Genomics of Root System Architecture Adaptation in Sorghum under Nitrogen and Phosphorus Deficiency

Joshua Magu[1], Joel Masanga[2], Catherin Muui[3], Steven Runo[1*]

[1]Department of Biochemistry, Microbiology and Biotechnology, Kenyatta University, P O Box 43844-00100, Nairobi, Kenya

[2]Department of Biology, The Pennsylvania State University, 432 Science Drive, University Park, PA 16802, USA

[3]Department of Agricultural Science and Technology, Kenyatta University, P O Box 43844-00100, Nairobi, Kenya

Corresponding author email: runo.steve@ku.ac.ke

Abstract

Plant root systems play a crucial role in water and nutrient uptake, yet the genetic mechanisms underlying root architecture adaptation to environmental stress remain poorly understood, particularly in non-model crops. Here, we investigate the diversity of root system architecture (RSA) in *Sorghum bicolor*, a climate-resilient cereal, under nitrogen (N) and phosphorus (P) starvation. We used a diverse sorghum reference set, and employed crown root imaging technology to quantify RSA traits and performed a genome-wide association study (GWAS) to identify genetic factors driving these responses. Our analysis revealed significant changes in six RSA traits—average root diameter, surface area, volume, total root length, and root length diameter classes (RLDR1 and RLDR2)—under N and P starvation. GWAS identified SNPs associated with these traits, pinpointing three candidate genes—*ILR3-like*, *bHLH*, and a *LEUNIG* homolog—with known roles in root growth regulation. These findings provide novel

genetic insights into sorghum root adaptation to nutrient limitations and offer potential targets for breeding climate-resilient, resource-efficient crop varieties.

Summary

Plants absorb water and nutrients from the soil through their roots, yet for most crops, little is known about how root shape and structure adapt to stressful conditions such as poor soil fertility. In this study, we used a globally diverse collection of sorghum genotypes to investigate how sorghum roots respond to nitrogen and phosphorus deficiency. Remarkably, some genotypes showed strong root adaptations under these conditions, leading us to identify the genetic factors driving these responses. Our findings not only improve our understanding of root adaptation to nutrient stress but also highlight promising genetic targets for breeding more nutrient-efficient crops.

Introduction

Root system architecture (RSA)—the spatial organization of root traits, or *phenes*—determines plant performance under both optimal and limiting conditions (Freschet et al. 2021). RSA also mediates interactions between roots and the rhizosphere, enhancing root foraging and physiological responses (York et al. 2013). Thus, understanding root *phene* development is crucial for advancing knowledge in rhizosphere communication, resource foraging, and plant physiological adaptation, with far-reaching implications for practical agriculture.

This understanding is particularly valuable when applied to crops such as sorghum (*Sorghum bicolor*), a globally important cereal for food, feed, and bioenergy. As a C_4 plant, sorghum can survive under adverse conditions such as high temperature and drought stress (Trouche et al. 2014); making it a strategic crop for mitigating adverse effects of climate change.

In addition to its tolerance to abiotic stresses, sorghum exhibits remarkable adaptability in RSA in response to nitrogen (N) and phosphorus (P) deficiencies; an important trait that

enhances nutrient foraging in low-input environments. Under N scarcity, sorghum increases root length to access mobile nitrate in deeper soil layers (Lynch 2013), and may also produce more lateral roots to enhance soil exploration (Saengwilai et al. 2014). In contrast, P deficiency typically induces a shallower root angle—likely due to P's immobility and accumulation in the topsoil—alongside increased root hair length and density to improve phosphorus uptake (Lynch 2011).

We sought to understand the distinct changes in sorghum RSA under N and P scarcity, as well as to identify the genetic regions underlying these traits for potential integration into breeding programs targeting nutrient-limited environments. To achieve this, we used the sorghum reference set, which includes all major cultivated, wild, and intermediate sorghum races originating from diverse agroecological regions across Africa, India, the Middle East, Europe, and North America (Casa et al. 2008) and has been genotyped sequencing (Morris et al. 2013). For phenotyping, we used a modified root crown phenotyping approach, which involved excavating the upper portion of the root system, removing the growth substrate (vermiculite in our case), and taking various measurements. The integrated RhizoVision Crown platform (Seethepalli et al. 2020) provided a reliable way to extract data from numerous roots. Similar systems have been successfully applied in studies of nutrient acquisition in crops like soybean (Glycine max L.), common bean (*Phaseolus vulgaris* L.), cowpea (*Vigna unguiculata* L.), wheat (*Triticum aestivum* L.), and maize (York 2018) .

We report on the results of changes in sorghum RSA under N and P scarcity, followed by association mapping of the traits to identify underlying genetic factors. We describe these findings and the implications thereof for breeding sorghum varieties better adapted to low-nutrient environments.

Results

Description of sorghum root traits

To extract root data from various sorghum genotypes, we used the RhizoVision platform, which provided 13 root traits. These traits are illustrated in Figure 1 and described in more detail in Table S1, with further illustrations in Figure 2. The traits can be grouped as follows: (i) Size: Total root length (TRL), Perimeter (P), Surface Area (SA), Network Area (NA), and Volume (V); (ii) Extent: Branch Points (BP), Branching Frequency (BF), and Number of Root Tips (NRT); (iii) Distribution: Average diameter (AD), Median diameter (MD), and Maximum diameter (MD).

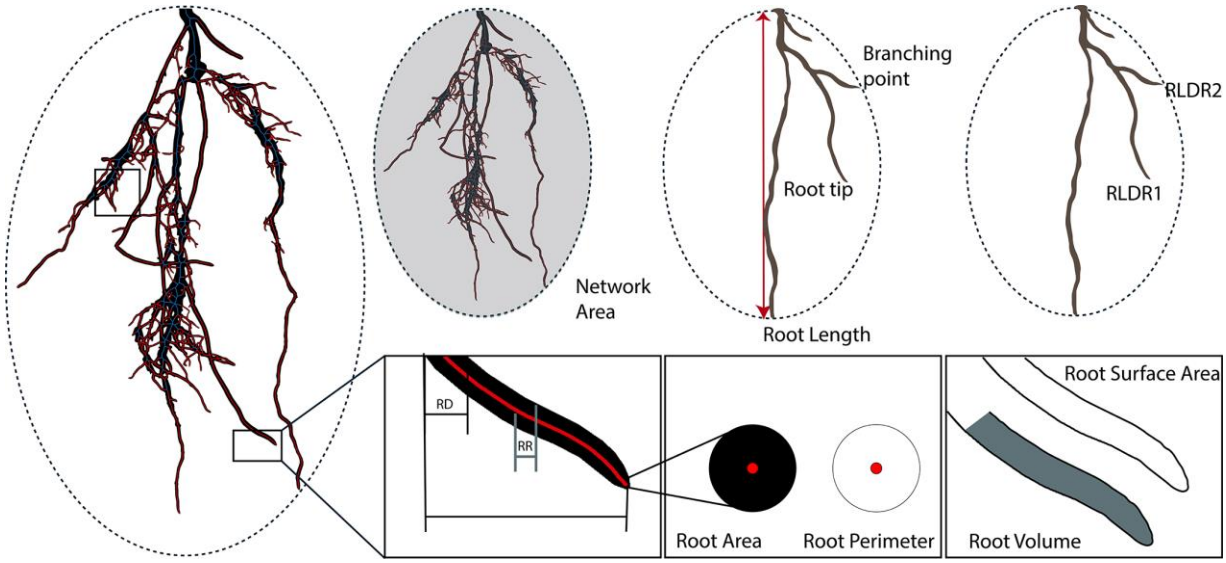

Figure 1. RhizoVision analysis of root system architecture showing a skeletonized sorghum root image used to extract measurements. The upper panel shows measurement for network area (left panel), root length, root tip and branching point (middle panel) and root length diameter range 1 (RLDR1) and RLDR2 (right panel). The lower panel shows a small root region that is amplified to show root diameter (RD), and root radius (RR) in the left panel, root area and root perimeter (middle panel) and root volume and root surface area (right panel). Scale of skeletonized root crown = 140 mm.

Genetic variation of sorghum root architecture and response to N and P nutrients starvation

To examine variation in the sorghum root system we used a subset of 123 genotypes of the sorghum diversity panel with a global representation (Figure S1a). The panel was highly diverse and comprised all major sorghum races. Phylogenetic analysis using Neighbor joining grouped the genotypes into 7 clades corresponding to races: bicolor, caudatum, durra, guinea, Guinea

subraces, kafir and wild genotypes (Figure S1b). These same groupings were reflected in Bayesian hierarchical clustering (Figure S1c) at the most optimal K (K=7). Additionally, the sorghum genotypes represented populations originating in diverse agro-ecological regions across Africa, Asia, America and Europe.

To determine if there were distinct differences in RSA that could be used to cluster all 123 genotypes used in the study, we computed Euclidean distances for genotypes across treatments. Under optimum nutrient availability, hierarchical clustering (HC) grouped the genotypes into 5 clusters that represented different root types (Figure 2). Cluster I represented in Figure 2 by IS19685 had the most root area while cluster V represented in the figure by IS23574 had the lowest root area. Under nutrient constraint (N and P starvation) HC grouped the genotypes into 2 clusters – representing genotypes with large (cluster I) or small (cluster II) root area (Figure S2). The 5 clusters under optimal conditions could point to high inter genotypic variation in root structure while the two structures are indicative of specific inter genotypic differences of roots in response to N and P starvation. Overall, nutrient availability significantly reshaped the relationships among these genotypes, causing them to cluster differently under stress conditions.

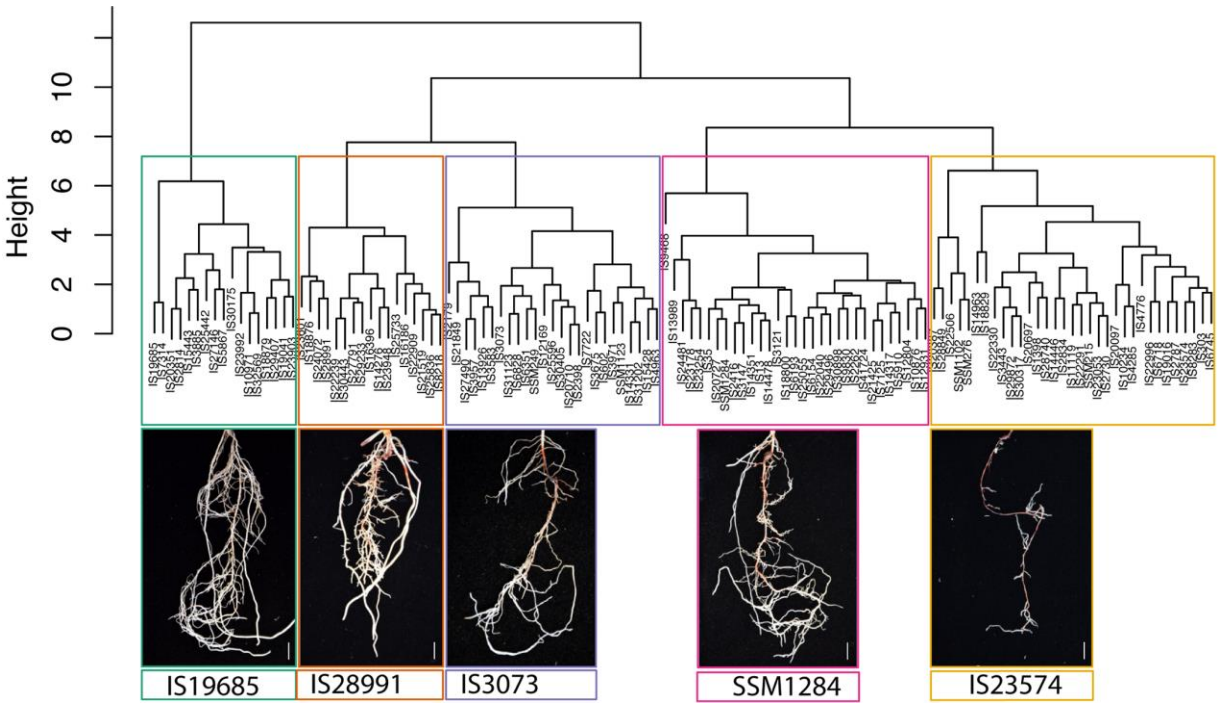

Figure 2. Variation of sorghum root architecture: Dendrogram showing hierarchical clustering of root traits from a dataset of 123 sorghum genotypes. Clustering was performed using the complete method and Euclidean distance. The colored boxes highlight five distinct root-type clusters, each represented with crown images corresponding to the genotypes within the cluster.

Partitioning of genotypic variation of sorghum root system architecture

Next, we examined how the variation of RSA was partitioned among the genotypes using principal component analysis (PCA) for each treatment condition: optimal, N and P starvation.

Under optimal conditions (Figure 3a), PCA1 and PCA2 accounted for 88.4% of the total variation (PCA1 = 56.9% and PCA2 = 31.5%). Variation in PCA1 was attributed to six traits: BF, RLDR1, NRTP, NBP, P, and TRL. Five of these traits clustered together, indicating a strong correlation between them. PCA2 comprised seven traits: RDL2, SA, NA, V, MaxD, AD, and MD. SA and NA overlapped, suggesting a high correlation between them. Distribution-related traits (AD, MaxD, and MD) formed a separate cluster with significant variance, indicating a lower correlation and weaker influence relative to other traits (Figure 3A).

In N starvation (Figure 3b), PCA1 accounted for 62.6% of the variation and PCA2 captured 28%, totaling 90.6% of the variance (Figure 3B). Under P starvation (Figure 3c), the total variance explained by PCA1 and PCA2 was 89.4% (Figure 3B). Both conditions exhibited greater variance compared to the optimal nutrient conditions (88.4%), suggesting distinct changes due to nitrogen and phosphorus removal. Notably, distribution traits (AD, MaxD, and MD) shifted to PCA3 under both N- and P-deficient treatments. This relocation indicates a significant change in their behavior and relationship with other traits, emphasizing a stronger negative or opposite correlation compared to their interactions under optimal conditions.

In all cases, dots representing different races did not form clusters but were widely dispersed within the PCAs, indicating high internal diversity. However, larger dots representing races with high influence on variance aligned with specific PCAs. Dura and wild genotypes, for example, showed strong associations with traits represented in PCA1 and PCA2, while kafir, guinea, and bicolor aligned were partitioned PCA3, which did not contribute to root variation. These findings suggest that dura, caudatum, and wild sorghum genotypes exhibit strong associations with root traits while bicolor, kafir, and guinea in PC3 and PC4 have less association with root trait variation. This implies that some genotypes exhibit greater resilience under nitrogen-limited conditions, making them potential candidates for breeding and further investigation.

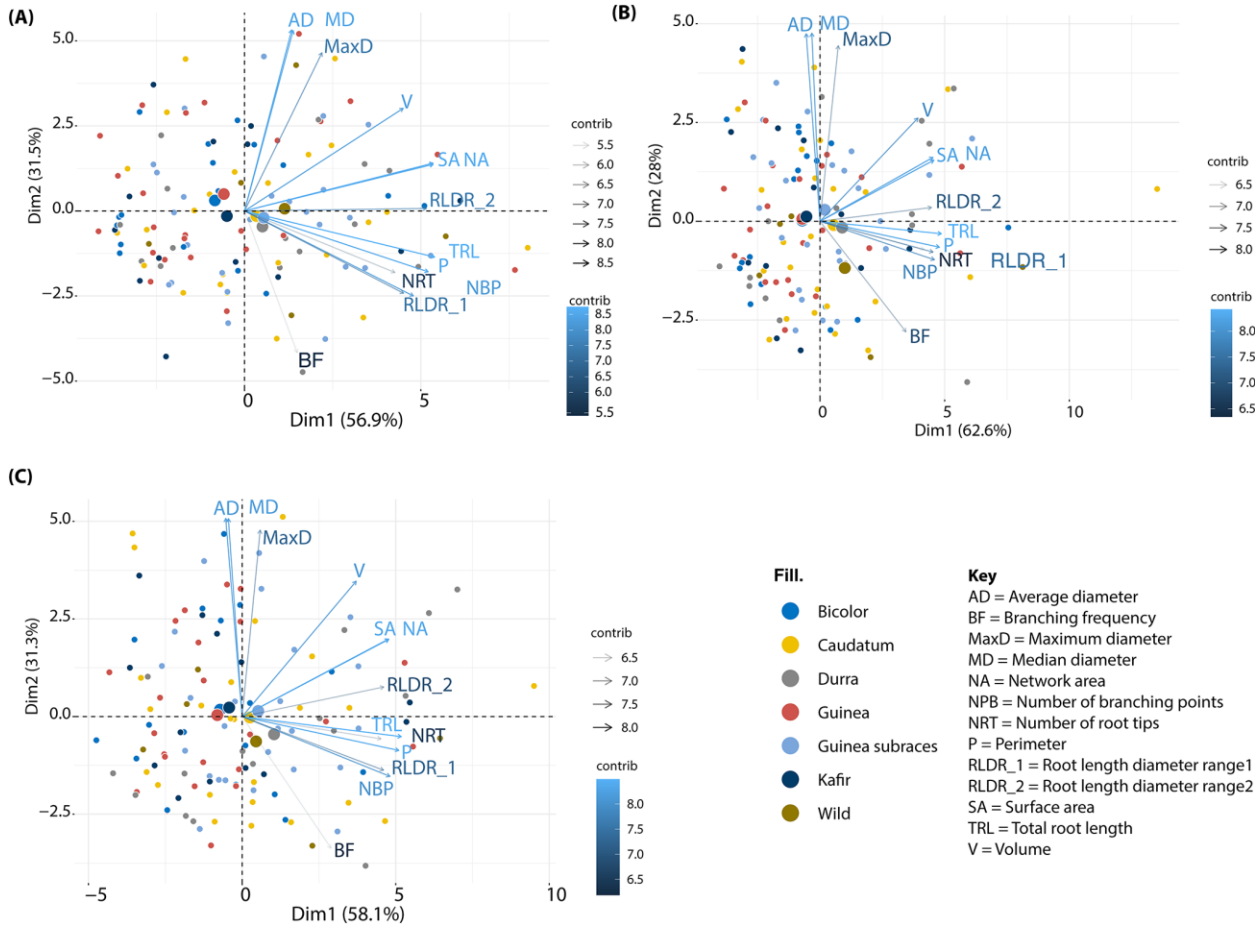

Figure 3. Partitioning genotypic variation in sorghum root system architecture (RSA) using biplot analysis of 123 genotypes under different treatments: (A) Under optimal conditions, variation was mainly explained by two principal components (PC1 = 23.02% (56.9), PC2 = 19% (31.5)). Distribution traits were weakly related to BF and other traits. (B) Under nitrogen-deficient conditions, variation shifted, with PCs showing a different contribution across experimental runs. (C) Under phosphorus-deficient conditions, distribution traits shifted similarly to the third quadrant. Colored dots represent sorghum races of the genotypes analyzed. Traits measured are shown on the key.

Interactions of sorghum root system architecture traits

To determine levels of correlations among various RSA traits under all experimental conditions, we performed a Person's correlation analasis. This analysis revealed a strong positive correlation between primary root length (P) and TRL (r = 0.99), as well as between RLDR1 and the number BPs (r = 0.91). Additionally, AD and MD were highly correlated (r = 0.99). Size-

related traits; MaxD, AD, and MD exhibited negative correlations with most other traits, particularly BF (Figure 4a).

Under N starvation conditions (Figure 4b), we observed a strong positive correlation between SA and NRT (r = 0.99). Positive correlations (Figure 4c) were also noted between P and TRL, AD, and MD, as well as between traits describing extent NBP and size (P and TRL). Conversely, a negative correlation was found between traits describing extent (BP) and distribution (MD, AD, MaxD). These trends were consistent under phosphorus-depleted conditions. Specifically, size traits (SA and NA), SA and V, SA and mass (MA), and AD were positively correlated, while negative correlations were observe d with distribution traits such as AD, MaxD, BF, NBP, and number of root tips (NRT). These findings highlight the inverse relationship between size and distribution in RSA traits.

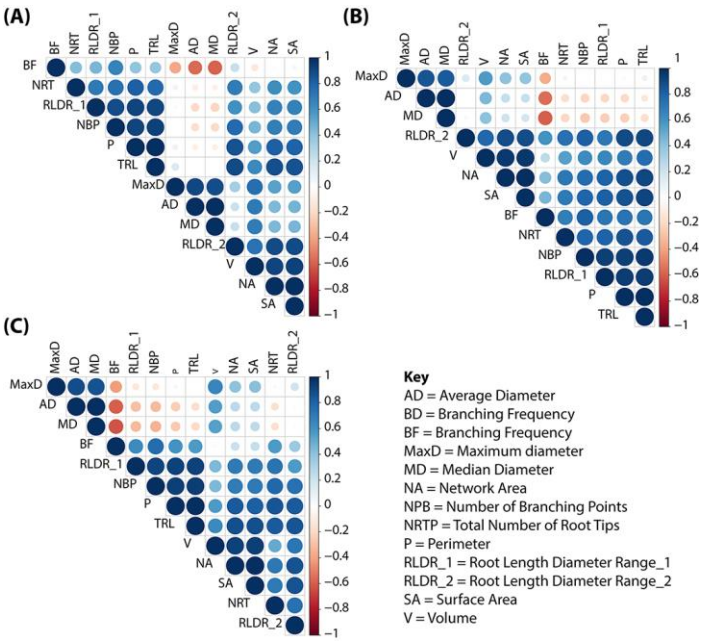

Figure 4: Pearson's correlation analysis between root system architecture (RSA) traits under: (A) Optimal (B) Minus nitrogen and (C) Minus phosphorus. Dark blue to dark brown indicates highly positive to highly negative correlations. Traits measured are shown on the key.

To further determine the correlation of RSA traits, we used the Gaussian graphical model described in (Bhushan et al. 2019) . In this model, a set of items or variables, are depicted by circles, and a set of lines that visualize relationships between the items or variables; where line

thickness represents the strength of the relationships between items while the absence of a line implies no or very weak relationships. The Gaussian graphical model also allows for examination of the conditional dependencies between studied traits. Greater values for outdegree and betweenness indicate greater centrality in a network, suggesting that a trait has influence on other traits. Results showed that the network area had the highest betweenness connecting other traits under all conditions – optimal, N and P starvation (Figure 5 a-c). Results also showed that BF weakly connected with NA, MD, P and RLDR1. Expectedly, traits linked to distribution (MaxD, AD, MD) clustered together and had weaker connection to traits associated with size (TRL, P, NA) and extent (NBP, NRT, BF) this was also observed in roots under N starvation (figure B). Together, these results suggest that NA has the highest conditional dependency, signifying its great influence on other traits.

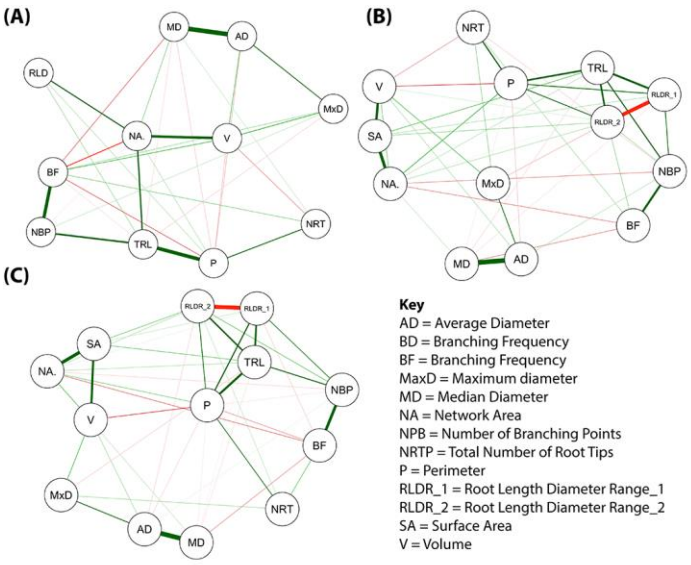

Figure 5. Gaussian graphical model showing strength of correlation between traits: A. Optimal nutrients, B, nitrogen starvation and C. Phosphorus starvation. Red and green lines show negative and positive correlations, respectively.

Impact of nutrient stress on sorghum root architecture

To assess the response of RSA traits under N and P starvation, we conducted a Multivariate Analysis of Variance (MANOVA). The analysis revealed significant differences in six RSA traits when conditions shifted from optimal to N and P starvation.

Specifically, root length traits—RLDR1, RLDR2, and TRL—increased markedly under P starvation, while remaining unchanged under N starvation. Conversely the RSA traits indicative of secondary growth—AD, SA and V—decreased under both N and P starvation, with a more pronounced reduction observed under P deficiency (Figure 6).

These findings suggest that root elongation traits are enhanced under phosphorus constraint, whereas root secondary growth diminish under both N and P limitations.

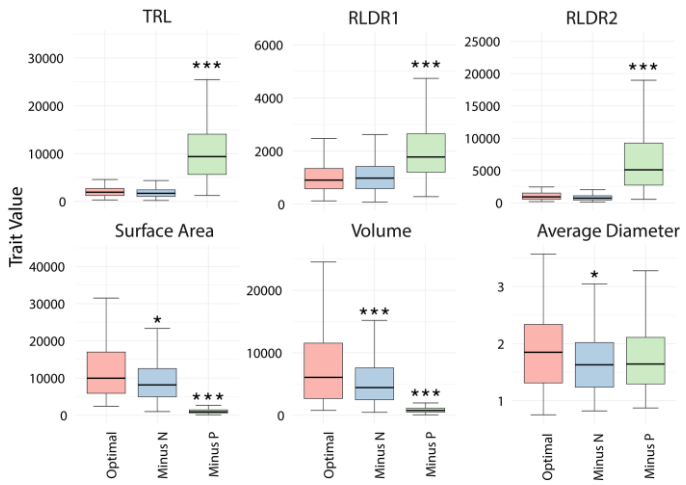

Figure 6: Impact of nutrient stress on sorghum root architecture
The traits that had statistical differences between treatments were total root length (TRL), root length diameter range 1 (RLDR1), and root length diameter range 2 (RLDR2), Surface area, volume and average diameter. Box plots show the distribution of each trait across the three nutrient conditions. Significant differences between treatments are indicated by asterisks (*$p < 0.05$, and ***$p < 0.001$).

Genome wide association studies to identify genetic factors underpinning root system architecture traits in sorghum

To investigate potential genetic factors underlying the observed RSA across different conditions, we performed a GWAS analysis. We analyzed all 13 traits individually using Fixed and Random Model Circulating Probability Unification (FarmCPU) and Multi-Locus Mixed Model (MLMM) models using Genome Association and Prediction Integration Tool (GAPIT). This analysis identified a total of 25 significant SNPs. Interestingly, significant SNPs were only detected under N and P starvation using RSA traits of V, SA, RLDR2, and NBP (N starvation) and RLDR2 (P starvation).

The most pominent SNP was S1_26205071 identified 11 times using V, SA, NBP, and RLDR2 under N starvation and RLDL2 under P starvation. This SNP did not fall on a gene region, but its highly significant P value and association with multiple traits in two models pointed to its importance in shaping the root architecture. Therefore, we characterized the phenotypic variation related to each haplotype of the SNP change. We determined that 92 and 8 accessions harbored the C and T haplotypes, respectively, whereas the remaining 37 accessions lacked genotype information. Under all environments, the accessions carrying the T allele showed significantly higher values of the RSA traits SA, RLDR2, and NRT than those carrying the C allele ($p \leq 0.05$), indicating that the T haplotype had a positive effect on RSA (Figure 5c).

Our analysis further identified several SNPs located on annotated genes with functions and phenotypes directly linked to RSA. For instance, S7_2952840 (p = 6.96E-08) was mapped to a gene annotated as *ILR3-like*. *ILR3* (also known as *bHLH105*) encodes a basic helix-loop-helix (bHLH) transcription factor involved in iron and auxin signaling—two pathways known to influence RSA (Tissot et al. 2019). Similarly, S10_11497177 (p = 3.08E-08) was located in the upstream region of a gene annotated as *bHLH18*. *bHLH18* is associated with stress response and hormone signaling pathways, including ethylene and gibberellins. These pathways regulate cell differentiation and patterning, affecting traits such as root branching and primary root length (Qin et al. 2022). Additionally, SNP 54899232 (p = 8.77E-09) was positioned in the exon of a gene annotated as a transcription corepressor of *LEUNIG*. Prior studies have shown that the *LEUNIG_HOMOLOG* (*LUH*) plays a critical role in root elongation. For example, *luh-1* mutants in Arabidopsis exhibit significantly shorter roots compared to wild-type plants (Sitaraman et al. 2008). The discovery of these root-associated genes highlights the robustness of GWAS analysis in uncovering the genetic mechanisms driving RSA trait adaptation under N and P starvation.

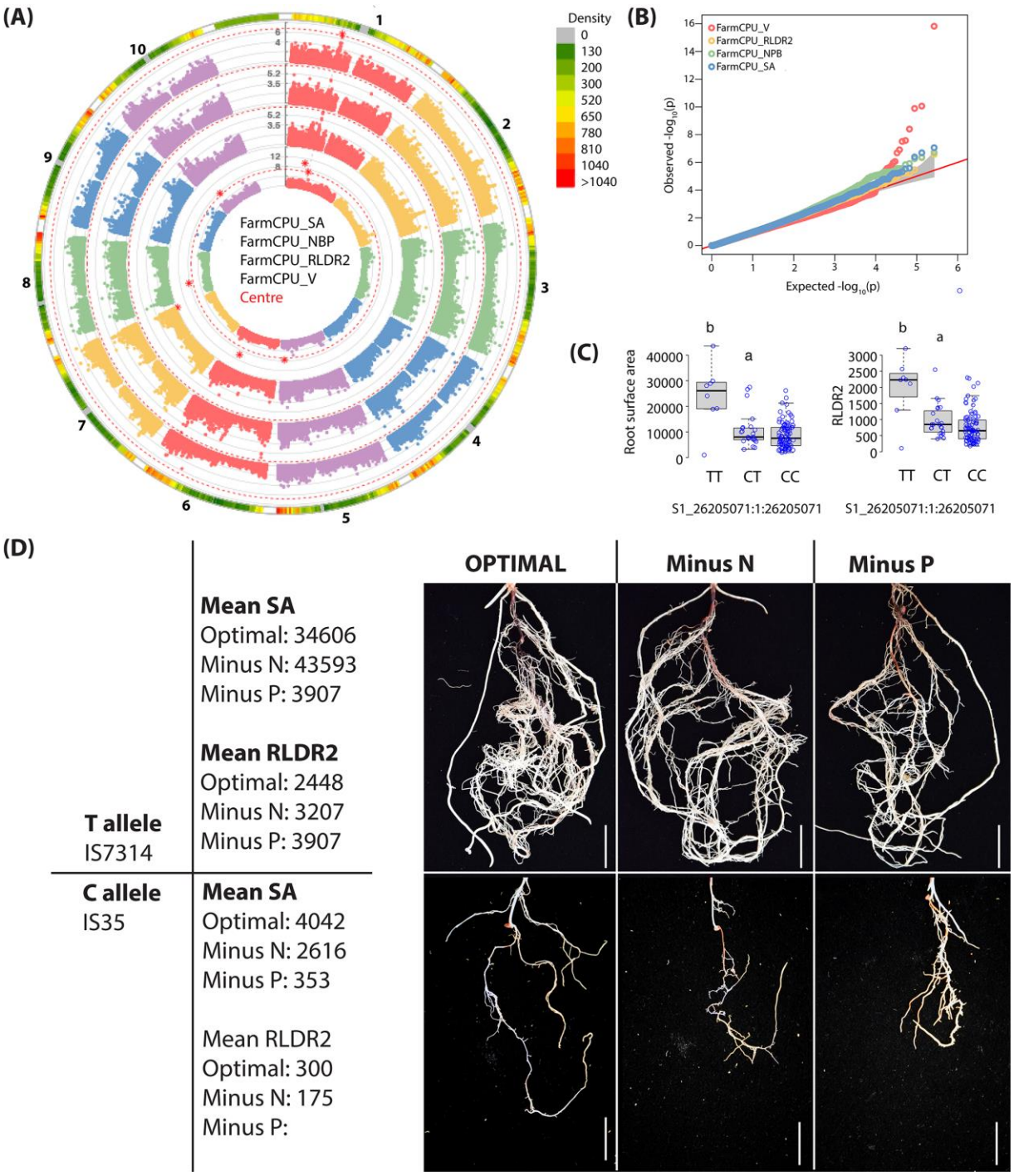

Figure 7: Genome-wide association studies (GWAS) identifying genetic factors influencing root system architecture (RSA) traits in sorghum. (A) Fixed and Random Model Circulating Probability Unification (FarmCPU) model circular Manhattan plot highlighting significant single nucleotide polymorphisms (SNPs) associated with root system architechure (RSA) traits, including volume (V), root length diameter 2 (RLD2), number of branch points (NBP), and surface area (SA). (B) Corresponding Quantile-quantile (QQ) plot comparing observed and expected p-values to assess model accuracy and false positives. (C) Boxplot illustrating the allelic effect of the most significant SNP (S1_26205071) on SA and

RLD2. Favorable alleles increase trait expression, while unfavorable alleles decrease it. (D) Representative root phenotypes comparing "C" and "T" allele genotypes under optimal, nitrogen-deficient, and phosphorus-deficient conditions. Genotypes with the "C" allele exhibited significantly higher SA and RLD2. Scale bar = 140 mm.

Table 1. Sorghum single nucleotide polymorphisms (SNPs) showing significant genome-wide associations with root system architecture (RSA) traits (FDR correction at ∝0.05). SNPs within genes are presented with their annotations in order of chromosomal location. AF: allele frequency; AdjP: P-value after FDR adjustment. *represent SNPs identified under nitrogen starvation, †represent SNPs identified under phosphorus starvation. Letters after SNP positions represent GWAS model used: a = BLINK, b = FarmCPU, c = MLMM.

Chr	Position	AF	AdjP	Associated Trait	Gene ID	Annotation
1	26205071*c	0.16	6.18E-08	Surface area	N/A	N/A
1	26205071*c	0.16	1.67E-07	Volume	N/A	N/A
1	26205071*b	0.16	4.05E-08	Surface area	N/A	N/A
1	26205071†c	0.16	1.89E-13	RLDL2	N/A	N/A
1	26205071*c	0.16	5.43E-08	Surface area	N/A	N/A
1	26205071*c	0.16	5.43E-08	Surface area	N/A	N/A
1	26205071*c	0.16	1.42E-09	RLDL2	N/A	N/A
1	26423820*c	0.20	1.05E-07	Volume	N/A	N/A
1	26731398*c	0.13	1.78E-07	Surface area	N/A	N/A
1	26731398*c	0.13	1.72E-08	Volume	N/A	N/A
1	26731398*b	0.13	1.19E-07	Surface area	N/A	N/A
1	26731398*b	0.13	3.70E-11	Volume	N/A	N/A
1	26731398*c	0.13	3.29E-08	Volume	N/A	N/A
1	27129528*c	0.10	2.20E-08	Volume	XM_002464586.2	CCRP 4 homolog
1	26205071*a	0.16	1.10E-10	Surface area	N/A	N/A
1	26423820*a	0.20	4.68E-08	Volume	N/A	N/A
1	26205071*a	0.16	1.56E-12	RLDL2	N/A	N/A
1	20102651*b	0.40	8.66E-11	Volume	N/A	N/A
2	63695738†c	0.42	2.03E-08	RLDL2	N/A	N/A
2	63695779†b	0.41	5.28E-09	RLDL2	N/A	N/A
3	15123886*b	0.10	1.97E-08	RLDL2	XP_002455526	Uncharacterised protein
3	54899232†c	0.08	8.77E-09	RLDL2	XP_02131268.1	*LEUNIG*
3	67879725*a	0.12	6.48E-11	Volume	XP_002456657.1	Cell Division Control Protein 6
4	12760823†b	0.28	6.82E-11	RLDL2	N/A	N/A

	5702394[*b]	0.09	8.74E-08	NBP	XP_002451690.2	Uncharacterised protein
4	58775820[*b]	0.16	3.12E-09	NBP	XP_021314639.1	*MAPK14-like*
5	17754110*b	0.04	3.38E-08	NBP	XP_002449381.2	1-aminocyclopropane-1-carboxylate oxidase 1
5	59062428[*b]	0.12	2.63E-08	Volume	N/A	N/A
6	55153514[†b]	0.08	6.48E-08	RLDL2	XP_021318054.1	E3 ubiquitin-protein ligase RNF14
6	56342501[*c]	0.03	2.55E-08	N NBP	XP_021318339.1	probable LRR receptor-like serine/threonine-protein kinase/hypothetical protein
6	56929524[*b]	0.03	4.41E-09	Volume	SORBI_3006G223101	Hypothetical protein
6	4077580[*a]	0.05	1.16E-09	NBP	XP_021318130.1	Uncharacterised protein
6	45806791[*b]	0.27	1.29E-10	Volume	XP_002446521.1	Transcription factor bHLH18
7	2952840[*c]	0.07	6.96E-08	NBP	XP_0213120712.1	ILR3
7	64170934[*b]	0.43	3.75E-13	Volume	N/A	N/A
8	15675606[*b]	0.05	3.96E-09	Volume	N/A	N/A
9	49587358[†b]	0.35	1.35E-11	NBP	N/A	N/A
9	49587358[*a]	0.35	1.44E-09	NBP	N/A	N/A
10	9879684[*b]	0.16	5.00E-09	NBP	XP_002438177.2	peroxygenase 4
10	17059423[*a]	0.35	2.99E-09	Volume	N/A	N/A
10	11497177[*b]	0.46	3.08E-08	Volume	N/A	N/A

Discussion

Our study sought to address the following questions: (i) what are the inter genotypic RSA variation and associated adaptation to N and P starvation?; (ii) how do RSA adopt to low N and P starvation? and (ii) which genetic factors underpin RSA trait adaptability?. The following sections provide an in-depth discussion of these questions, exploring the implications of our findings and their relevance to sorghum improvement.

What are the inter genotypic RSA variation and associated adaptation to N and P starvation? Data from neighbor joining phylogenetic analysis grouped our sample size of 123

sorghum genotypes into all representative races of sorghum. The results were further confirmed with HBC which grouped the races into 7 clusters. Expectedly, genetic diversity of the panel was further enhanced by the broad range from where the sorghum genotypes were sourced – owing to agroecological adaptations.

We found that all RSA traits varied significantly among the genotypes. A hierarchical clustering (HC) dendrogram grouped the genotypes into five clusters, each representing distinct root types. However, when N and P were removed, only two clusters emerged. This reduction in the number of clusters suggests that the differences in root architecture are primarily driven by N and P deprivation. It also indicates that most genotypes exhibit similar responses to N and P starvation. Genotypes that maintain RSA traits under these conditions may possess specialized mechanisms for N and P uptake or utilization, making them unique even under low nitrogen availability. These genotypes are of particular interest for breeding programs focused on improving N and P use efficiency or enhancing tolerance to nutrient stress.

PCA analysis revealed that NA, SA and V accounted for most phenotypic variations, implying that these traits account for most of the diversity observed between the sorghum genotypes tested. This observation was also highlighted in network analysis which showed NA and BF to have the most significant effect. There are studies that both support and contrast these findings depending on the species, environment, and focus of the study. (Lynch 2011) emphasizes that root surface area and volume are critical for resource acquisition and account for significant variation in root functionality, particularly under stress conditions like low N or P availability. Similarly, (York et al. 2013) showed that traits like root network area and volume explain much of the variation in resource acquisition, especially under water-limited conditions. (Postma and Lynch 2011) showed that root length and branching density can account for more variation in nutrient acquisition than surface area or volume, depending on the soil nutrient distribution.

For both PCA, and correlation analysis, extent traits (MaxD, AD, and MD) on PC 2 showed a weak correlation with other traits – especially distribution traits. This negative correlation highlights trade-offs among RSA traits as suggested by York et al., (2013). The authors argue that thicker roots (higher root diameter) tend to be less branched or have lower branching frequency, particularly in resource-constrained environments. This observation aligns with the idea that plants optimize root traits for specific soil conditions—favoring fewer but thicker roots for penetrating compacted soils or thinner, more branched roots for nutrient exploration in less compacted soils.

How do RSA adopt to low N and P starvation? Results showed that both N and P starvation led to enhanced root length and diminished secondary growth. Enhanced root growth under P starvation is consistent with results obtained in rice (Wissuwa 2003; Shimizu et al. 2004; Yi et al. 2005) and in maize (Mollier and Pellerin 1999); but in contrast to shorter root growth observed in Arabidopsis grown in low P conditions. The varying effects of P deficiency on root development may be attributed to species and genotypic variations, as Reviewed in (Niu et al. 2013). For this study, the increased elongation of lateral results in a more branched root system with a greater efficiency for P uptake.

Root elongation under low P was also accompanied by reduced secondary growth. This is consistent with the notion that reduced root secondary growth decreases maintenance and construction costs, allowing greater root elongation and soil exploration, thereby improving P acquisition and plant growth under P stress. The afore hypothesis was advanced by (Strock et al. 2017) and supported using in silico and root measurement studies in beans.

Which genetic factors underpin RSA trait adaptability? GWAS results revealed a significant SNP that colocalized with SA and RLDR2. Further analysis of phenotypic variation linked to the different haplotypes of S1_26205071 (C allele vs. T allele) demonstrated the phenotypic impact of this SNP where C Allele accessions exhibited significantly higher values of SA and RLDR2. The haplotype analysis confirms that this SNP is not just statistically significant

but biologically relevant in shaping important root traits, making it a strong candidate for further validation and incorporation into breeding strategies aimed at improving root system architecture. Although no gene was associated with the SNP, this SNP could be linked to a causative variant in the vicinity, functioning through linkage disequilibrium. Additional fine-mapping or functional studies, such as transcriptome analysis, could reveal specific genes affected by this SNP.

Our GWAS analysis revealed significant associations with additional SNPs located in gene-coding regions, with three noteworthy genes identified: (i) The *ILR3-like* gene belongs to the basic helix-loop-helix (*bHLH*) family of transcription factors, which are known to play pivotal roles in iron homeostasis, nutrient acquisition, and root development. In *Arabidopsis thaliana*, the *ILR3* gene regulates iron deficiency responses, influencing root architecture by modulating root hair development and lateral root growth (Tissot et al. 2019). Given its involvement in iron regulation, *ILR3-like* may mediate crosstalk between nitrogen signaling and other nutrients, such as iron. Under conditions of nutrient imbalance, such as nitrogen deficiency, *ILR3-like* might trigger structural changes in roots, enhancing branching and improving nutrient foraging efficiency. (ii) *bHLH18*, another member of the bHLH transcription factor family, is associated with diverse plant growth and developmental processes (Qin et al. 2022), including root development, stress responses, and hormonal signaling pathways (e.g., auxin and gibberellin). The association between *bHLH18* and root volume suggests that this gene may influence root growth by regulating cell proliferation, expansion, or overall biomass accumulation in roots. (iii) Transcriptional Repressor of *LEUNIG* (*LUH*) likely acts as a positive regulator of root elongation, influencing cell division and expansion processes in the root meristem or elongation zone (Sitaraman et al. 2008). This hypothesis is further supported by the identification of a SNP linked to LUH using root length diameter range 2 (RLDR2) as a trait, suggesting its role in promoting elongation under specific root system adaptations.

The strong associations of these genes with RSA traits suggest they may function in the same or complementary pathways that modulate root structure and function. Further experimental validation—such as expression analysis, mutant studies, and co-expression network analysis—will provide critical insights into their roles and interactions, offering new avenues for understanding and improving root system adaptations to nutrient stress.

In conclusion, our study underscores the vast genetic diversity within the sorghum gene pool, making it a valuable resource for identifying alleles for germplasm improvement. By integrating high-throughput root architecture phenotyping with freely accessible genomic data, enabled us to pinpoint genetic regions associated with RSA adaptation under nutrient-limiting conditions and to infer, plausibly the biological processes involved. This knowledge has far reaching implications for advancing crop breeding strategies aimed at enhancing resilience to both current and future challenges associated with climate change.

Materials and Method

Plant material

A subset comprising 123 accessions of the sorghum reference set [Dataset S1 and (https://genebank.icrisat.org/Common/Viewer?ctg=Referenceset&ref=RS1)] was obtained from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) Nairobi, Kenya and maintained at the Plant Transformation Laboratory (Kenyatta University, Nairobi, Kenya).

Genetic description of the sorghum accession panel

Sorghum distribution maps were generated using MapTool package (Lemmond, 1994) version 2 in R based on metadata available at https://www.morrislab.org/data (Lasky et al. 2015) . To perform population genetic structure analysis, we retrieved SNP data from the repository of Colorado State University (Morris et al. 2013) , available at https://www.morrislab.org/data. Two

methods were used for population genetic structure analysis; phylogenetic relationships and Bayesian-based clustering approaches.

For phylogenetic analysis, we first converted the retrieved SNP HapMap file into Variant Call Format (VCF) using TASSEL version 5.0 (Bradbury et al. 2007) . The resulting VCF file was then utilized to construct a neighbor-joining (NJ) tree with the 'ape' package (Paradis et al. 2004) in R.

For Bayesian clustering, the SNP HapMap file was transformed into PLINK format using TASSEL. We then used ADMIXTURE software to analyze population clusters, evaluating K values ranging from 1 to 10. The optimal number of clusters was determined based on the K value with the lowest cross-validation error. The ADMIXTURE results were visualized using plots generated with the 'ggplot2' package (Villanueva and Chen 2019) in R.

Seedling growth

Sorghum seedlings were sowed in pots measuring 13.5 cm × 11 cm × 11 cm containing vermiculite. For each of the 123 sorghum accessions, 5 seedlings were selected and planted in 3 replicates and 3 treatments of watering with; Long Ashton media (Hudson 1967) , Long Ashton media minus N, and Long Ashton Media minus P. Pots were laid out in a complete randomized block and maintained in the greenhouse for 14 days before phenotyping. Greenhouse conditions were set at a light/ dark cycle of 12/12 h and a temperature of 28℃ day and 24℃ nights.

Root imaging and analysis

On the 14th day, fresh root samples were photographed using a Nikon DX VR 18–140 mm camera at a resolution of 1400 dpi. The acquired images were processed and analyzed using RhizoVision Explorer,

(https://academic.oup.com/aobpla/article/13/6/plab056/6366354?utm_source=chatgpt.com) an open-source software designed for root phenotyping.

Images were firstly scaled and cropped to produce grayscale versions, followed by the application of a global image threshold to distinguish roots (foreground) from the background using the software's 'Broken Roots' mode. The pixel-to-centimeter conversion was calibrated to 41.02 pixels per millimeter. Image thresholding was adjusted to achieve clear root boundaries while minimizing external noise. Non-root artifacts smaller than 2 mm^2 were filtered out to prevent interference with the analysis. Edge smoothing was enabled and set to level 2 to reduce false root appearances during skeletonization, thereby enhancing the delineation of root features. The root pruning threshold was configured to remove false root segments where tips are shorter than the diameter of the primary root, ensuring accurate feature extraction.

Consequently, RhizoVision Explorer output 13 RSA traits, grouped by size (total root length, perimeter, volume, surface area, and network area), distribution (average, maximum, and median diameters), and extent (number of branching points, number of root tips, and branching frequency). These parameters are detailed in Table 1 and illustrated in Figure 1. Raw data is provided in Dataset S2.

Data analysis

Data of the above-mentioned parameters was subjected to analysis of variance (ANOVA). Means and standard deviations of the means were generated and compared across the three study conditions for all the genotypes. The Tukey's HSD test at a 95% confidence interval ($P<0.05$) was used to separate the means and to ascertain whether there is a significant difference in the data set. Pearson correlation coefficients, principal component analysis (PCA) were implemented in R Software version 3.5.10. Root system architecture classification of sorghum under the three conditions was achieved following the procedure as described by (Bodner et al. 2013) based on PCA and computed with FactoMineR (Lê et al. 2008).

Heritability of each trait under the tree conditions was estimated according to (Hill et al., 1998).

$$h^2 = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_{G\times E}^2 \frac{1}{n} + \sigma^2 \frac{1}{n \times r}}$$

Where σ_G^2 represent the genotypic variance, $\sigma_{G\times E}^2$ denotes variance of genotype by environment interaction, σ^2 error of variance, n number of environments and r number of replicates.

Genome wide association studies (GWAS)

Genome-wide association study (GWAS) was performed on root system architecture (RSA) phenotypic data across three experimental conditions: optimal, nitrogen-deficient (minus N), and phosphorus-deficient (minus P). 32,955 single nucleotide polymorphisms (SNPs) were used to perform a multivariate analysis with the Genome Association and Prediction Integration Tool (GAPIT) Version 2 (Tang et al. 2016) , employing two models: the Fixed and Random Model Circulating Probability Unification (FarmCPU) (Kusmec and Schnable 2018) and the Multi-Locus Mixed Model (MLMM) (Segura et al. 2012) .

Markers were filtered to retain those with a minor allele frequency (MAF) of at least 0.05. To correct for multiple testing, we applied a 5% false discovery rate (FDR) implemented in R. The results were visualized using quantile-quantile (QQ) plots and Manhattan plots.

Significant SNPs associated with the phenotypic traits were mapped onto the *Sorghum bicolor* genome using Phytozome v3.1.1's JBrowse tool (Goodstein et al. 2012). Candidate genes of interest were selected based on their proximity to the SNPs and their potential involvement in the regulation of root development.

Author contributions

SR and second author JM conceived and designed the study. First author JM performed the experiments and all analysis guided by SR and CM. SR, and first author JM, wrote the manuscript. All authors read and approved the final manuscript.

Data availability statement

The data that support the findings were derived from the following resources available in the public domain: [https://www.morrislab.org/data and https://genebank.icrisat.org/Common/Viewer?ctg=Referenceset&ref=RS1]. The data that supports the findings of this study are available in the supplementary material of this article. Dataset S1 contain the sorghum accessions metadata and Dataset S2 contain the root system architecture (RSA) data of sorghum accessions used in the study.

Acknowledgements

This work was funded by the National Research Fund grant to S.R. contract number KU/DVCR/NRF/VOL1/27 under the Multi Disciplinary Research Grants. The authors thank and acknowledge the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) through Dr. Eric Manyasa for availing the sorghum germplasm used in this project.

Supplemental information

The following Supporting Information is available for this article:

Figure S1. Genetic diversity of sorghum accessions used for root system architecture analysis.

Table S1: Description of root system architecture (RSA) traits extracted from RhizoVision analysis.

Figure S2. Sorghum root clusters generated under different nutrient conditions.

Figure S3. Multivariate detection of Single Nucleotide Polymorphisms (SNPs) associated with Root System Architecture (RSA) using multiple traits and models

Dataset S1: Sorghum accessions metadata.

Dataset S2. Root system architechure (RSA) data of sorghum accessions used in the study.

Refferences

Bhushan N, Mohnert F, Sloot D, Jans L, Albers C, Steg L. 2019. Using a Gaussian Graphical Model to Explore Relationships Between Items and Variables in Environmental Psychology Research. Front Psychol. 10:1050. doi:10.3389/fpsyg.2019.01050.

Bodner G, Leitner D, Nakhforoosh A, Sobotik M, Moder K, Kaul H-P. 2013. A statistical approach to root system classification. Front Plant Sci. 4:292. doi:10.3389/fpls.2013.00292.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics. 23(19):2633–2635. doi:10.1093/bioinformatics/btm308.

Casa AM, Pressoir G, Brown PJ, Mitchell SE, Rooney WL, Tuinstra MR, Franks CD, Kresovich S. 2008. Community Resources and Strategies for Association Mapping in Sorghum. Crop Sci. 48(1):30–40. doi:10.2135/cropsci2007.02.0080.

Freschet GT, Roumet C, Comas LH, Weemstra M, Bengough AG, Rewald B, Bardgett RD, Deyn GBD, Johnson D, Klimešová J, et al. 2021. Root traits as drivers of plant and ecosystem functioning: current understanding, pitfalls and future research needs. N Phytol. 232(3):1123–1158. doi:10.1111/nph.17072.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, et al. 2012. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 40(Database issue):D1178–D1186. doi:10.1093/nar/gkr944.

Hudson JP. 1967. Sand and Water Culture Methods Used in the Study of Plant Nutrition. :104--104.

Kusmec A, Schnable PS. 2018. FarmCPUpp: Efficient large-scale genomewide association studies. Plant Direct. 2(4):e00053. doi:10.1002/pld3.53.

Lasky JR, Upadhyaya HD, Ramu P, Deshpande S, Hash CT, Bonnette J, Juenger TE, Hyma K, Acharya C, Mitchell SE, et al. 2015. Genome-environment associations in sorghum landraces predict adaptive traits. Sci Adv. 1(6):e1400218. doi:10.1126/sciadv.1400218.

Lê S, Josse J, Husson F. 2008. FactoMineR : An R Package for Multivariate Analysis. J Stat Softw. 25(1). doi:10.18637/jss.v025.i01.

Lynch JP. 2011. Root Phenes for Enhanced Soil Exploration and Phosphorus Acquisition: Tools for Future Crops. Plant Physiol. 156(3):1041–1049. doi:10.1104/pp.111.175414.

Lynch JP. 2013. Steep, cheap and deep: an ideotype to optimize water and N acquisition by maize root systems. Ann Bot. 112(2):347–357. doi:10.1093/aob/mcs293.

Mollier A, Pellerin S. 1999. Maize root system growth and development as influenced by phosphorus deficiency. J Exp Bot. 50(333):487–497. doi:10.1093/jxb/50.333.487.

Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, et al. 2013. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proc National Acad Sci. 110(2):453–458. doi:10.1073/pnas.1215985110.

Niu YF, Chai RS, Jin GL, Wang H, Tang CX, Zhang YS. 2013. Responses of root architecture development to low phosphorus availability: a review. Ann Bot. 112(2):391–408. doi:10.1093/aob/mcs285.

Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics. 20(2):289–290. doi:10.1093/bioinformatics/btg412.

Postma JA, Lynch JP. 2011. Root Cortical Aerenchyma Enhances the Growth of Maize on Soils with Suboptimal Availability of Nitrogen, Phosphorus, and Potassium. Plant Physiol. 156(3):1190–1201. doi:10.1104/pp.111.175489.

Qin T, Kazim A, Wang Y, Richard D, Yao P, Bi Z, Liu Y, Sun C, Bai J. 2022. Root-Related Genes in Crops and Their Application under Drought Stress Resistance—A Review. Int J Mol Sci. 23(19):11477. doi:10.3390/ijms231911477.

Saengwilai P, Nord EA, Chimungu JG, Brown KM, Lynch JP. 2014. Root Cortical Aerenchyma Enhances Nitrogen Acquisition from Low-Nitrogen Soils in Maize. Plant Physiol. 166(2):726–735. doi:10.1104/pp.114.241711.

Seethepalli A, Guo H, Liu X, Griffiths M, Almtarfi H, Li Z, Liu S, Zare A, Fritschi FB, Blancaflor EB, et al. 2020. RhizoVision Crown: An Integrated Hardware and Software Platform for Root Crown Phenotyping. Plant Phenomics. 2020:3074916. doi:10.34133/2020/3074916.

Segura V, Vilhjálmsson BJ, Platt A, Korte A, Seren Ü, Long Q, Nordborg M. 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat Genet. 44(7):825–830. doi:10.1038/ng.2314.

Shimizu A, Yanagihara S, Kawasaki S, Ikehashi H. 2004. Phosphorus deficiency-induced root elongation and its QTL in rice (Oryza sativa L.). Theor Appl Genet. 109(7):1361–1368. doi:10.1007/s00122-004-1751-4.

Sitaraman J, Bui M, Liu Z. 2008. LEUNIG_HOMOLOG and LEUNIG Perform Partially Redundant Functions during Arabidopsis Embryo and Floral Development. Plant Physiol. 147(2):672–681. doi:10.1104/pp.108.115923.

Strock CF, Riva LM de la, Lynch JP. 2017. Reduction in Root Secondary Growth as a Strategy for Phosphorus Acquisition. Plant Physiol. 176(1):691–703. doi:10.1104/pp.17.01583.

Tang Y, Liu X, Wang J, Li M, Wang Q, Tian F, Su Z, Pan Y, Liu D, Lipka AE, et al. 2016. GAPIT Version 2: An Enhanced Integrated Tool for Genomic Association and Prediction. Plant Genome. 9(2):0. doi:10.3835/plantgenome2015.11.0120.

Tissot N, Robe K, Gao F, Grant-Grant S, Boucherez J, Bellegarde F, Maghiaoui A, Marcelin R, Izquierdo E, Benhamed M, et al. 2019. Transcriptional integration of the responses to iron

availability in Arabidopsis by the bHLH factor ILR3. N Phytol. 223(3):1433–1446. doi:10.1111/nph.15753.

Trouche G, Bastianelli D, Hamadou TVC, Chantereau J, Rami J-F, Pot D. 2014. Exploring the variability of a photoperiod-insensitive sorghum genetic panel for stem composition and related traits in temperate environments. Field Crop Res. 166:72–81. doi:10.1016/j.fcr.2014.06.008.

Villanueva RAM, Chen ZJ. 2019. ggplot2: Elegant Graphics for Data Analysis (2nd ed.). Meas Interdiscip Res Perspectives. 17(3):160–167. doi:10.1080/15366367.2019.1565254.

Wissuwa M. 2003. How Do Plants Achieve Tolerance to Phosphorus Deficiency? Small Causes with Big Effects. Plant Physiol. 133(4):1947–1958. doi:10.1104/pp.103.029306.

Yi K, Wu Z, Zhou J, Du L, Guo L, Wu Y, Wu P. 2005. OsPTF1, a Novel Transcription Factor Involved in Tolerance to Phosphate Starvation in Rice. Plant Physiol. 138(4):2087–2096. doi:10.1104/pp.105.063115.

York LM. 2018. Phenotyping Crop Root Crowns: General Guidance and Specific Protocols for Maize, Wheat, and Soybean. Methods Mol Biol (Clifton, NJ). 1761:23–32. doi:10.1007/978-1-4939-7747-5_2.

York LM, Nord EA, Lynch JP. 2013. Integration of root phenes for soil resource acquisition. Front Plant Sci. 4:355. doi:10.3389/fpls.2013.00355.