified in one-

418 by-one analyses of environmental descriptors, including candidates whose functional role in

419 root anatomy is supported by previous studies of inbred maize (Fig. 5; Klein et al. 2020;

420 Schneider et al. 2020). The combined use of field evaluation and *in silico* modeling has

421 allowed great progress to be made in defining the functional impact of root anatomical

422 variation (Heymans et al. 2020; Lynch et al. 2021; Sidhu et al. 2023). The further study of

423 locally adapted native varieties has the potential to complement these other approaches. The

424 history of native crop diversity is a natural experiment that has run for thousands of years,

425 selection imposed by environmental conditions being integrated over many generations. As

426 such, subtle signals that can be hard to detect in experimental evaluation may be amplified

427 and detected as patterns of GEA.

428

429 **MATERIALS AND METHODS**

430 *Phenotyped Burton panel*

431 Phenotypic data from previous characterizations of greenhouse-grown native Mexican maize

432 were obtained from Burton et al. 2013. After filtering for accessions of Mexican origin,

433 subsequent analyses were completed with data from 39 georeferenced individuals. Additional

434 root anatomical features including *total metaxylem vessel area* (MVA), *individual metaxylem*

435 *diameter* (MD), *individual metaxylem area* (MA), *number metaxylem vessels* (NMV), and

436 *cortical cell size* (CCS) were measured from the original Burton et al. cross-section images

437 generated using *RootScan* v2.4, an imaging software designed to measure anatomical features

438 of root cross-sections from digital images (Burton et al. 2012).

439

440 *Genotyped CIMMyT FOAM panel*

441 Genotypes from a collection of 1791 native Mexican maize accessions from the CIMMyT

442 Maize Germplasm Bank (CIMMyT panel) were obtained from (Navarro *et al.* 2017; Gates *et*

443 *al.* 2019). In brief, sequences were generated using an Illumina HiSeq, and genotypes were

444 called in TASSEL. Missing SNPs were imputed using BEAGLE4, and SNPs were further

445 filtered for minor allele frequency >1%. The genotype data was uplifted to coordinates on the

446 B73 v4 reference genome using Crossmap.

447

448 *Environmental Data*

449

450 We compiled climatic and soil data for each representative of the long-term averages

451 experienced by an accession's point of origin for both the Burton and CIMMyT panel. All

452 data used was sourced from publicly available sources with global coverage. Climate data

453 was extracted using R/raster::extract (Hijmans 2023) following the methods described in

454 (Lasky et al. 2015). Briefly, the first set of climate variables come from WorldClim and

455 include information on monthly minimum, maximum, and mean temperatures; mean monthly

456 precipitation; and other derived parameters of biological importance that take into account

457 temperature and precipitation dynamics (Hijmans et al. 2005). Monthly and annual average

458 potential evapotranspiration (PET), and a measure of aridity (mean annual precipitation

459 divided by mean annual PET) that is calculated from WorldClim data were collected from the

460 CGIAR-CSI Globality-Arbitdty database (Zomer et al. 2008). Information on inter-annual

461 variability in precipitation, which may representative of areas where drought acclimation is

462 important (Lasky et al. 2012) were calculated with data from the NCEP/NCAR Reanalysis

463 project (https://psl.noaa.gov/data/reanalysis/reanalysis.shtml; Kalnay et al. 1996). Inter-

464 annual variability in precipitation was obtained by calculating each calendar month's

465 coefficient of variation (CV) across years for each month's surface precipitation rate.

466 Information on estimated photosynthetically active radiation (PAR) for each quarter were

467   averaged for data collected from NASA SRB (https://asdc.larc.nasa.gov/project/SRB).

468   Dynamics of evaporative demand on individuals in the form of vapor pressure deficit (VPD),

469   or the difference between partial pressure of water vapor and maximum potential pressure

470   was collected from the Climate Research Unit (New et al. 2002). In addition to climate data,

471   we also included edaphic chemical and physical properties representative of the long term

472   averages experienced by accessions. The soil data was collected from two sources, SoilGrids

473   (Hengl et al. 2017) and the Global Soil Dataset (GSD; Shangguan et al. 2014). Data from

474   GSD includes soil features of the topsoil and 1 meter below the surface. We found high

475   concordance of values for topsoil and 1 meter below the surface and excluded the topsoil data

476   from our dataset. All soil variables were cleaned by removing outliers and imputed missing

477   values using the MICE package (van Buuren and Groothuis-Oudshoorn 2011; Fox et al.

478   2017).

479

480   *GRANAR representations and MECHA estimation of emergent hydraulic properties*

481   Generator of Root Anatomy in R (GRANAR; https://granar.github.io), and the model of

482   explicit cross-section hydraulic architecture (MECHA; https://mecharoot.github.io) are open-

483   sourced computational tools. The first tool uses anatomical parameters as inputs to generate

484   digital root anatomies. Once constructed, GRANAR root anatomies can be used for digital

485   visualizations of anatomical parameters, and the anatomical network can be written as a XML

486   file with the same format as CellSet output (Pound et al. 2012). The second tool uses the

487   anatomical networks, such as the one generated by GRANAR to estimate emergent hydraulic

488   properties. We used GRANAR to reconstruct virtual anatomies for all observed accessions of

489   the Burton panel, our predictions of the Burton panel Mexican lines, predicted CIMMyT

490   clusters, and predicted and observed novel germplasm grown in this study. For all

491   aforementioned individuals, we estimated the root hydraulic conductance (K_r and K_x) and

492   conductivity (k_r) with MECHA. The subcellular hydraulic parameters are the same as in

493   (Heymans et al. 2020), and the chosen hydraulic scenario accounts for the hydrophobic

494   structures of an endodermal Casparian strip. The script used is available on a GitHub

495   repository HydraulicViper/RootDiversity (doi: 10.5281/zenodo.10104521) under a GPL-3

496   license.

497

498   As not all anatomical features required for the GRANAR-MECHA pipeline were predictable

499   with our RF models, a few transformations were required for predicted anatomical data to be

500   input into GRANAR. With the exception of traits where RF predictions could directly be

501   used as inputs (*number of metaxylem vessels, metaxylem vessel area*, *aerenchyma area*), we

502   used constant values of the mean Burton panel (*root cross-section area*) or extrapolated

503   values from RF predictions (*total stele area* calculated from RF predicted *total stele*

504   *area:root cross-section area* using the Burton panel *root cross-section area* mean).

505

506   *Principal component analysis/initial trait/env association*

507   We used principal component analysis (PCA) to initially explore internal anatomical

508   variation captured in the greenhouse grown Burton panel. Phenotypes were constrained to

509   only include "pure" traits, excluding proportional and percentage traits which are likely

510   redundant and may be representative of more emergent properties. Principal components

511   were built with R/ade4::dudi.pca and visualized with R/factoextra. We first compared

512   phenotypic PC axes to calculated hydraulic properties to explore the combinations of

513   anatomical traits that are related to variation in hydraulic properties. As an initial attempt to

514   identify relationships between root anatomical traits and environmental variation, we

515   compared the first three phenotypic PCs loadings and calculated hydraulic properties to core

516   environmental features of the accessions' point of origin (elevation, annual precipitation,

517   annual mean temperature, soil pH). For all environmental, hydraulic, and phenotypic PC

518   correlations, p-values were adjusted for multiple testing using the Holm method with

519   R/stats::p.adjust ($\alpha = 0.05$). As correlation between both root anatomical traits and derived

520   hydraulic properties and environment of origin were weak, we then considered if combined

521   environmental features were more able to describe variation in root anatomical traits and

522   hydraulic properties.

523

524   *Feature Selection*

525   We sought to determine if aspects of accessions' home environment predict variation in root

526   anatomical traits using a machine learning approach. Importantly, not all variables in the

527   environmental dataset are related to root anatomical variation and not all tested root anatomy

528   traits are significantly associated with variation in environmental features. Feature selection

529   was employed to identify the anatomical traits that had relationships with environmental

530   features ("environmentally related traits" from here on) and the environmental descriptors

531   that described variation in those traits through eliminating unimportant variables. We used

532   the function R/Boruta::boruta to obtain all important and tentatively important environmental

533   features for each trait in our dataset (Kursa and Rudnicki 2010). Root anatomical traits which

534   had significant variation described by at least two environmental features were considered

535   environmentally related and retained for further analysis.

536

537   *Random Forest Models*

538

539   We employed random forest (RF) to determine if variation in response variables (observed

540   anatomical traits of the Burton panel) could be described by several explanatory variables

541   (feature-selected environmental descriptors). For each environmentally related trait, we built

542   a RF model that summarized how trait values are predicted to change across feature-selected

543   environmental space. RF is a non-parametric classification method that constructs decision

544   trees using subsets of input data to select the predictor variables that limit variance for the

545   response variable predictive model (Breiman 2001). RF models have been found to have high

546   predictive performance when tested data include a large number of predictor variables with

547   little to no relationship to response variables (Fox et al. 2017); however, best practice is to

548   limit the number of predictor variables to avoid model overfitting. As such, we limited

549   response variables for each environmentally related trait model to be the specific

550   environmental features identified in the Boruta feature selection step. As the environmental

551   variables used in this study were continuous, RF models were built as regression trees. RF

552   models were built using R/randomForest::randomForest, 5000 trees were built per model and

553   one third the number of explanatory variables were tried at each split (Liaw and Wiener

554   2002). We increased our number of trees from the default value (500), to account for models

555   with a large number of predictors and for increased stability of variable importance. Model

556   success was evaluated with the percent variance explained output extracted from the

557   randomForest package and the correlation coefficient between observed and RF predicted

558   trait values. To determine the contribution of each boruta-identified environmental descriptor

559   for constructed RF models, we calculated SHapley Additive exPlanations (SHAP) values

560   (Lundberg and Lee 2017).

561

562   *Predicting traits using environmental relationships*

563

564   Using the constructed RF models, unknown phenotypes of 1791 georeferenced, genotyped

565   CIMMyT accessions (CIMMyT panel) were predicted from environmental descriptors of

566   accession point of origin. Using R/caret::predict (Kuhn 2021), values for all nine

567   environmentally associated traits were predicted for the 1791 CIMMyT panel. For each

568   trained RF model, the full environmental dataset summarizing the source environment of the

569   CIMMyT panel was constrained to include the environmental descriptors used to train each

570   RF model. These constrained environmental descriptors were used as inputs in each trained

571   RF model to calculate environmentally related traits for all individuals of the genotyped

572   panel.

573

574   *Clustering analysis*

575   CIMMyT FOAM accessions were clustered by predicted root anatomical traits using

576   Partitioning Around Medoids (PAM) as previously described (Klein et al. 2020). Briefly,

577   anatomical trait values were centered and scaled using R/caret (Kuhn 2021) and outlying

578   values (> 3 standard deviations from the mean) removed. Within cluster sums of squares

579   (WSS) were visualized as a function of cluster number using R/factoextra::fviz_nbclust

580   (Kassambara and Mundt 2020). From inspection of the resulting curve, the accessions were

581   grouped into seven clusters using R/cluster::pam (Maechler et al. 2021) under default

582   settings. Primary variety (landrace) designations were assigned using data available from

583   CIMMyT (www.mgb.cimmyt.org; 1454 accessions assigned), and these matched to existing

584   morphological-isozymatic (Sanchez G. et al. 2000) and environmental (Ruiz Corral et al.

585   2008) classifications. Testing for enrichment of a given variety in a given cluster was

586   performed using Fisher tests with R/stats::fisher.test, under a contingency table formed by

587   partitioning the 1454 accessions by membership of the cluster and assignment to the variety.

588   Results were adjusted using the Holm method with R/stats::p.adjust ($\alpha = 0.05$).

589

590   *Greenhouse evaluation of root anatomy in selected accessions*

591   We selected eight novel native Mexican maize accessions, representative of environmental

592   diversity within the CIMMyT panel for phenotypic validation of RF anatomical predictions.

593   Ten biological replicates of each accession were grown in a greenhouse in State College, PA

594   (40.8028708, -77.8640406) from April to May of 2022. Plants were grown in 2.83 L pots (4

595   in x 4 in x 14 in, Greenhouse Megastore). The growth media was a mix of silica sand (50%),

596   turface (30%), and field soil (20%) sourced from Rock Springs, PA. Pots were watered to

597   field capacity the night before planting and watered every day after sowing until germination.

598   Once germinated, plants were watered every other day. One week after germination, plants

599   were fertigated with Peters Excel 15 - 5 - 15 Cal Mag Special with Black Iron 200 ppm N

600   recipe and supplemented with an extra 5 ppm Fe (Sprint 330), fed at 1:100 dilution, two times

601   per week until harvest. Greenhouse settings were set at 16 hour days, with a minimum

602   temperature of 21 degrees C and a maximum temperature of 28 degrees C.

603   Following methods from Burton et al. 2013, 28 days after planting, plants were destructively

604   harvested. Two representative axial roots from nodes two and three were collected. From

605   each axial root, a 4-cm root sample was excised five to nine cm from the most basal portion

606   of the sample. Root samples were stored in 75% ethanol until sectioned by laser ablation

607   tomography (LAT; Strock et al. 2022). In LAT, a sample is moved via an automatic stage

608   towards a 355-nm Avia 7000 pulsed laser and ablated in the focal plane of a camera. A

609   Canon T3i camera with a 53micro lens (MP-E 65 mm) was used to capture images of the root

610   cross-section. Two representative images for each root sample sectioned 1 to 3 cm apart were

611   saved for later image analysis with _RootScan_. Anatomical phenotypes were averaged for each

612   nodal root of a plant, where each value is an average of two roots from each node and two

613   LAT image sections of each root.

614

615   *Estimation of genotypic effects on anatomical traits*

616   To estimate the effect of landrace genotype (race designation) on each measured trait across

617   growth stages (nodes), we used linear mixed models to calculate the best linear unbiased

618   predictions (BLUPs) with the equation:

619    $$Trait_{rwt} = \beta_0 + R_{0r} + N_{0n} + T_{0t} + e_{rnt}$$

620   Where for each modeled trait, $Trait_{rwt}$, β_0 is the overall intercept, R_{0r} is the random effect of

621   plant race, N_{0n} is the random effect of node, T_{0t} is the random effect of the tray the plant was

622   grown in, and e_{rwt} is the error term. BLUPs were calculated for each measured anatomical

623   trait of a given landrace and extracted using R/lme4::ranef.

624

625   *Validation of multivariate fit of trait predictions and observations*

626   We used procrustes analysis to determine concordance between all predicted RF anatomical

627   traits and observed BLUPs. The procrustes analysis relates the overall shape of two sets of

628   multivariate matrices by minimizing the total distance between the two distributions and

629   quantifying how much the relationship between variables in the matrices differ after this

630   alignment (Goodall 1991). The algorithm was implemented using the function

631   R/vegan::procrustes (Oksanen et al. 2022).

632

633   *Genome-Wide Association in the Wisconsin Diversity Panel*

634   We used genetic and phenotypic data for 175 inbred maize lines from the expanded

635   Wisconsin Diversity Panel (Schneider et al. 2020). We filtered published SNP data for minor

636   allele frequency >5%, resulting in a total of 370,991 SNPs. Thirteen root anatomical traits

637   were extracted from the published study: MA, MD, MVA, NMV, RXSA, TCA,TSA, AA,

638   X.A, CCFN, CCS, TSA.RXSA, TSA.TCA. The full design and experimental protocol are

639   described in (Schneider et al. 2020). Briefly, maize genotypes were grown in two replicates

640   under well-watered and water-stressed conditions in a randomized complete block design.

641   Plots were irrigated with a center pivot system, and water stress was applied 4 weeks after

642   planting. At anthesis, one representative plant from each plot was excavated from the soil

643   using a standard shovel. Root crowns were soaked and washed to remove soil particles. A

644   representative fourth node root (5-8 cm from the base of the root) was excised and imaged

645   and phenotyped for root anatomical traits using LAT and *RootScan* software. We fitted a

646   linear mixed effect model using R/lme4 (Bates et al. 2007) for the well-watered and water-

647   stressed condition with overall mean as the fixed effect and genotype and block as random

648   effects and extracted BLUPs for genotypes with the R/lme4::ranef function. Broad sense

649   heritability for each root trait was estimated as the genotype variance divided by the sum of

650   genotype variance and error variance from linear mixed effect models. Root trait BLUPs

651   were used to fit linear mixed models using R/GridLMM, a package for fitting linear mixed

652   models with multiple random effects (Runcie & Crawford 2018). We used the function

653   R/GridLMM::GridLMM_GWAS to run the GWA study and set the environmental vector to -

654   1 or 1 in the model to represent the water-stressed and well-watered treatments. The p-values

655   for the genotype main effect and the genotype by environment interaction effect were

656   calculated using Wald tests. The SNP level P-values were combined into the gene level

657   associations using Multi-marker Analysis of GenoMic Annotation (MAGMA) (de Leeuw et

658   al. 2015). MAGMA uses a multiple regression model to aggregate all SNP information into a

659   gene while accounting for linkage disequilibrium (LD). SNPs were annotated to genes using

660   a 2.5 kilobase window around each gene, resulting in 24,099 genes.

661

662   *Environmental GWA*

663   We performed MAGMA to measure the gene-level associations between CIMMyT genotypes

664   and selected environmental variables of their native locations. The first five eigenvectors of

665   the genetic relationship matrix were included in the model to control for population structure.

666   SNPs were annotated to genes using a 2.5 kilobase window around each gene. The final

667   dataset contained 1656 genotypes and 28898 genes for CIMMyT panel accessions.

668

669   *Common genes shared between WIDP and CIMMyT*

670   After gene annotation, we obtained 21883 genes that are shared by the WIDP panel and the

671   CIMMyT panel. To evaluate if genes that are highly associated with root anatomical traits

672   also showed associations with our predicted root traits and environmental variables, we

673   extracted and pooled candidate genes from the top 100 genes for all WIDP root traits,

674   predicted root traits, and related environmental variables identified by MAGMA. The final

675   gene list contains WIDP root anatomical genes (576 genotype main effect genes and 542

676   WIDP genotype x treatment interaction genes), 636 RF predicted anatomical genes, and 1282

677   environmental genes.

678

679   *Minor allele frequency*

680   To understand the relationship between allele variation, environment, and root traits, we

681   extracted the genotypic information of top SNPs of the target genes. We divided maize

682   landraces into PAM clusters, and calculated the mean elevation, the minor allele frequencies

683   (MAF) of the target SNPs, and the mean predicted root traits for each cluster. Pearson

684   correlation was conducted to test the correlations between MAF and elevation, and between

685   MAF and predicted root traits.

686

687   **ACKNOWLEDGEMENTS**

697

AUTHOR CONTRIBUTIONS

699 R.J.H.S., C.M.M., and M.L. conceived and designed the analysis. C.M.M. built models used

700 to predict root anatomy; M.L. and R.J.H.S. performed the genetic analysis; A.H. constructed

701 digital anatomies and estimated hydraulic components; C.M.M, M.L., M.P., and R.J.H.S.

702 grew and harvested the pot experiment; C.M.M. obtained anatomical characterizations for

703 novel material using LAT. C.M.M., R.J.H.S., H.S , M.L, and J.L. wrote the paper. All authors

704 read and approved the final version of the manuscript.

705

REFERENCES

707 Aguilar-Rangel MR, Chávez Montes RA, González-Segovia E, Ross-Ibarra J, Simpson
708 JK, Sawers RJH. 2017. Allele specific expression analysis identifies regulatory variation
709 associated with stress-related genes in the Mexican highland maize landrace Palomero
710 Toluqueño. *PeerJ* 5:e3737.

711 Allah AA, Abd Allah A, Badawy S, Zayed B, El-Gohary A. 2010. THE ROLE OF
712 ROOT SYSTEM TRAITS IN THE DROUGHT TOLERANCE OF RICE (Oryza sativa
713 L.). *Journal of Plant Production* [Internet] 1:621–631. Available from:

714   http://dx.doi.org/10.21608/jpp.2010.86384

715   Arteaga MC, Moreno-Letelier A, Mastretta-Yanes A, Vázquez-Lobo A, Breña-Ochoa A,
716   Moreno-Estrada A, Eguiarte LE, Piñero D. 2016. Genomic variation in recently
717   collected maize landraces from Mexico. *Genom Data* 7:38–45.

718   Atkinson JA, Rasmussen A, Traini R, Voß U, Sturrock C, Mooney SJ, Wells DM,
719   Bennett MJ. 2014. Branching out in roots: uncovering form, function, and regulation.
720   *Plant Physiol.* 166:538–550.

721   Bates D, Sarkar D, Bates MD, Matrix L. 2007. The lme4 package. *R package version*
722   2:74.

723   Bayuelo-Jiménez JS, Gallardo-Valdéz M, Pérez-Decelis VA, Magdaleno-Armas L,
724   Ochoa I, Lynch JP. 2011. Genotypic variation for root traits of maize (Zea mays L.)
725   from the Purhepecha Plateau under contrasting phosphorus availability. *Field Crops*
726   *Research* [Internet] 121:350–362. Available from:
727   http://dx.doi.org/10.1016/j.fcr.2011.01.001

728   Bellon MR, Mastretta-Yanes A, Ponce-Mendoza A, Ortiz-Santamaría D, Oliveros-
729   Galindo O, Perales H, Acevedo F, Sarukhán J. 2018. Evolutionary and food supply
730   implications of ongoing maize domestication by Mexican. *Proc. Biol. Sci.* [Internet]
731   285. Available from: http://dx.doi.org/10.1098/rspb.2018.1049

732   Bennett AE, Groten K. 2022. The Costs and Benefits of Plant-Arbuscular Mycorrhizal
733   Fungal Interactions. *Annu. Rev. Plant Biol.* 73:649–672.

734   Bomfim NN, Graciano-Ribeiro D, Nassar NMA. 2011. Genetic diversity of root
735   anatomy in wild and cultivated Manihot species. *Genetics and Molecular Research*
736   [Internet] 10:544–551. Available from: http://dx.doi.org/10.4238/vol10-2gmr1093

737   Breiman L. 2001. Random Forests. *Mach. Learn.* 45:5–32.

738   Burton AL, Brown KM, Lynch JP. 2013. Phenotypic diversity of root anatomical and
739   architectural traits in Zea species. *Crop Sci.* 53:1042–1055.

740   Burton AL, Williams M, Lynch JP, Brown KM. 2012. RootScan: Software for high-
741   throughput analysis of root anatomical traits. *Plant Soil* 357:189–203.

742   van Buuren S, Groothuis-Oudshoorn K. 2011. mice: Multivariate Imputation by Chained
743   Equations inR. *J. Stat. Softw.* [Internet] 45. Available from:
744   http://www.jstatsoft.org/v45/i03/

745   Chen C, DeClerck G, Tian F, Spooner W, McCouch S, Buckler E. 2012. PICARA, an
746   analytical pipeline providing probabilistic inference about a priori candidates genes
747   underlying genome-wide association QTL in plants. *PLoS One* 7:e46596.

748   Chen Z, Sun J, Li D, Li P, He K, Ali F, Mi G, Chen F, Yuan L, Pan Q. 2022. Plasticity
749   of root anatomy during domestication of a maize-teosinte derived population. *J. Exp.*
750   *Bot.* 73:139–153.

751   Chimungu JG, Loades KW, Lynch JP. 2015. Root anatomical phenes predict root

penetration ability and biomechanical properties in maize (Zea Mays). *Journal of Experimental Botany* [Internet] 66:3151–3162. Available from: http://dx.doi.org/10.1093/jxb/erv121

Collins GN. 1914. Pueblo Indian maize breeding. *J. Hered.* 5:255–268.

Colmer TD. 2003. Long-distance transport of gases in plants: a perspective on internal aeration and radial oxygen loss from roots. *Plant, Cell & Environment* [Internet] 26:17–36. Available from: http://dx.doi.org/10.1046/j.1365-3040.2003.00846.x

Couvreur V, Faget M, Lobet G, Javaux M, Chaumont F, Draye X. 2018. Going with the Flow: Multiscale Insights into the Composite Nature of Water Transport in Roots. *Plant Physiol.* 178:1689–1703.

Couvreur V, Vanderborght J, Javaux M. 2012. A simple three-dimensional macroscopic root water uptake model based on the hydraulic architecture approach. *Hydrol. Earth Syst. Sci.* 16:2957–2971.

Des Marais DL, Hernandez KM, Juenger TE. 2013. Genotype-by-environment interaction and plasticity: Exploring genomic responses of plants to the abiotic environment. *Annu. Rev. Ecol. Evol. Syst.* 44:5–29.

Doussan C, Vercambre G, Pagè LC. 1998. Modelling of the Hydraulic Architecture of Root Systems: An Integrated Approach to Water Absorption–Distribution of Axial and Radial Conductances in Maize. *Ann. Bot.* 81:225–232.

Eagles HA, Lothrop JE. 1994. Highland maize from central Mexico—its origin, characteristics, and use in breeding programs. *Crop Sci.* 34:11–19.

Feng W, Lindner H, Robbins NE 2nd, Dinneny JR. 2016. Growing Out of Stress: The Role of Cell- and Organ-Scale Growth Control in Plant Water-Stress Responses. *Plant Cell* 28:1769–1782.

Fischer KS, Johnson EC, Edmeades GO. 1983. Breeding and Selection for Drought Resistance in Tropical Maize. In: Centro Internacional de Mejoramiento de Maiz y Trigo (Cimmyt).

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM. 2011. A map of local adaptation in Arabidopsis thaliana. *Science* 334:86–89.

Fox EW, Hill RA, Leibowitz SG, Olsen AR, Thornbrugh DJ, Weber MH. 2017. Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology. *Environ. Monit. Assess.* 189:316.

Frensch J, Steudle E. 1989. Axial and Radial Hydraulic Resistance to Roots of Maize (Zea mays L.). *Plant Physiol.* 91:719–726.

Fumagalli M, Sironi M, Pozzoli U, Ferrer-Admettla A, Pattini L, Nielsen R. 2011. Signatures of Environmental Genetic Adaptation Pinpoint Pathogens as the Main Selective Pressure through Human Evolution. *PLoS Genet.* 7:e1002355.

Galindo-Castañeda T, Brown KM, Lynch JP. 2018. Reduced root cortical burden

improves growth and grain yield under low phosphorus availability in maize. *Plant Cell Environ.* 41:1579–1592.

Gaudin ACM, McClymont SA, Soliman SSM, Raizada MN. 2014. 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.

Goodall C. 1991. Procrustes methods in the statistical analysis of shape. *J. R. Stat. Soc.* 53:285–321.

G Viana W, Scharwies JD, Dinneny JR. 2022. Deconstructing the root system of grasses through an exploration of development, anatomy and function. *Plant Cell Environ.* 45:602–619.

Hayano-Kanashiro C, Calderón-Vázquez C, Ibarra-Laclette E, Herrera-Estrella L, Simpson J. 2009. Analysis of gene expression and physiological responses in three Mexican maize landraces under drought stress and recovery irrigation. *PLoS One* 4:e7531.

von Heckel K, Stephan W, Hutter S. 2016. Canalization of gene expression is a major signature of regulatory cold adaptation in temperate Drosophila melanogaster. *BMC Genomics* 17:574.

Hengl T, Mendes de Jesus J, Heuvelink GBM, Ruiperez Gonzalez M, Kilibarda M, Blagotić A, Shangguan W, Wright MN, Geng X, Bauer-Marschallinger B, et al. 2017. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One* 12:e0169748.

Hereford J. 2009. A quantitative survey of local adaptation and fitness trade-offs. *Am. Nat.* 173:579–588.

Heymans A. 2022. In silico analysis of the influence of root hydraulic anatomy on maize (Zea mays) water uptake.

Heymans A, Couvreur V, LaRue T, Paez-Garcia A, Lobet G. 2020. GRANAR, a Computational Tool to Better Understand the Functional Importance of Monocotyledon Root Anatomy. *Plant Physiol.* 182:707–720.

Hijmans RJ. 2023. raster: Geographic Data Analysis and Modeling. Available from: https://CRAN.R-project.org/package=raster

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* 25:1965–1978.

Hoban S, Kelley JL, Lotterhos KE, Antolin MF, Bradburd G, Lowry DB, Poss ML, Reed LK, Storfer A, Whitlock MC. 2016. Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *Am. Nat.* 188:379–397.

Hochholdinger F. 2009. The Maize Root System: Morphology, Anatomy, and Genetics. In: Bennetzen JL, Hake SC, editors. Handbook of Maize: Its Biology. New York, NY: Springer New York. p. 145–160.

Hu B, Henry A, Brown KM, Lynch JP. 2014. Root cortical aerenchyma inhibits radial nutrient transport in maize (Zea mays). *Annals of Botany* [Internet] 113:181–189. Available from: http://dx.doi.org/10.1093/aob/mct259

Jackson MB, Fenning TM, Drew MC, Saker LR. 1985. Stimulation of ethylene production and gas-space (aerenchyma) formation in adventitious roots of Zea mays L. by small partial pressures of oxygen. *Planta* [Internet] 165:486–492. Available from: http://dx.doi.org/10.1007/bf00398093

Janzen GM, Aguilar-Rangel MR, Cíntora-Martínez C, Blöcher-Juárez KA, González-Segovia E, Studer AJ, Runcie DE, Flint-Garcia SA, Rellán-Álvarez R, Sawers RJH, et al. 2022. Demonstration of local adaptation in maize landraces by reciprocal transplantation. *Evol. Appl.* 15:817–837.

Jaramillo RE, Nord EA, Chimungu JG, Brown KM, Lynch JP. 2013. Root cortical burden influences drought tolerance in maize. *Ann. Bot.* 112:429–437.

Joswig JS, Wirth C, Schuman MC, Kattge J, Reu B, Wright IJ, Sippel SD, Rüger N, Richter R, Schaepman ME, et al. 2022. Climatic and soil factors explain the two-dimensional spectrum of global plant trait variation. *Nat Ecol Evol* 6:36–50.

Jung JKH, McCouch S. 2013. Getting to the roots of it: Genetic and hormonal control of root architecture. *Front. Plant Sci.* 4:186.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, et al. 1996. The NCEP/NCAR 40-year reanalysis project. vol. 77, issue 3. *Bull. Am. Meteorol. Soc.*:1520–0477.

Kassambara A, Mundt F. 2020. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. Available from: https://CRAN.R-project.org/package=factoextra

Klein SP, Schneider HM, Perkins AC, Brown KM, Lynch JP. 2020. Multiple Integrated Root Phenotypes Are Associated with Improved Drought Tolerance. *Plant Physiol.* 183:1011–1025.

Kuhn M. 2021. caret: Classification and Regression Training. Available from: https://CRAN.R-project.org/package=caret

Kursa MB, Rudnicki WR. 2010. Feature Selection with theBorutaPackage. *J. Stat. Softw.* [Internet] 36. Available from: http://www.jstatsoft.org/v36/i11/

Kusmec A, Srinivasan S, Nettleton D, Schnable PS. 2017. Distinct genetic architectures for phenotype means and plasticities in Zea mays. *Nat Plants* 3:715–723.

Lasky JR, Des Marais DL, Lowry DB, Povolotskaya I, McKay JK, Richards JH, Keitt TH, Juenger TE. 2014. Natural variation in abiotic stress responsive gene expression and local adaptation to climate in Arabidopsis thaliana. *Mol. Biol. Evol.* 31:2283–2296.

Lasky JR, Des Marais DL, McKay JK, Richards JH, Juenger TE, Keitt TH. 2012. Characterizing genomic variation of Arabidopsis thaliana: the roles of geography and climate. *Mol. Ecol.* 21:5512–5529.

867 Lasky JR, Josephs EB, Morris GP. 2023. Genotype-environment associations to reveal
868 the molecular basis of environmental adaptation. *Plant Cell* 35:125–138.

869 Lasky JR, Upadhyaya HD, Ramu P, Deshpande S, Hash CT, Bonnette J, Juenger TE,
870 Hyma K, Acharya C, Mitchell SE, et al. 2015. Genome-environment associations in
871 sorghum landraces predict adaptive traits. *Sci Adv* 1:e1400218.

872 de Leeuw CA, Mooij JM, Heskes T, Posthuma D. 2015. MAGMA: generalized gene-set
873 analysis of GWAS data. *PLoS Comput. Biol.* 11:e1004219.

874 Leitner D, Meunier F, Bodner G, Javaux M, Schnepf A. 2014. Impact of contrasted
875 maize root traits at flowering on water stress tolerance – A simulation study. *Field*
876 *Crops Res.* 165:125–137.

877 León J, Gayubas B, Castillo M-C. 2020. Valine-Glutamine Proteins in Plant Responses
878 to Oxygen and Nitric Oxide. *Front. Plant Sci.* 11:632678.

879 Levins R. 1968. Evolution in Changing Environments. Available from:
880 http://dx.doi.org/10.1515/9780691209418

881 Liaw A, Wiener M. 2002. Classification and Regression by Randomforest. *R News*
882 2:18–22.

883 Lopez-Valdivia I, Perkins AC, Schneider HM, Vallebueno-Estrada M, Burridge JD,
884 González-Orozco E, Montufar A, Montiel R, Lynch JP, Vielle-Calzada J-P. 2022.
885 Gradual domestication of root traits in the earliest maize from Tehuacán. *Proc. Natl.*
886 *Acad. Sci. U. S. A.* 119:e2110245119.

887 Louette D, Smale M. 2000. *Euphytica* 113:25–41.

888 Lowry DB. 2012. Ecotypes and the controversy over stages in the formation of new
889 species. *Biol. J. Linn. Soc. Lond.* 106:241–257.

890 Lundberg S, Lee S-I. 2017. A Unified Approach to Interpreting Model Predictions.
891 *arXiv [cs.AI]* [Internet]. Available from: http://arxiv.org/abs/1705.07874

892 Lynch JP. 2019. Root phenotypes for improved nutrient capture: an underexploited
893 opportunity for global agriculture. *New Phytol.* 223:548–564.

894 Lynch JP, Strock CF, Schneider HM, Sidhu JS, Ajmera I, Galindo-Castañeda T, Klein
895 SP, Hanlon MT. 2021. Root anatomy and soil resource capture. *Plant Soil* 466:21–63.

896 Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. 2021. cluster: Cluster
897 Analysis Basics and Extensions. Available from: https://CRAN.R-
898 project.org/package=cluster

899 Magalhaes JV, Liu J, Guimarães CT, Lana UGP, Alves VMC, Wang Y-H, Schaffert RE,
900 Hoekenga OA, Piñeros MA, Shaff JE, et al. 2007. A gene in the multidrug and toxic
901 compound extrusion (MATE) family confers aluminum tolerance in sorghum. *Nat.*
902 *Genet.* 39:1156–1161.

903 Mano Y, Muraki M, Fujimori M, Takamizo T, Kindiger B. 2005. Identification of QTL

controlling adventitious root formation during flooding conditions in teosinte (Zea mays ssp. huehuetenangensis) seedlings. *Euphytica* [Internet] 142:33–42. Available from: http://dx.doi.org/10.1007/s10681-005-0449-2

Mano Y, Nakazono M. 2021. Genetic regulation of root traits for soil flooding tolerance in genus Zea. *Breed. Sci.* 71:30–39.

Mano Y, Omori F, Takamizo T, Kindiger B, McK. Bird R, Loaisiga CH, Takahashi H. 2007. QTL mapping of root aerenchyma formation in seedlings of a maize × rare teosinte "Zea nicaraguensis" cross. *Plant and Soil* [Internet] 295:103–113. Available from: http://dx.doi.org/10.1007/s11104-007-9266-9

Markesteijn L, Poorter L. 2009. Seedling root morphology and biomass allocation of 62 tropical tree species in relation to drought- and shade-tolerance. *J. Ecol.* 97:311–325.

Ma Z, Guo D, Xu X, Lu M, Bardgett RD, Eissenstat DM, Luke McCormack M, Hedin LO. 2018. Evolutionary history resolves global organization of root functional traits. *Nature* [Internet] 555:94–97. Available from: http://dx.doi.org/10.1038/nature25783

Mercer KL, Perales H. 2019. Structure of local adaptation across the landscape: flowering time and fitness in Mexican maize (Zea mays L. subsp. mays) landraces. *Genetic Resources and Crop Evolution* [Internet] 66:27–45. Available from: http://dx.doi.org/10.1007/s10722-018-0693-7

M New, Lister D, Hulme M, Makin I. 2002. A high-resolution data set of surface climate over global land areas. *Clim. Res.* 21:1–25.

Nardini A, Battistuzzo M, Savi T. 2013. Shoot desiccation and hydraulic failure in temperate woody angiosperms during an extreme summer drought. *New Phytol.* 200:322–329.

Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Solymos P, Stevens MHH, Szoecs E, et al. 2022. vegan: Community Ecology Package. Available from: https://CRAN.R-project.org/package=vegan

Pace J, Gardner C, Romay C, Ganapathysubramanian B, Lübberstedt T. 2015. Genome-wide association analysis of seedling root development in maize (Zea mays L.). *BMC Genomics* 16:47.

Passioura JB. 1972. The effect of root geometry on the yield of wheat growing on stored water. *Aust. J. Agric. Res.* 23:745.

Perales H, Golicher D. 2014. Mapping the diversity of maize races in Mexico. *PLoS One* 9:e114657.

Pound MP, French AP, Wells DM, Bennett MJ, Pridmore TP. 2012. CellSeT: novel software to extract and analyze structured networks of plant cells from confocal images. *Plant Cell* 24:1353–1361.

Priatama RA, Heo J, Kim SH, Rajendran S, Yoon S, Jeong D-H, Choo Y-K, Bae JH, Kim CM, Lee YH, et al. 2022. Narrow Metaxylems Enhance Drought Tolerance and Optimize Water Use for Grain Filling in Dwarf Rice. *Front. Plant Sci.* 13:894545.

943 Purushothaman R, Zaman-Allah M, Mallikarjuna N, Pannirselvam R, Krishnamurthy L,
944 Gowda CLL. 2013. Root Anatomical Traits and Their Possible Contribution to Drought
945 Tolerance in Grain Legumes. *Plant Production Science* [Internet] 16:1–8. Available
946 from: http://dx.doi.org/10.1626/pps.16.1

947 Ren W, Zhao L, Liang J, Wang L, Chen L, Li P, Liu Z, Li X, Zhang Z, Li J, et al. 2022.
948 Genome-wide dissection of changes in maize root system architecture during modern
949 breeding. *Nat Plants* 8:1408–1422.

950 Richards RA, Passioura JB. 1989. A breeding program to reduce the diameter of the
951 major xylem vessel in the seminal roots of wheat and its effect on grain yield in rain-fed
952 environments. *Australian Journal of Agricultural Research* [Internet] 40:943. Available
953 from: http://dx.doi.org/10.1071/ar9890943

954 Romero Navarro JA, Willcox M, Burgueño J, Romay C, Swarts K, Trachsel S, Preciado
955 E, Terron A, Delgado HV, Vidal V, et al. 2017. A study of allelic diversity underlying
956 flowering-time adaptation in maize landraces. *Nat. Genet.* 49:476–480.

957 Ruiz Corral JA, Durán Puga N, Sánchez González J de J, Ron Parra J, González
958 Eguiarte DR, Holland JB, Medina García G. 2008. Climatic adaptation and ecological
959 descriptors of 42 Mexican maize races. *Crop Sci.* 48:1502–1512.

960 Runcie DE, Crawford L. 2019. Fast and flexible linear mixed models for genome-wide
961 genetics. *PLoS Genet.* 15:e1007978.

962 Saengwilai P, Nord EA, Chimungu JG, Brown KM, Lynch JP. 2014. Root Cortical
963 Aerenchyma Enhances Nitrogen Acquisition from Low-Nitrogen Soils in Maize. *PLANT
964 PHYSIOLOGY* [Internet] 166:726–735. Available from:
965 http://dx.doi.org/10.1104/pp.114.241711

966 Sanchez G. JJ, Goodman MM, Stuber CW. 2000. Isozymatic and morphological
967 diversity in the races of maize of Mexico. *Econ. Bot.* 54:43–59.

968 Sawers RJH, Gutjahr C, Paszkowski U. 2008. Cereal mycorrhiza: an ancient symbiosis
969 in modern agriculture. *Trends Plant Sci.* 13:93–97.

970 Schneider HM, Klein SP, Hanlon MT, Kaeppler S, Brown KM, Lynch JP. 2020. Genetic
971 control of root anatomical plasticity in maize. *Plant Genome* 13:e20003.

972 Schneider HM, Postma JA, Wojciechowski T, Kuppe C, Lynch JP. 2017. Root Cortical
973 Senescence Improves Growth under Suboptimal Availability of N, P, and K. *Plant
974 Physiol.* 174:2333–2347.

975 Schneider HM, Wojciechowski T, Postma JA, Brown KM, Lücke A, Zeisler V,
976 Schreiber L, Lynch JP. 2017. Root cortical senescence decreases root respiration,
977 nutrient content and radial water and nutrient transport in barley. *Plant Cell Environ.*
978 40:1392–1408.

979 Schönemann PH. 1966. A generalized solution of the orthogonal procrustes problem.
980 *Psychometrika* 31:1–10.

981 Shangguan W, Dai Y, Duan Q, Liu B, Yuan H. 2014. A global soil data set for earth

982 system modeling. *J. Adv. Model. Earth Syst.* 6:249–263.

983 Shimono Y, Watanabe M, Hirao AS, Wada N, Kudo G. 2009. Morphological and
984 genetic variations of Potentilla matsumurae (Rosaceae) between fellfield and snowbed
985 populations. *Am. J. Bot.* 96:728–737.

986 Sidhu JS, Ajmera I, Arya S, Lynch JP. 2023. RootSlice - a novel functional-structural
987 model for root anatomical phenotypes. *Plant Cell Environ.* [Internet]. Available from:
988 http://dx.doi.org/10.1111/pce.14552

989 Song W, Zhao H, Zhang X, Lei L, Lai J. 2015. Genome-Wide Identification of VQ
990 Motif-Containing Proteins and their Expression Profiles Under Abiotic Stresses in
991 Maize. *Front. Plant Sci.* 6:1177.

992 Stebbins GL Jr. 1952. Aridity as a stimulus to plant evolution. *Am. Nat.* 86:33–44.

993 Stinchcombe JR, Weinig C, Ungerer M, Olsen KM, Mays C, Halldorsdottir SS,
994 Purugganan MD, Schmitt J. 2004. A latitudinal cline in flowering time in Arabidopsis
995 thaliana modulated by the flowering time gene FRIGIDA. *Proc. Natl. Acad. Sci. U. S. A.*
996 101:4712–4717.

997 Strock CF, Burridge JD, Niemiec MD, Brown KM, Lynch JP. 2021. Root metaxylem
998 and architecture phenotypes integrate to regulate water use under drought stress. *Plant*
999 *Cell Environ.* 44:49–67.

1000 Strock CF, Schneider HM, Galindo-Castañeda T, Hall BT, Van Gansbeke B, Mather
1001 DE, Roth MG, Chilvers MI, Guo X, Brown K, et al. 2019. Laser ablation tomography
1002 for visualization of root colonization by edaphic organisms. *J. Exp. Bot.* 70:5327–5342.

1003 Strock CF, Schneider HM, Lynch JP. 2022. Anatomics: High-throughput phenotyping of
1004 plant anatomy. *Trends Plant Sci.* 27:520–523.

1005 Wahl S, Ryser P. 2000. Root tissue structure is linked to ecological strategies of grasses.
1006 *New Phytologist* [Internet] 148:459–471. Available from:
1007 http://dx.doi.org/10.1046/j.1469-8137.2000.00775.x

1008 Walley JW, Sartor RC, Shen Z, Schmitz RJ, Wu KJ, Urich MA, Nery JR, Smith LG,
1009 Schnable JC, Ecker JR, et al. 2016. Integration of omic networks in a developmental
1010 atlas of maize. *Science* 353:814–818.

1011 Wang W, Guo W, Le L, Yu J, Wu Y, Li D, Wang Y, Wang H, Lu X, Qiao H, et al.
1012 2022. Integration of high-throughput phenotyping, GWAS, and predictive models
1013 reveals the genetic architecture of plant height in maize. *Mol. Plant* [Internet]. Available
1014 from: http://dx.doi.org/10.1016/j.molp.2022.11.016

1015 Wellhausen EJ, Roberts LM, Hernandez-Xoconostle E, in collaboration with
1016 Mangelsdorf PC. 1952. Races of Maize in Mexico: their origin, characteristics and
1017 distribution. Jamaica Plain: Bussey Institution of Harvard University

1018 Yamauchi T, Tanaka A, Mori H, Takamure I, Kato K, Nakazono M. 2016. Ethylene-
1019 dependent aerenchyma formation in adventitious roots is regulated differently in rice
1020 and maize. *Plant, Cell & Environment* [Internet] 39:2145–2157. Available from:

1021   http://dx.doi.org/10.1111/pce.12766

1022   Zhu J, Brown KM, Lynch JP. 2010. Root cortical aerenchyma improves the drought
1023   tolerance of maize (Zea mays L.). *Plant Cell Environ.* 33:740–749.

1024   Zomer RJ, Trabucco A, Bossio DA, Verchot LV. 2008. Climate change mitigation: A
1025   spatial analysis of global land suitability for clean development mechanism afforestation
1026   and reforestation. *Agric. Ecosyst. Environ.* 126:67–80.

1027

1028

1029 **Table 1** Description of root anatomical traits used in this study.

Abbreviation	Trait description
RXSA	Root Cross-Section Area (mm^2)
TCA	Total Cortical Area (mm^2)
TSA	Total Stele Area (mm^2)
TSA.RXSA	Total Stele Area:Root Cross-Section Area
TSA.TCA	Total Stele Area:Total Cortical Area
AA	Aerenchyma Area (mm^2)
X.A	Percent of Cortex as Aerenchyma
CCA	Cortical Cell Area (mm^2)
CCS	Cortical Cell Size (mm^2)
X.CCA	Percent Living Cortical Area
MA	Individual Metaxylem Vessel Area (mm^2)
MD	Individual Metaxylem Vessel Diameter (mm)
MVA	Total Metaxylem Vessel Area (mm^2)
NMV	Number of Metaxylem Vessels
CCN	Number of Cortical Cells
CCFN	Cortical Cell File Number

1030

1031

1032
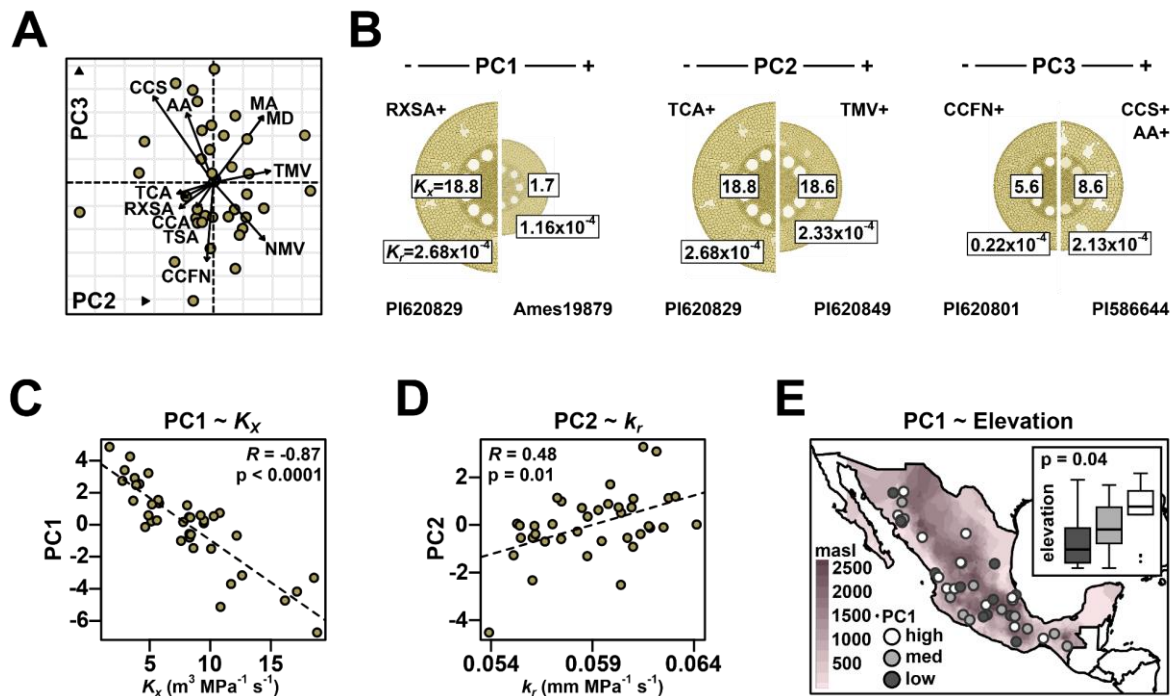
Figure 1. Root anatomy varies in Mexican native maize. A) PC2 and PC3 loadings for root anatomical traits of accessions from Burton et al. 2013. Trait description and codes as Table 1. B) Representative cross sections of extreme high- and low- loading individuals for PCs 1-3 rendered using GRANAR, scaled to the measured *root cross-section area*. Trait codes indicate broad trends seen in trait loading on the PCs. Boxed numbers adjacent to the central stele show modeled axial conductance (K_x). Boxed numbers on the outer epidermis show modeled radial conductance (K_r). Accession numbers are given at the base of the images. C) Correlation between modeled K_x and anatomical PC1. D) Correlation between modeled radial conductivity (k_r) and anatomical PC2. E) Accession source labeled by loadings on PC1, divided into terciles as low, medium (med) or high. Base map shaded by elevation. Inset box plots show the median and quartile elevation for the low, med and high PC loading groups. Whiskers extend to the most extreme points within 1.5x box length; outlying values beyond this range are shown as points. Stated p-value refers to an ANOVA for differences in elevation among the PC1 tercile groups.
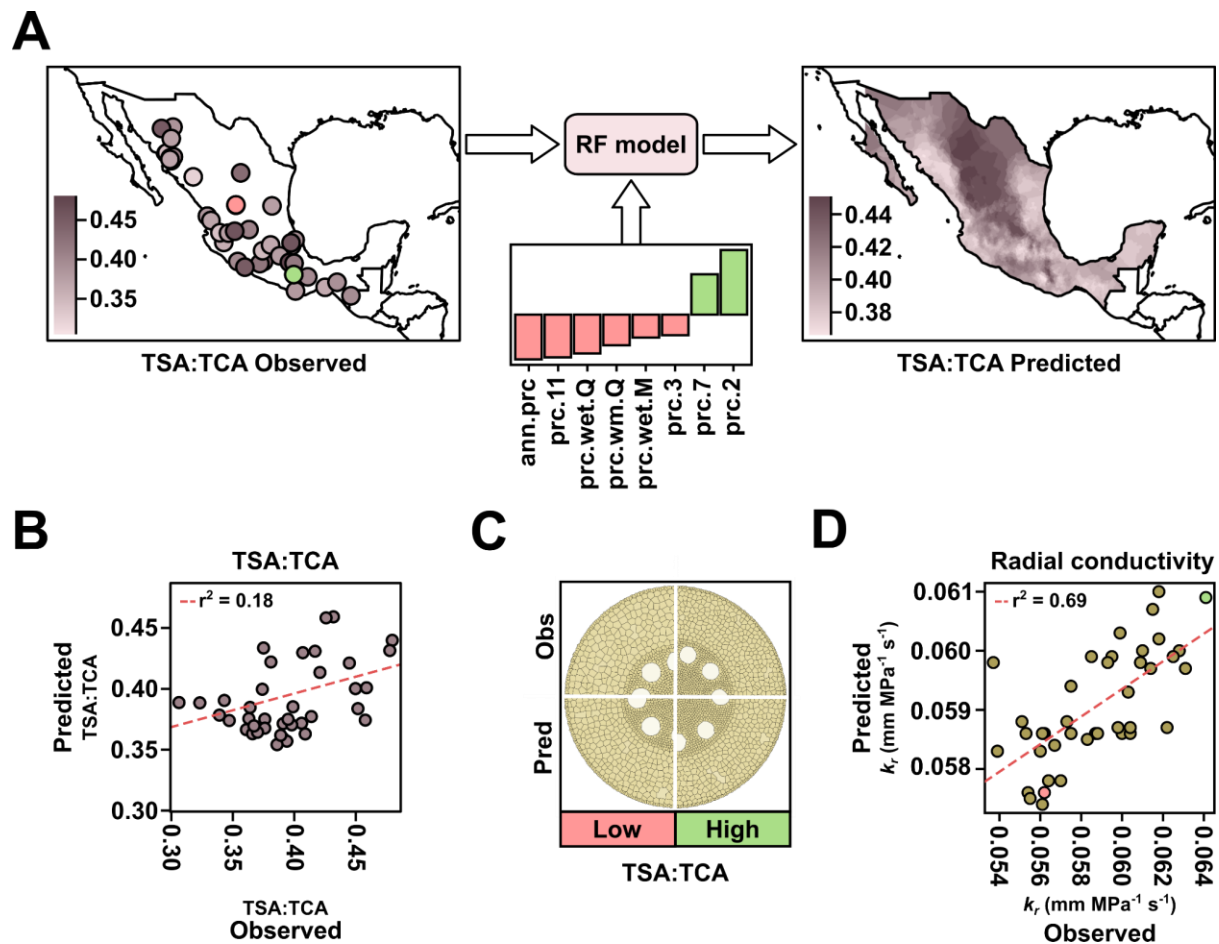
1048

Figure 2. Home environment predicts root anatomy in Mexican native maize. A)

Random forest (RF) modeling for the ratio of *total stele area:total cortical area* (TSA:TCA).

Accession point of origin colored by observed TSA:TCA from Burton 2013. Individuals

colored pink and green denote the accessions with the lowest (AMES19907) and highest

(PI629263) observed TSA:TCA, respectively. Trait-specific significant environmental

descriptors identified by the Boruta method used for RF model construction, displayed as

SHapley Additive exPlanations (SHAP) contributions. Smoothed RF predicted TSA:TCA for

native Mexican maize. B) RF predicted vs observed TSA:TCA values for all individuals used

in model training and validation. C) Composite GRANAR representation of observed (obs)

and predicted (pred) GRANAR sections for the accessions with the lowest (low, pink) and

highest (high, green) observed TSA:TCA. Predicted GRANAR cross-sections use predictions

for all traits for which RF models were constructed and are rendered at the same size. D)

1061   Predicted vs observed radial conductivity (k_r) for all individuals used in model training and

1062   validation. Predicted k_r was calculated using RF anatomical predictions and observed k_r was

1063   calculated using observed anatomical values from Burton et al. 2013. The individuals with

1064   the lowest and highest observed TSA:TCA are colored pink and green, respectively. Dashed

1065   line is the coefficient of determination for all plotted points.

1066
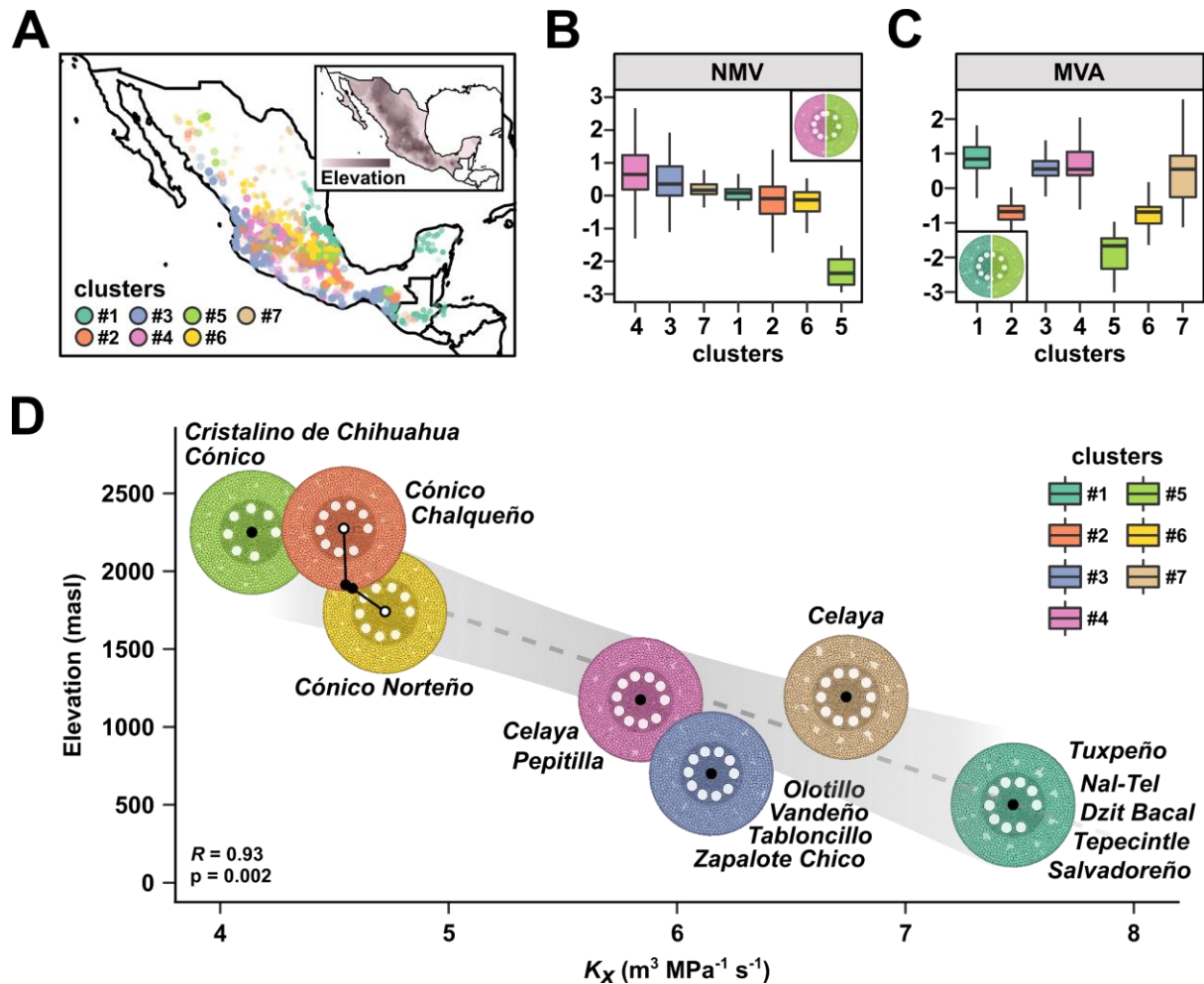
1067

Figure 3. Groups defined by shared root anatomical characteristics originate from

distinct environments. 1791 maize accessions forming the CIMMyT panel were grouped

into seven clusters based on eight RF predicted root anatomical traits. A) Geographical

distribution of the clusters. Inset shows elevation, with darker shading corresponding to

higher values. B) Centered and scaled *Number of metaxylem vessels* (NMV) and C) *Total*

metaxylem vessel area (MVA) in the seven clusters. Inset shows composite GRANAR

representation generated from the median trait values of highest (left) and lowest (right)

scoring clusters. D) Variation in mean cluster predicted axial conductance across elevation.

Black points indicate the mean axial conductance calculated using RF predicted anatomy vs

the mean elevation at the point of collection for each of the seven clusters. For clusters 2 and

1078   6, root-cross section renderings are plotted off the regression line for easier view of digitized

1079   anatomy. Root cross-section images are GRANAR representations generated using the

1080   median trait values for each cluster, all rendered at the same size. Native varieties

1081   overrepresented in each cluster are listed adjacent to the GRANAR images.

1082
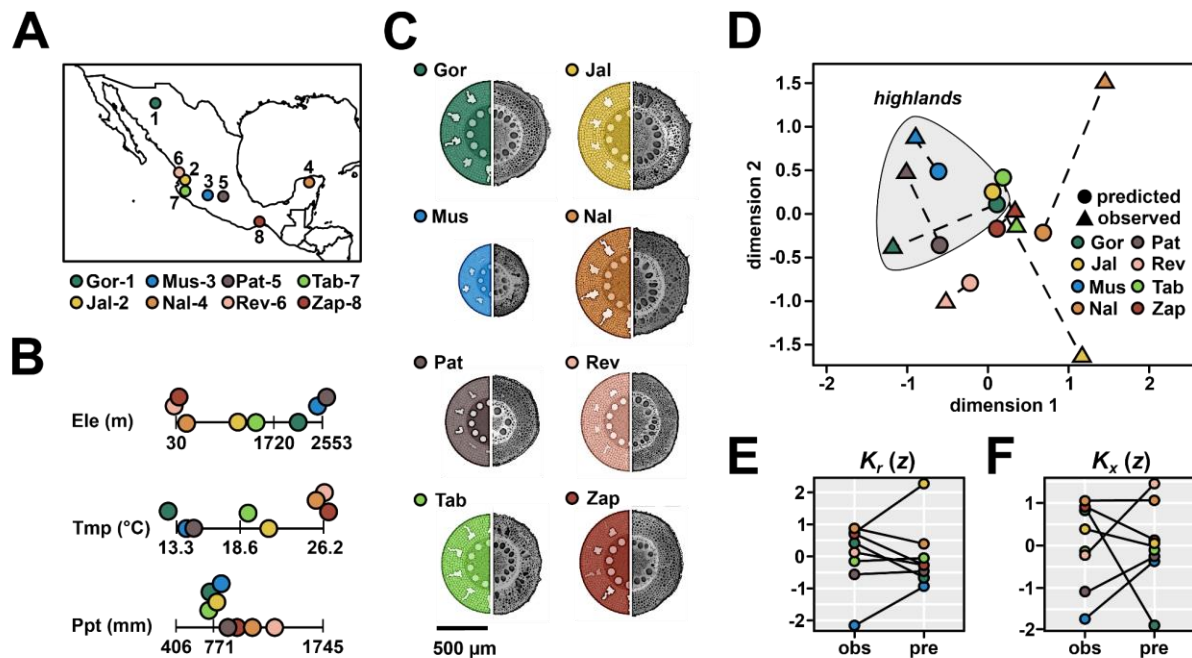
1083

1084 **Figure 4**. **Novel phenotypic data is consistent with model predictions.** A) Source of eight

1085 selected native maize varieties: 1) Gordo, 2) Jala, 3) Mushito, 4) Nal Tel, 5) Palomero

1086 Toluqueño, 6) Reventador, 7) Tabloncillo, 8) Zapalote Chico. B) Elevation (Ele), annual

1087 mean temperature (Tmp), and annual precipitation (Ppt) at source locality for the eight native

1088 varieties. Bars on the line plots represent the 5%, 50%, and 95% quantiles for each

1089 environmental descriptor across all CIMMyT panel individuals included in this study (1791).

1090 Points color-coded as A. C) GRANAR renderings of observed anatomical BLUPs across

1091 node two and three roots (colored) and photographs of representative cross-sections of third-

1092 node roots, scaled to the mean measured *root cross-section area.* D) Procrustes analysis

1093 comparing distribution of RF-predicted (circles) and observed (triangles) anatomical traits.

1094 For each variety, predicted and observed projections are linked with a dotted line. An

1095 arbitrary ellipse was added around three varieties sourced from high elevation. E)

1096 Comparison of standardized observed (obs) and predicted (pre) values of modeled radial

1097 conductance (K_r). Values for each variety are connected to illustrate the level of consistency

1098 in ranking. F) as E, showing modeled axial conductance (K_x).
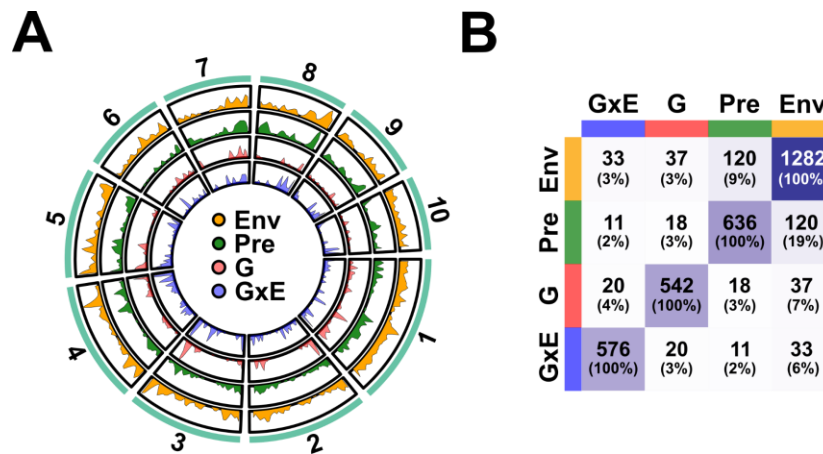
1099

1100

1101

1102

Figure 5. Evidence for a shared genetic basis of root anatomical variation between inbred breeding lines and Mexican native maize. A) Distribution of top 100 genes from GWA analyses of WIDP and CIMMyT accessions across the genome. The shaded area represents the density of top genes overlapped with window regions (1×10^7 bp). Sector names represent the number of chromosomes. B) The number of pairwise overlapping genes among the GWA gene sets. The darker the color of the squares indicates a higher number of genes. Totals in parentheses show the percentage of the row set in the other sets.
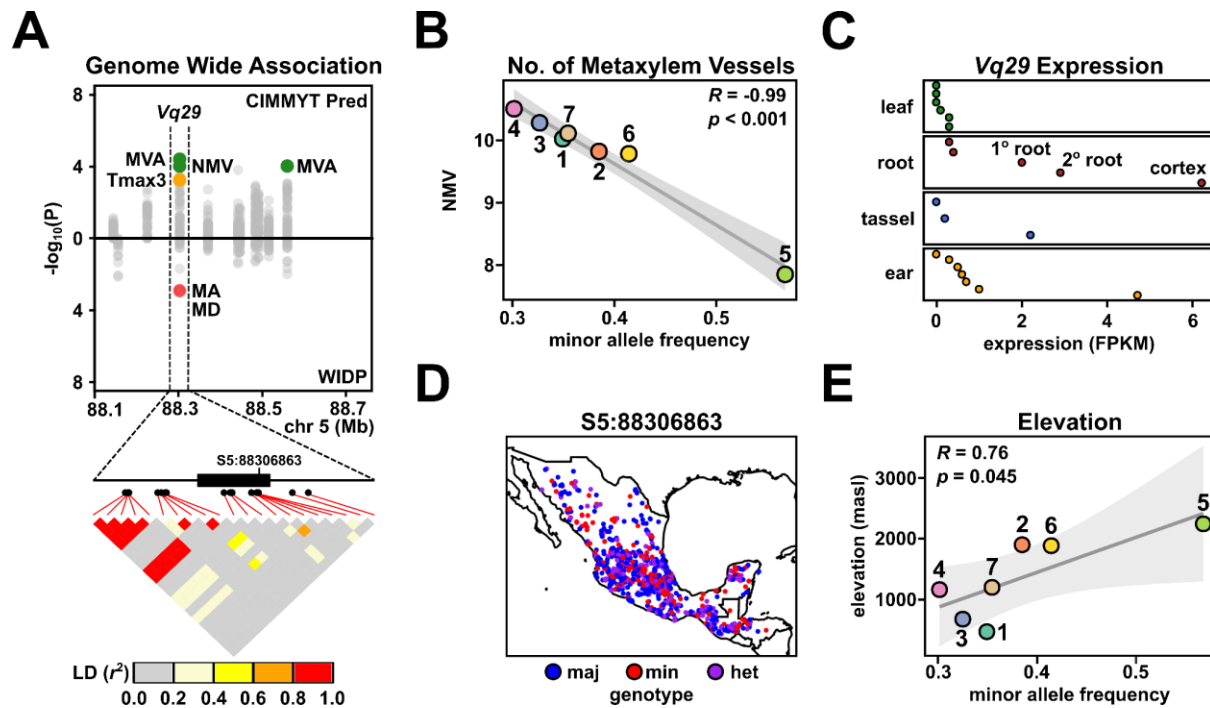
Figure 6. The gene *Vq29* is linked to variation in root anatomy and source environment.

A) Miami plot showing GWA support ($-\log_{10}P$) for association with root anatomy for genes on a region of chromosome 5. Points above the x-axis show support for predicted phenotypes in the CIMMyT panel; points below the x-axis show support for observed phenotypes for the WIDP panel. The gene *Vq29* is associated with *total metaxylem vessel area* (MVA), *number of metaxylem vessels* (NMV), *individual metaxylem vessel area* (MA) and *individual metaxylem vessel diameter* (MD) across the two analyses. Image below the Miami plot shows the *Vq29* gene model (CDS as filled box), SNP position (filled circles) and pairwise linkage disequilibrium (LD). The position of the focal SNP S5:88306863 is highlighted. B) Correlation between frequency of the minor allele at S5:88306863 in the previously defined CIMMyT clusters and mean predicted NMV. C) Expression of *Vq29* in four named tissues from publicly available expression data. Points show different subsamples. The root cortex, corresponding to the highest expression, is highlighted. D) Geographic allele-distribution of S5:88306863 in the CIMMyT panel. E) Correlation between frequency of the minor allele at S5:88306863 and mean elevation at source in the CIMMyT clusters.