GRAS Family Transcription Factor Binding Behaviors in Sorghum bicolor, Oyrza, and Maize

Nicholas Gladman[1,2], Sunita Kumari[2], Audrey Fahey[2], Michael Regulski[2], Doreen Ware[1,2].

1. USDA-ARS Robert Holley Center, Ithaca, NY 14853
2. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724

Abstract

Identifying non-coding regions that control gene expression has become an essential aspect of understanding gene regulatory networks that can play a role in crop improvements such as crop manipulation, stress response, and plant evolution. Transcription Factor (TF)-binding approaches can provide additional valuable insights and targets for reverse genetic approaches such as EMS-induced or natural SNP variant screens or CRISPR editing techniques (e.g. promoter bashing). Here, we present the first ever DAP-seq profiles of three GRAS family TFs (SHR, SCL23, and SCL3) in the crop *Sorghum bicolor*, *Oryza sativa japonica*, and *Zea mays*. The binding behaviors of the three GRAS TFs display unique and shared gene targets and categories of previously characterized DNA-binding motifs as well as novel sequences that could potentially be GRAS family-specific recognition motifs. Additional transcriptomic and chromatin accessibility data further facilitates the identification of root-specific GRAS regulatory targets corresponding to previous studies. These results provide unique insights into the GRAS family of TFs and novel regulatory targets for further molecular characterization.

Keywords

GRAS
Transcription Factor
DAP-seq
Sorghum
Maize
Oryza

Introduction

GRAS transcription factors (TFs) form a large family of plant-specific TFs. Named from three members of the family, GIBBERELLIN-ACID INSENSITIVE (GAI), REPRESSOR of GA1 (RGA), and SCARECROW (SCR), these TFs serve a multitude of developmental and environmental response functions and comprise several subfamilies and number in the several dozens of individual proteins across plant lineages(Jaiswal et al., 2022). Members of the GRAS contingent have been shown to be involved in meristem development and axial initiation in tomato, petunia, and arabidopsis(Schumacher et al., 1999; Stuurman et al., 2002; Goldy et al., 2021); influencing meiotic progression in pollen development (Morohashi et al., 2003); fruit development and ripening(Huang et al., 2015; Liu et al., 2021b); seed germination(Lim et al., 2013); arbuscular mycorrhizal symbiosis(Gobbato et al., 2012; Floss et al., 2013; Xue et al., 2015); and light signal transduction as well as plant growth and fertility(Peng et al., 1999; Fukazawa et al., 2014; Fukazawa et al., 2017). This broad onus for GRAS TFs is due, in part, to its expansive cross-talk with numerous hormone signaling pathways, including gibberellic acid(Peng et al., 1997;

Silverstone et al., 1998; Dill and Sun, 2001; King et al., 2001; Fu et al., 2002; Niu et al., 2019), jasmonic acid(Hou et al., 2010), brassinosteroids(Tong et al., 2009; Tong et al., 2012), and auxin(Gao et al., 2004; Sánchez et al., 2007); which also relates to GRAS genes having involvement in numerous stress responses like drought, heat, salinity, cold(Ma et al., 2010; Yang et al., 2011; Yuan et al., 2016), light(Chen et al., 2015), disease resistance(Fode et al., 2008; Wild et al., 2012; Li et al., 2018), and flavonoid production(Pillet et al., 2015; Huang et al., 2021). Initially, it was shown that many GRAS TFs might require the interaction of other proteins like Indeterminate Domain (IDD) TFs to regulate transcription(Welch et al., 2007; Hirano et al., 2017; Aoyanagi et al., 2020), but other structural studies demonstrated the innate ability of certain GRAS TFs to bind DNA without heterodimerization(Li et al., 2016).

Gene regulatory networks (GRNs) have been useful to identify modules that influence plant growth and development(Tu et al., 2020; Zhu et al., 2023; Fu et al., 2024; Khan et al., 2024). Incorporating multiple -omics datasets into these networks improves the power and resolution of their conclusions. While gene expression, epigenetic, and phenotypic profiling information have been useful to unveil regulatory schema, the addition of TF binding information and DNA-binding motif fingerprinting can add significant improvements to GRN construction and interpretation(Savadel et al., 2021; Shojaee and Huang, 2023) and bioengineering targets(Rodríguez-Leal et al., 2017; Liu et al., 2021a; Cao et al., 2022; Yang et al., 2023). Despite the known importance that many GRAS TFs play throughout plants, few TF binding experiments have been conducted on members of this family in model or crop species(Yoon et al., 2016; Tu et al., 2020). By generating TF binding profiles of GRAS proteins, selected regulatory candidate promoters, enhancers, and gene targets can be identified and modified via CRISPR editing for breeding efforts pertaining to root and shoot development(Ron et al., 2014; Triozzi et al., 2021). These targeted approaches can show greater phenotypic variability than creating null or hypomorphic alleles of the TFs themselves(Aguirre et al., 2023).

Sorghum [*Sorghum bicolor (L.) Moench*] is a globally important C_4 grass crop with observed drought, heat, and high-salt tolerances with a completely sequenced genome (2x=2n=10; ~720 Mb)(Paterson et al., 2009; McCormick et al., 2018; Cooper et al., 2019). There are significant discrepancies in understanding the targets and regulatory regions of important TF families in this monocot despite the wealth of genetic resources such as natural diversity panels(Casa et al., 2008) and EMS-mutagenized populations(Jiao et al., 2016; Addo-Quaye et al., 2017) to be used for forward genetics and functional characterization(Jiao et al., 2017; Jiao et al., 2018; Dampanaboina et al., 2019; Gladman et al., 2019) as well as increased genomic profiling of sorghum root, leaf, flower, and seed. In the work presented here, we demonstrate the first DAP-seq profiles of three GRAS family TFs: SHORT ROOT (SHR), SCARECROW-LIKE23 (SCL23), and SCARECROW-LIKE3 (SCL3) in *Sorghum bicolor*; characterize their binding behavior in maize (B73) and oryza (Nipponbare); and demonstrate the identification of conserved binding sites in both promoter and intergenic space through the incorporation of publicly available histone methylation data from root tissues. Further, we extend the potentially novel TF binding motifs discovered through the DAP-seq pipeline via a cross-species projection model using position weight matrices (PWM) to improve the validation of new binding sites. Ultimately, this combined information strengthens the model that some GRAS TFs can bind DNA without interacting with

other TFs while also strengthening tissue-specific candidates for genome editing approaches and can help further refine the functions of GRAS protein GRNs within the root system.

Results

GRAS Family Transcription Factor Selection and Expression

TF profiling was conducted via DNA Affinity Purification (DAP-seq)(O'Malley et al., 2016) using *Sorghum bicolor* BTx623 gDNA. Three GRAS family TFs were chosen to be profiled via DAP-seq based on their ability to be stably expressed and remain soluble in bacterial expression systems and to represent different clades within the GRAS family as defined by Fan et al. 2021: *SHR* (SORBI_3001G327900) from the SHR clade, *SCL23* (SORBI_3002G342800) from the SCR clade, and *SCL3* (SORBI_3005G029600) from the SCL3 clade. Sufficient levels of protein were able to be generated for DNA pulldowns with the addition of mannitol in the bacterial cultures (**Figure 1a**). Bound DNA fragments were eluted, sequenced, and peaks were called using input DNA fragments as controls.

GRAS Transcription Factor Behaviour in Sorghum bicolor

The peak calling from the SHR, SCL23, and SCL3 DAP-seq yielded tens of thousands of peaks genome-wide for all three TFs, with SHR having the most at >100,000 significant peaks throughout the genome (**Supplemental Data Files 1,2, and 3**). A total of 473 genes had all three GRAS TF binding events in their promoter region (<2,000 bp from transcriptional start site (TSS) to 1000bp after the TSS) (**Figure 1b)**. Most peaks for all three TFs were contained within intergenic space, but around 20% of peaks could be discreetly classified to be within proximal promoter regions (**Figure 1c).** Greater than 6000 peaks between SHR, SCL23, and SCL3 overlapped at least somewhat throughout the genome (**Figure 1d**). While not much DNA-binding information is available for GRAS family proteins in monocots to compare against our results, there are ChIP-chip data from Arabidopsis for SHR, SCL23, and SCL3 orthologs (Cui et al., 2014) that was used for comparison. Based on this, there is ~60% overlap with the Arabidopsis genes that have peaks within the promoter of a corresponding sorghum ortholog gene and 75% overlap when including distal peaks (2kb-25Kb upstream from the TSS).

When evaluating functional ontologies for genes with GRAS peaks, it was determined that genes with multiple DAP-seq peaks within the promoter did show enrichment in several interesting categories. Namely, gene promoters with multiple SHR peaks in this 3000 bp promoter window had ontologies associated with root hair development (GO:0080147), salicylic acid mediated signaling (GO:2000031), sulfate transport (GO:1902358), hypoxia detection (GO:0070483), amino acid biosynthesis (GO:0000162, GO:0055129), acyl-CoA metabolism (GO:0006637), menaquinone biosynthesis (GO:0009234), and others. The SCL23 cohort of gene promoters with multiple peaks include biological process of electron transport coupled proton transport (GO:0015990), acyl-CoA metabolism (GO:0006637), amino acid and nutrient transport (GO:0015808, GO:1903401, GO:0015813, GO:0009749, GO:0070574, GO:0071805), and very long-chain fatty acid and sphingolipid biosynthesis (GO:0042761, GO:0006665). Similar to SCL23, the genes with multiple SCL3 peaks had ontological processes of electron transport coupled proton transport (GO:0015990) and amino acid transport (GO:0015808) as well as glutathione metabolic process (GO:0006749), cell wall biogenesis (GO:0042546), and

cellular oxidant detoxification (GO:0098869). When evaluating all the biological process ontologies of the nearest genes annotated as being associated with a region where all three TFs have overlapping peaks, there is an enrichment for detection of abiotic stimulus (GO:0009582), triterpenoid biosynthetic process (GO:0016104), polyketide biosynthesis (GO:0030639), and phosphorelay signal transduction (GO:0000160). Most of the gene targets accounting for the polyketide biosynthesis ontology are chalcone synthases, which have been shown to be influenced by upstream GRAS activity(Pillet et al., 2015; Huang et al., 2021).

When evaluating whether GRAS peaks exist in important root developmental and stress response pathways, we found that all three TFs had peaks associated with genes involved in gibberellic acid, jasmonic acid, phosphate starvation response, and arbuscular mycorrhizal symbiosis. Notably, either SHR, SCL23, or SCL3 have peaks associated with 89% of all sorghum GRAS TFs; at least one of the three TFs have a peak associated with another GRAS family (**Supplemental Data Table 1**).

Sorghum GRAS Transcription Factor Binding in Maize and Oryza

To evaluate the consistency of binding targets of the sorghum GRAS TFs in other monocots, we used maize (B73) and oryza (Nipponbare) DNA in the DAP-seq pulldowns using the sorghum SHR, SCL23, and SCL3 proteins. For both Nipponbare and B73, there were fewer peaks called for all three GRAS TFs compared to the sorghum BTx623 gDNA pulldowns (**Figure 2a,b,d,e and Supplemental Data Table 2 and 3**). When focusing on the sorghum genes with GRAS peaks in the promoter, SHR shared 76 gene orthologs with annotated peaks in B73 and 17 in Nipponbare, SCL23 shared 15 gene orthologs with annotated peaks in B73 and 9 in Nipponbare, and SCL3 shared 22 gene orthologs with annotated peaks in B73 and 6 in Nipponbare. This disparity in peak enrichment using the SbGRAS proteins could be due to the divergence in protein sequence identity between the closest orthologs of SbSHR, SbSCL23, and SbSCL3 in oryza and maize: SHR ortholog sequence identity is ~72-87%; SCL23 ortholog sequence identity is ~85-97%; and SCL3 ortholog identity is ~59-67%. Only the SbSCL3 peaks in B73 were enriched for DNA-binding motifs including bHLH, Myb, bZIP and MADS-box. All three sorghum GRAS TF peaks in Nipponbare were enriched for TF-binding motifs. SbSHR had bZIP, LOB, GATA, ABI3, B3, and TCP motifs. SbSCL23 promoter peaks were enriched for WRKY, NAC, GATA, Myb and GRF motifs. SbSCL3 oryza promoter peaks were enriched for ABI3, WRKY, bZIP, Myb and GRF motifs (**Supplemental Data File 4 and 5**). The SHR, SCL23 and SCL3 motif enrichment analyses in either oryza or maize also yielded uncharacterized motifs that shared similar promoter frequency profiles to those identified from the DAP-seq pulldowns that used sorghum gDNA (see following section) **(Figure 2c,f)**.

Discovery of Novel DNA-binding Motifs in GRAS Family Transcription Factors

A DNA motif enrichment analysis was performed to determine what other TFs were active in the same regulatory space as SHR, SCL23, and SCL3. All GRAS peaks in the promoter region were analyzed using the MEME suite and several different classes of TF binding sites were identified as co-populating the peaks for either SHR, SCL23, or SCL3 (**Supplemental Figure 1 and Supplemental Data File 6**). The TF binding motifs observed to occur within the SHR promoter peaks included family members of AP2/EREBP, ABI3, TCP, NAC, and MYB

(**Supplemental Figure 1a**). TF binding sequences found in SCL23 promoter peaks are AP2/EREBP, ABI3, GRF, DOF, Heat Shock Factor, bZIP, and MYB **(Supplemental Figure 1b**). For SCL3 promoter peaks, DNA-binding motifs were discovered for the TF families AP2/EREBP, ABI3, GARP G2-like, NAC, and MYB **(Supplemental Figure 1c**).

A subset of the enriched DNA motifs in the SHR and SCL23 peaks could be classified as novel GRAS motifs as they did not appear in existing TF-binding profile databases like JASPAR(Rauluseviciute et al., 2024) or Catalog of Inferred Sequence Binding Preferences (CIS-BP)(Weirauch et al., 2014). Additional bioinformatic analyses were conducted to determine if any of these putatively novel GRAS motifs were legitimate TF-binding sites due to the noisy nature of DAP-seq as a TF-profiling experiment. The position weight matrix (PWM) for each of the novel motifs discovered in SHR and SCL23 were projected across the promoter regions of all genes within sorghum (BTx623 v3), maize (B73 v5), and oryza (Nipponbare v1) genomes to determine their frequency of occurrence relative to the TSS (**Figure 3a**). The promoter region was binned into 40 bp intervals and each motif occurrence was counted. The motif sequence frequency in the promoter region, which we propose is unique to the three GRAS family TFs in this analysis, gradually reaches apogee at the -1000 bp position relative to the TSS. Then it decreases abruptly and then has a sharp, bimodal occurrence near the TSS (~-400->+60 bp region), with some SCL23 profiles possibly displaying a 4th peak well into the coding sequence. These motifs had a unique frequency profile compared to other TF family DNA-binding motifs like WRKY, AP2/EREB, and bZIP proteins that can be more unimodal around the TSS but sometimes still display abrupt changes in frequency at different positions relative to the TSS **(Figure 3b).** There is variation of motif frequency conservation across sorghum, maize, and oryza for the SHR motifs; motif #29 shows fairly conserved frequencies across the three monocots, whereas SHR motif #26 shows strong frequency similarity between maize and sorghum, but the oryza profile seems to be ~50% of that. The frequency of these novel, previously uncharacterized motifs from the maize and oryza DAP-seq analysis (previous section) also displayed similar frequency profiles to that of sorghum-derived pulldowns in oryza and sorghum genomes. All sorghum genes that have GRAS TF DAP-seq peaks as well as motif frequency occurrences from the PWM projections can be found in **Supplemental Table 4**. This suggests that 1) the gene modules that could be regulated by these TFs via their promoter regions are not completely conserved, 2) there is some incompleteness in the motif fingerprint projection, 3) GRAS TFs can recognize multiple promiscuous sequences, and/or 4) the motifs we identified only represent a portion of a larger recognition motif that is defined by *in vivo* binding activity. Ultimately, this analysis likely indicates that these previously uncharacterized motifs are more likely to be real DNA-binding sites due to their unique frequency and conserved occurrence around coding sequences in different monocots.

SCL23 Occupancy Around 3'UTR regions

While the focus of DNA-binding profiles tends to fixate around the promoters, TSS, and enhancer regions, TFs can bind in and around the 3' untranslated regions (3'UTR) of gene models (**Supplemental data table 5**). The sorghum SHR, SCL23, and SCL3 TFs all displayed 3'UTR binding across hundreds of genes, often with no other nearby gene models. When evaluating DNA-recognition sites within 3'UTRs that host GRAS peaks, several environmentally

responsive TF binding sites come out, specifically ARR-B/HHO, ABI3, and GRF motifs. There was no enrichment for the putative GRAS motifs identified within promoter peaks at 3'UTR regions. There are no consistent ontology enrichments between the three GRAS TFs for the genes with peaks in their 3'UTRs. However, when manually validating SCL23 peaks that 1) occur within 3'UTRs and 2) are not upstream of a neighboring gene TSS, numerous growth and developmental genes emerge. Some examples are SORBI_3001G261545, a DNA photolyase that is essential for phosphorus starvation response in roots (Nilsson et al., 2007); a strigolactone biosynthesis gene SORBI_3005G168200 that is strongly upregulated in response to limiting phosphorus conditions in roots(Gladman et al., 2022); SORBI_3002G075900, a dolichyl-diphosphooligosaccharide-protein glycosyltransferase protein whose orthologs plays a role in proper primary and lateral root formation in oryza(Qin et al., 2013), a cysteine desulfurase domain (SORBI_3002G174900) and a MATE family transporter (SORBI_3002G232600) that are strongly upregulated in roots in response to abiotic stress(McCormick et al., 2018; Gladman et al., 2022), and an exocyst complex component, SORBI_3003G158700, whose orthologs are involved in a variety of targeted cell secretions for proper cell progression and polarization(Pecenková et al., 2017; Synek et al., 2021).

Shared Peaks Highlight Tissue-specific Gene Expression

Evaluating peaks that are shared between all three GRAS TFs revealed conserved promoter and enhancer elements within the sorghum BTx623 genome. Incorporating epigenetic information with the SHR, SCL23, and SCL3 binding locations also gives more power in identifying real versus spurious binding sites as well as yielding tissue-specific gene expression targets. To do this, all promoter DAP-seq peaks that have partial or near total overlap between SHR, SCL23, and SCL3 and were compared to Histone 3 trimethylation on the 4th lysine (H3K4me3) peaks that were derived from whole BTx623 roots grown in hydroponic conditions during a limiting phosphate experiment (data from Gladman, et al., 2022). H3K4me3 peaks generally indicate active gene expression and are often localized in a narrower fashion around the TSS, however H3K4me3 peaks can also be more broad and exist in the upstream promoter regions or distal enhancer space, and are likely indicative of tissue-specific expression in plants and other eukaryotes(Benayoun et al., 2014; Zhang et al., 2021).

Of the 256 overlapping peak regions that occur within promoters between all three GRAS TFs, 240 of them were associated with unique gene models (16 overlapping peaks occurred multiple times around the same gene element). Of the 240 unique shared peaks, 157 were manually confirmed to have near perfect overlap when visualized on a genome browser and 79 of those had almost complete overlap and exist on a region with notably higher H3K4me3 marks relative to the surrounding genome space (**Supplemental Data Table 6**). While gene expression doesn't always correspond with upstream H3K4me3 peaks, 77 of the 257 overlapping peaks did show statistically different gene expression somewhere within the root system (data from Gladman, et al., 2022) during limiting phosphorus growth conditions. The majority of this significant differential gene expression occurred in the lateral root region, but there were also genes that showed differential expression in the root apex and elongation zone (**Figure 4a)**. Both up- and down-regulated genes were downstream of these strong GRAS TF peaks, so the

presence alone of these TFs and H3K4me3 marks do not fully explain the regulatory nature of the native DNA binding capability of GRAS family proteins during nutrient stress response.

To assess the accuracy of the PWM projections on these shared peaks, 186 of the 256 genes (72.7%) were present in at least one of the PWM projected gene groups from the above analysis (**Figure 4b**). Many of the genes that had near perfect overlap with all GRAS peaks; H3K4me3 marks; significant differential gene expression in the lateral, apex, or elongation root region; and existed in the PWM projections, are involved in cell growth, development, signaling, and environmental response. Some examples are the chromatin remodeler SORBI_3006G038400; the fasciclin-like arabinogalactan protein SORBI_3004G137200; the ABA/WDS protein SORBI_3008G049200 that is induced by water/ABA stress in oryza(Li et al., 2017); and the LysM domain-coding protein SORBI_3002G222500 (**Figure 4c**), which is involved in arbuscular mycorrhizal symbiosis(Yu et al., 2023). SORBI_3005G053501, a defense response gene, shows a greater H3K4me3 signal and is more strongly expressed in the root apex and elongation zone during sufficient phosphorus conditions. SORBI_3002G076100, a G-box TF that has been identified as playing roles in both photomorphogenesis with HY5(Singh et al., 2012) and also root hair development in Arabidopsis(Richter et al., 2011) has shared promoter binding between all three GRAS TFs and was shown to be upregulated in sorghum lateral root regions in response to limiting phosphorus(Gladman et al., 2022). When comparing the gene promoters that contained shared peaks with the PWM projection data on the putative SHR and SCL23-specific motifs, 50 out of 256 promoters (19.5%) had hits for at least one of the two SHR motifs and 156 out of the 256 promoters (49.2%) had hits for at least one of the four SCL23 motifs. Incorporating data from the STRING database, a network could be constructed based on existing protein-protein and co-expression information as well, which resulted in a smaller network that was statistically enriched for purine biosynthesis (GO:0009205 and GO:0006164) and translation (GO:0006412) (**Figure 4d and Supplemental Data Table 7**).

Additionally, this multi-omics integration identified multiple short gene models that have no current functional domain annotation (e.g. SORBI_3007G226900, SORBI_3010G238300, SORBI_3005G145800, SORBI_3010G201332, SORBI_3002G149600, SORBI_3006G024450), yet have significant expression in root tissue or abundant H3K4me3 marks with the trio of GRAS peaks in the promoter, or both. This provides evidence that these are real gene models and likely active in the root transcriptome (**Supplemental Figure 2**). Importantly, when comparing DAP-seq peaks with the conserved cis-regulatory element information from the Conservatory Project(Hendelman et al., 2021), the GRAS-specific regulatory regions did not usually coincide with an evolutionarily constrained cis-regulatory elements, suggesting that while GRAS TFs are quite old in plants, their regulatory sites can undergo re-wiring in a species-specific manner.

Discussion

Prior experimentation on GRAS DNA-binding is scant, especially considering the importance of the family to multiple functions across plant systems. Our DAP-seq profiling demonstrated good agreement with other work in *Arabidopsis*, and targeted genes yielded functional ontologies that could correspond with prior genetic and molecular characterizations of *SHR* and *SCL23* in other species. Specifically, all three GRAS TFs displayed gene target ontology associations for

previously characterized functions like flavonoid biosynthesis(Pillet et al., 2015; Huang et al., 2021) and as well as basal molecular functions, especially with essential cofactor metabolism, cellular respiration, and cell wall/membrane-associated processes. This broad involvement of genetic function for these TFs could reflect the long co-evolution of GRAS genes since they first emerged in pre-vascularized plants.

TF profiling in combination with identifying conserved non-coding sequences has become a powerful method to identify targets for gene editing approaches(Rodriguez-Leal et al., 2019; Hendelman et al., 2021; Liu et al., 2021a; Aguirre et al., 2023) as well as improve gene regulatory networks inference through machine learning approaches(Shojaee and Huang, 2023). Despite DAP-seq being a 'noisy' assay to profile TF binding sites within a genome, it provides a notable benefit over similar assays like ChIP-seq and CUT&RUN: cell-type naive binding. While DAP-seq profiling precludes binding behavior of TFs that require or are modified by cell-type specific cofactors and DNA and histone methylation, native DNA binding behavior of TFs can be assessed and winnowed for high-confidence peaks through the inclusion of tissue-specific transcriptomic and epigenetic data. This creates the capacity to generate GRNs across multiple tissues by performing DAP-seq once and then layering that data with tissue-specific gene expression information and epigenetic marks. We have used this paradigm in our novel characterization of the important plant-specific GRAS family of TFs and combine additional expression and epigenetic data to improve the confidence in calling biologically significant binding sites for SHR, SCL23, and SCL3 in *Sorghum bicolor*.

Confirming the presence of TF binding motifs is more straightforward when working with TFs that have previously characterized binding sequences, such as WRKYs, NACs, bHLH, and others. It is hard to confidently assert that a previously uncharacterized motif sequence is a true TF binding recognition site with DAP-seq data. However, we found that filtering for SHR and SCL23 peaks, for example, that occur in the promoter regions of genes, then evaluating the frequency of those motifs across the promoters of the entire genomes for sorghum, oryza, and maize created additional supportive evidence that 1) those DNA-binding motifs are likely real and 2) they are likely recognized by the GRAS TFs due to the unique pattern of their occurrence relative to other well characterized TFs. This type of bioinformatic approach allows for additional means of identifying putative regulatory targets for TFs that can bind to projected motifs as well as provide insight into the evolutionary conservation of them across species.

Our first-ever DAP-seq profiling of the sorghum SHR, SCL23, and SCL3 proteins using B73 and Nipponbare input DNA yielded significantly fewer peaks in both maize and oryza compared to the original sorghum gDNA template. This disparity in peak enrichment using the SbGRAS proteins could be due to the divergence in protein sequence identity between the closest orthologs of SbSHR, SbSCL23, and SbSCL3 in oryza and maize. Another explanation for this peak enrichment disparity is there is significant rewiring in the promoter and enhancer space of recognition sites for these three GRAS family motifs. However, both of these explanations rely upon the non-conservation of the regulatory motifs across these three species, which confounds our motif occurrence projections as described above and in **Figure 3**. Ultimately, this could reflect that the putative GRAS-specific motifs we identified for SbSHR and SbSCL23 are only

representative of a part of a larger cis-regulatory element sequence that cannot be resolved through the MEME suite program using DAP-seq peak inputs. There is evidence that single DAP-seq might only capture a small window of a larger biologically active binding site for TFs that require complexing with other TFs or cofactors(Li et al., 2023). This hypothesis could absolutely apply to the GRAS TFs, and SHR and SCL23 in particular, since it is well known that they interact with other TFs and protein cofactors as well as each other to heterodimerize and modulate binding functions in the nucleus.

Ultimately, the power of this research lies in the GRNs that are generated. They reveal potential GRAS gene targets that can be leveraged by and for breeding programs and functional research through identification of disruptive alleles and CRISPR editing approaches. Furthermore, this analysis affirms that GRAS TFs do have innate DNA-binding activity without interacting with other protein cofactors or TFs like IDDs. Identifying these regulatory upstream promoter sequences are useful for EMS mutagenesis or CRISPR genome editing approaches to fine tune gene expression for a spectrum of agronomically valuable phenotypes. Additionally, these types of profiling experiments can also reveal tissue-specific regulatory DNA regions that are being acted upon by promiscuous TFs families, which could permit more precise genome editing that could impact intended organs and not have systemic effects across the plant. For example, SHR has been shown to be involved in arbuscular mycorrhizal symbiosis, and indeed through our GRNs and PWM projections, we show that all three GRAS TFs bind to a very narrow promoter region upstream of a gene involved in arbuscular mycorrhizal symbiosis, LysM (SORBI_3002G222500). This promoter region also displays strong epigenetic signals for open chromatin; yielding a strong candidate for EMS mutagenesis or CRISPR genome editing for those interested in nutrient use efficiency. Ultimately, this first-ever GRAS profiling in sorghum combined with additional -omics data has generated a list of useful targets for additional agronomic characterization that span everything from nutrient use efficiency to growth and developmental modification.

Methods

DAP-seq Pulldown and Sequencing

DAP-seq was performed in a modified fashion from the original methods(O'Malley et al., 2016). *Scarecrow Like 3 (SCL3; SORBI_3005G029600), Scarecrow Like 23 (SCL23; SORBI_3002G342800),* and *Short Hair Root (SHR; SORBI_3001G327900)* full length coding sequences (CDSs) were synthesized from TwistBiosciences (pTwist-ENTR Kozak vector) and cloned into the pDEST15 Gateway vector (N-terminal GST-tag). The resulting plasmids were transformed into BL21 competent cells. Expression of GST-tagged TF proteins was induced between OD_{600}=0.500-0.600 by the addition of 1mM isopropyl-beta-D-thiogalactoside (IPTG) (Goldbio: I2481C25) and 0.05M or 0.1M D-mannitol to bacteria in Lysogeny Broth. A concentration of 0.05M D-mannitol was added to SCL3 expressing bacteria, and 0.1M D-mannitol was added to both SCL23 and SHR expressing bacteria. The bacteria were then shaken at 16°C at 220 rpm for 16 hours. Cultures were spun down and pellets resuspended in 0.5M D-mannitol PBS. Cell membranes and plasmid DNA were disrupted by sonicating at 4°C in 10 cycles of 30 sec on, 30 sec off. Soluble fractions of lysate were added to triplicates of MagneGST beads (Promega) suspended in equal volumes of 0.5M D-mannitol PBS and

incubated at 4°C for 1.5hrs undergoing end-over-end mixing to bind GST-tagged transcription factor proteins. The protein-bound beads were then incubated in sheared adaptor-ligated high-purity, sheared genomic DNA of sorghum BTx623, oryza Nipponbare, or maize B73, using between 100-250 ng of genomic DNA per sample.

Genomic DNA from these plants were sheared by Covaris S220 sonicator. Adaptors were ligated to genomic DNA fragments according to NEBNext Ultra II DNA Library Prep Kit and size-selected using AMPure XP beads for fragments larger than 300bp. Transcription factor-bound DNA fragments were then washed three times with the mannitol PBS solution and eluted off the beads by incubating at 98℃ for 5 minutes. qPCR was performed to determine the number of cycles needed during PCR to amplify eluted fragments and add barcodes; this PCR was performed with both elutes and fragmented input DNA as control, using 10-18 cycles depending on the library. PCR products were cleaned up using AMPure XP magnetic beads. Concentration was determined by Qubit HS dsDNA kit. Quality check was performed and average size of the libraries determined by Bioanalyzer on High Sensitivity DNA Chips. qPCR was performed to determine concentration of the barcoded libraries. Samples were pooled and sent for sequencing. *S*equencing was performed at paired end 150 high output using the Illumina NextSeq2000 platform for the sorghum BTx623 samples and on the Illumina Novaseq 6000 platform for the maize B73 and oryza Nipponbare samples. All 36 libraries (3 replicates of SHR, SCL23, and SCL3 pulldowns + Input control) from each sorghum, oryza and maize were multiplexed into 3 total pools (12 libraries per pool), yielding between 22.1-66.3 million reads per sample for the *Sorghum bicolor* DNA query (BTx623), 9.7- 27.2 million reads per sample for the *Zea mays* query (B73), and 8.9-33.8 million reads per sample for the *Oryza sativa (*Nipponbare*)* query.

DAP-seq Bioinformatic Analysis

The FASTQ files were aligned and merged as follows: Trimmomatic(Bolger et al., 2014) was used for FASTQ trimming, followed by BWA mem alignment to the BTx623 v3 genome and MACS3 callpeak (v3.0.0b1) peak calling (using the input controls for background subtraction), and finally the annotatePeaks program from the ChipSeeker package(Wang et al., 2022) was used to associate peaks with gene models from the reference genome files. For ChipSeeker, default values for intergenic space was defined was used >10,000 bp upstream or >300 bp downstream of gene elements. Promoter regions were defined as <2,000 bp from the TSS.

The *Sorghum bicolor* BTx623 v3 reference genome files housed by Gramene(Release 66)(Tello-Ruiz et al., 2022) were used for annotating peaks. Sorghum GFF and GTF files were both used for Chipseeker features functionality; SAMtools was used for various file formatting and manipulation steps, including sorting and merging of the 150-bp paired-end read files. The version 1 *Oryza sativa* Nipponbare reference genome files housed by Gramene (Release 66) (Tello-Ruiz et al., 2022) were used for oryza mapping and peak calling. The version v5 *Zea mays* B73 reference genome files housed by Gramene (Release 66)(Tello-Ruiz et al., 2022) were used for maize mapping and peak calling. Motifs were compared between the results of each genomic DNA background. Motif enrichment analysis was performed using the MEME suite(Bailey et al., 2015). Identification of DNA-recognition motifs were done by comparing DAP-

seq PWMs to the JASPAR nonredundant plant database(Rauluseviciute et al., 2024). Gene ontology analysis was done by submitting genes to the Gene Ontology Consortium online tool (https://geneontology.org/)(Ashburner et al., 2000) and only the Fisher's Exact statistical test was used to calculate for significant enrichment without any correction.

Cis-regulatory Motif Frequency Projections

To determine cis-regulatory elements for GRAS TFs within the promoters of sorghum, maize and oryza, we used a previously described computational prediction pipeline(Eveland et al., 2014; Knauer et al., 2019) (Eveland et al, 2014; Knauer et al 2019) that uses the Search Tool for Occurrence of Regulatory Motifs (STORM) from the Comprehensive Regulatory Element Analysis and Detection (CREAD) suite of tools (Smith et al., 2006; Schones et al., 2007). We identified GRAS PWMs from the MEME analysis (see above) and used these PWMs to identify motifs within the promoter region spanning 3kb upstream and 1kb downstream from the transcription start sites of all of the protein coding genes of sorghum BTx623, maize b73 and oryza Nipponbare. We considered only those motifs that were over-represented (*p*-value <0.001) in promoter sequences of protein coding genes as compared with a background set of the same number of random genomic sequences and these predictions were further filtered based on a PWM-specific median score threshold (i.e., quality score greater than or equal to the median score passed the filter) and a motif occurrence frequency of two or more per promoter.

Supplemental Data Information

Supplemental Data File 1. shr_peaks_sorghum_BTx623.zip
Supplemental Data File 2. scl23_peaks_sorghum_BTx623.zip
Supplemental Data File 3. scl3_peaks_sorghum_BTx623.zip
Supplemental Data File 4. b73_3kb_to_1kb_downstream_tss_meme_output.zip
Supplemental Data File 5. nipponbare_3kb_to_1kb_downstream_tss_meme_output.zip
Supplemental Data File 6. Btx623_3kb_to_1kb_downstream_tss_meme_output.zip
Supplemental Figure 1. DNA binding motifs in GRAS transcription factor peaks
Supplemental Figure 2. Combining GRAS DAP-seq peaks and root H3K4me3 marks
Supplemental Data Table 1. Sorghum GRAS transcription factor geness that have associated peaks from SHR, SCL23, and/or SCL3.
Supplemental Data Table 2. Oryza Nipponbare-mapped peaks for SHR, SLC23, and SLC3
Supplemental Data File 3. Maize B73-mapped peaks for SHR, SLC23, and SLC3
Supplemental Data Table 4. Genes with all thre GRAS TF peaks (SHR, SCL23, SLC3) 2kb upstream to TSS and also belong to the subset of promoters that have the PWM projections from either SHR or SCL23.
Supplemental Data Table 5. 3'UTR region binding as annotated by the ChIPseaker program.
Supplemental Data File 6. Gene promoter peaks that contain SHR, SLC23, and SLC3 within 2kb to TSS.
Supplemental Data Table 7. Shared GRAS peak network using STRING-DB protein-protein interaction and co-expression data.

Data Availability Statement

The DAPseq reads are available at the NCBI SRA repository, BioProject ID = PRJNA1162020.

Author Contributions

NG and DW experimental design. NG performed the peak calling, annotation, and motif analysis. SK performed the PWM frequency projection analysis. AF and MR performed the DNA template preparation and DAP-seq pulldowns. All authors contributed to the manuscript. No AI/LLM was used in the writing of this manuscript.

Acknowledgements

The authors would like to thank members of the plant research community at Cold Spring Harbor Laboratory for their feedback and support. This work was performed with assistance from the US National Institutes of Health Grant S10OD028632-01. This project was funded by the USDA-ARS award number 8062-21000-044-000D and 8062-21000-051-000D.

Work Cited

Addo-Quaye C, Buescher E, Best N, Chaikam V, Baxter I, Dilkes BP (2017) Forward Genetics by Sequencing EMS Variation-Induced Inbred Lines. G3 **7**: 413–425

Aguirre L, Hendelman A, Hutton SF, McCandlish DM, Lippman ZB (2023) Idiosyncratic and dose-dependent epistasis drives variation in tomato fruit size. Science **382**: 315–320

Aoyanagi T, Ikeya S, Kobayashi A, Kozaki A (2020) Gene Regulation via the Combination of Transcription Factors in the INDETERMINATE DOMAIN and GRAS Families. Genes . doi: 10.3390/genes11060613

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet **25**: 25–29

Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME Suite. Nucleic Acids Res **43**: W39–49

Benayoun BA, Pollina EA, Ucar D, Mahmoudi S, Karra K, Wong ED, Devarajan K, Daugherty AC, Kundaje AB, Mancini E, et al (2014) H3K4me3 breadth is linked to cell identity and transcriptional consistency. Cell **158**: 673–688

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**: 2114–2120

Cao L, Wang Z, Ma H, Liu T, Ji J, Duan K (2022) Multiplex CRISPR/Cas9-mediated raffinose synthase gene editing reduces raffinose family oligosaccharides in soybean. Front Plant Sci **13**: 1048967

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**: 30–40

Chen K, Li H, Chen Y, Zheng Q, Li B, Li Z (2015) TaSCL14, a novel wheat (Triticum aestivum L.) GRAS gene, regulates plant growth, photosynthesis, tolerance to photooxidative stress, and senescence. J Genet Genomics **42**: 21–32

Cooper EA, Brenton ZW, Flinn BS, Jenkins J, Shu S, Flowers D, Luo F, Wang Y, Xia P, Barry K, et al (2019) A new reference genome for Sorghum bicolor reveals high levels of sequence similarity between sweet and grain genotypes: implications for the genetics of sugar metabolism. BMC Genomics **20**: 420

Cui H, Kong D, Liu X, Hao Y (2014) SCARECROW, SCR-LIKE 23 and SHORT-ROOT control bundle sheath cell fate and function in Arabidopsis thaliana. Plant J **78**: 319–327

Dampanaboina L, Jiao Y, Chen J, Gladman N, Chopra R, Burow G, Hayes C, Christensen SA, Burke J, Ware D, et al (2019) Sorghum MSD3 Encodes an ω-3 Fatty Acid Desaturase that Increases Grain Number by Reducing Jasmonic Acid Levels. Int J Mol Sci. doi: 10.3390/ijms20215359

Dill A, Sun T (2001) Synergistic derepression of gibberellin signaling by removing RGA and GAI function in Arabidopsis thaliana. Genetics **159**: 777–785

Eveland AL, Goldshmidt A, Pautler M, Morohashi K, Liseron-Monfils C, Lewis MW, Kumari S, Hiraga S, Yang F, Unger-Wallace E, et al (2014) Regulatory modules controlling maize inflorescence architecture. Genome Res **24**: 431–443

Floss DS, Levy JG, Lévesque-Tremblay V, Pumplin N, Harrison MJ (2013) DELLA proteins regulate arbuscule formation in arbuscular mycorrhizal symbiosis. Proc Natl Acad Sci U S A **110**: E5025–34

Fode B, Siemsen T, Thurow C, Weigel R, Gatz C (2008) The Arabidopsis GRAS protein SCL14 interacts with class II TGA transcription factors and is essential for the activation of stress-inducible promoters. Plant Cell **20**: 3122–3135

Fu J, McKinley B, James B, Chrisler W, Markillie LM, Gaffrey MJ, Mitchell HD, Riaz MR, Marcial B, Orr G, et al (2024) Cell-type-specific transcriptomics uncovers spatial regulatory networks in bioenergy sorghum stems. Plant J **118**: 1668–1688

Fukazawa J, Mori M, Watanabe S, Miyamoto C, Ito T, Takahashi Y (2017) DELLA-GAF1 Complex Is a Main Component in Gibberellin Feedback Regulation of GA20 Oxidase 2. Plant Physiol **175**: 1395–1406

Fukazawa J, Teramura H, Murakoshi S, Nasuno K, Nishida N, Ito T, Yoshida M, Kamiya Y, Yamaguchi S, Takahashi Y (2014) DELLAs function as coactivators of GAI-ASSOCIATED FACTOR1 in regulation of gibberellin homeostasis and signaling in Arabidopsis. Plant Cell **26**: 2920–2938

Fu X, Richards DE, Ait-Ali T, Hynes LW, Ougham H, Peng J, Harberd NP (2002) Gibberellin-mediated proteasome-dependent degradation of the barley DELLA protein SLN1 repressor. Plant Cell **14**: 3191–3200

Gao M-J, Parkin I, Lydiate D, Hannoufa A (2004) An auxin-responsive SCARECROW-like transcriptional activator interacts with histone deacetylase. Plant Mol Biol **55**: 417–431

Gladman N, Hufnagel B, Regulski M, Liu Z, Wang X, Chougule K, Kochian L, Magalhães J, Ware D (2022) Sorghum root epigenetic landscape during limiting phosphorus conditions. Plant Direct **6**: e393

Gladman N, Jiao Y, Lee YK, Zhang L, Chopra R, Regulski M, Burow G, Hayes C, Christensen SA, Dampanaboina L, et al (2019) Fertility of Pedicellate Spikelets in Sorghum Is Controlled by a Jasmonic Acid Regulatory Module. Int J Mol Sci. doi: 10.3390/ijms20194951

Gobbato E, Marsh JF, Vernié T, Wang E, Maillet F, Kim J, Miller JB, Sun J, Bano SA, Ratet P, et al (2012) A GRAS-type transcription factor with a specific function in mycorrhizal signaling. Curr Biol **22**: 2236–2241

Goldy C, Pedroza-Garcia J-A, Breakfield N, Cools T, Vena R, Benfey PN, De Veylder L, Palatnik J, Rodriguez RE (2021) The *Arabidopsis* GRAS-type SCL28 transcription factor controls the mitotic cell cycle and division plane orientation. Proc Natl Acad Sci U S A. doi: 10.1073/pnas.2005256118

Hendelman A, Zebell S, Rodriguez-Leal D, Dukler N, Robitaille G, Wu X, Kostyun J, Tal L, Wang P, Bartlett ME, et al (2021) Conserved pleiotropy of an ancient plant homeobox gene uncovered by cis-regulatory dissection. Cell **184**: 1724–1739.e16

Hirano Y, Nakagawa M, Suyama T, Murase K, Shirakawa M, Takayama S, Sun T-P, Hakoshima T (2017) Structure of the SHR-SCR heterodimer bound to the BIRD/IDD transcriptional factor JKD. Nat Plants **3**: 17010

Hou X, Lee LYC, Xia K, Yan Y, Yu H (2010) DELLAs modulate jasmonate signaling via competitive binding to JAZs. Dev Cell **19**: 884–894

Huang W, Xian Z, Kang X, Tang N, Li Z (2015) Genome-wide identification, phylogeny and expression analysis of GRAS gene family in tomato. BMC Plant Biol **15**: 209

Huang X, Li W, Zhang X (2021) Flavonoid scutellarin positively regulates root length through NUTCRACKER. Plant Divers **43**: 248–254

Jaiswal V, Kakkar M, Kumari P, Zinta G, Gahlaut V, Kumar S (2022) Multifaceted roles of GRAS transcription factors in growth and stress responses in plants. iScience **25**: 105026

Jiao Y, Burke J, Chopra R, Burow G, Chen J, Wang B, Hayes C, Emendack Y, Ware D, Xin Z (2016) A Sorghum Mutant Resource as an Efficient Platform for Gene Discovery in Grasses. Plant Cell **28**: 1551–1562

Jiao Y, Burow G, Gladman N, Acosta-Martinez V, Chen J, Burke J, Ware D, Xin Z (2017) Efficient Identification of Causal Mutations through Sequencing of Bulked F 2 from Two Allelic Bloomless Mutants of Sorghum bicolor. Front Plant Sci **8**: 2267

Jiao Y, Lee YK, Gladman N, Chopra R, Christensen SA, Regulski M, Burow G, Hayes C, Burke J, Ware D, et al (2018) MSD1 regulates pedicellate spikelet fertility in sorghum through the jasmonic acid pathway. Nat Commun **9**: 822

Khan A, Tian R, Bean SR, Yerka M, Jiao Y (2024) Transcriptome and metabolome analyses reveal regulatory networks associated with nutrition synthesis in sorghum seeds. Commun Biol **7**: 841

King KE, Moritz T, Harberd NP (2001) Gibberellins are not required for normal stem growth in Arabidopsis thaliana in the absence of GAI and RGA. Genetics **159**: 767–776

Knauer S, Javelle M, Li L, Li X, Ma X, Wimalanathan K, Kumari S, Johnston R, Leiboff S, Meeley R, et al (2019) A high-resolution gene expression atlas links dedicated meristem genes to key architectural traits. Genome Res **29**: 1962–1973

Li J, Li Y, Yin Z, Jiang J, Zhang M, Guo X, Ye Z, Zhao Y, Xiong H, Zhang Z, et al (2017) OsASR5 enhances drought tolerance through a stomatal closure pathway associated with ABA and H2 O2 signalling in rice. Plant Biotechnol J **15**: 183–196

Lim S, Park J, Lee N, Jeong J, Toh S, Watanabe A, Kim J, Kang H, Kim DH, Kawakami N, et al (2013) ABA-insensitive3, ABA-insensitive5, and DELLAs Interact to activate the expression of SOMNUS and other high-temperature-inducible genes in imbibed seeds in Arabidopsis. Plant Cell **25**: 4863–4878

Li M, Yao T, Lin W, Hinckley WE, Galli M, Muchero W, Gallavotti A, Chen J-G, Huang S-SC (2023) Double DAP-seq uncovered synergistic DNA binding of interacting bZIP transcription factors. Nat Commun **14**: 2600

Li S, Zhao Y, Zhao Z, Wu X, Sun L, Liu Q, Wu Y (2016) Crystal Structure of the GRAS Domain of SCARECROW-LIKE7 in Oryza sativa. Plant Cell **28**: 1025–1034

Liu L, Gallagher J, Arevalo ED, Chen R, Skopelitis T, Wu Q, Bartlett M, Jackson D (2021a) Enhancing grain-yield-related traits by CRISPR-Cas9 promoter editing of maize CLE genes. Nat Plants **7**: 287–294

Liu Y, Shi Y, Su D, Lu W, Li Z (2021b) SlGRAS4 accelerates fruit ripening by regulating ethylene biosynthesis genes and SlMADS1 in tomato. Hortic Res **8**: 3

Li X, Liu W, Li B, Liu G, Wei Y, He C, Shi H (2018) Identification and functional analysis of cassava DELLA proteins in plant disease resistance against cassava bacterial blight. Plant Physiol Biochem **124**: 70–76

Ma H-S, Liang D, Shuai P, Xia X-L, Yin W-L (2010) The salt- and drought-inducible poplar GRAS protein SCL7 confers salt and drought tolerance in Arabidopsis thaliana. J Exp Bot **61**: 4011–4019

McCormick RF, Truong SK, Sreedasyam A, Jenkins J, Shu S, Sims D, Kennedy M, Amirebrahimi M, Weers BD, McKinley B, et al (2018) The Sorghum bicolor reference genome: improved assembly, gene annotations, a transcriptome atlas, and signatures of genome organization. Plant J **93**: 338–354

Morohashi K, Minami M, Takase H, Hotta Y, Hiratsuka K (2003) Isolation and characterization of a novel GRAS gene that regulates meiosis-associated gene expression. J Biol Chem **278**: 20865–20873

Nilsson L, Müller R, Nielsen TH (2007) Increased expression of the MYB-related transcription factor, PHR1, leads to enhanced phosphate uptake in Arabidopsis thaliana. Plant Cell Environ **30**: 1499–1512

Niu X, Chen S, Li J, Liu Y, Ji W, Li H (2019) Genome-wide identification of GRAS genes in Brachypodium distachyon and functional characterization of BdSLR1 and BdSLRL1. BMC Genomics **20**: 635

O'Malley RC, Huang S-SC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR (2016) Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. Cell **165**: 1280–1292

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, et al (2009) The Sorghum bicolor genome and the diversification of grasses. Nature **457**: 551–556

Pecenková T, Janda M, Ortmannová J, Hajná V, Stehlíková Z, Žárský V (2017) Early Arabidopsis root hair growth stimulation by pathogenic strains of Pseudomonas syringae. Ann Bot **120**: 437–446

Peng J, Carol P, Richards DE, King KE, Cowling RJ, Murphy GP, Harberd NP (1997) The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses. Genes Dev **11**: 3194–3205

Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, et al (1999) "Green revolution" genes encode mutant gibberellin response modulators. Nature **400**: 256–261

Pillet J, Yu H-W, Chambers AH, Whitaker VM, Folta KM (2015) Identification of candidate flavonoid pathway genes using transcriptome correlation network analysis in ripe strawberry (Fragaria × ananassa) fruits. J Exp Bot **66**: 4455–4467

Qin C, Li Y, Gan J, Wang W, Zhang H, Liu Y, Wu P (2013) OsDGL1, a homolog of an oligosaccharyltransferase complex subunit, is involved in N-glycosylation and root development in rice. Plant Cell Physiol **54**: 129–137

Rauluseviciute I, Riudavets-Puig R, Blanc-Mathieu R, Castro-Mondragon JA, Ferenc K, Kumar V, Lemma RB, Lucas J, Chèneby J, Baranasic D, et al (2024) JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. Nucleic Acids Res **52**: D174–D182

Richter S, Müller LM, Stierhof Y-D, Mayer U, Takada N, Kost B, Vieten A, Geldner N, Koncz C, Jürgens G (2011) Polarized cell growth in Arabidopsis requires endosomal recycling mediated by GBF1-related ARF exchange factors. Nat Cell Biol **14**: 80–86

Rodríguez-Leal D, Lemmon ZH, Man J, Bartlett ME, Lippman ZB (2017) Engineering quantitative trait variation for crop improvement by genome editing. Cell **171**: 470–480.e8

Rodriguez-Leal D, Xu C, Kwon C-T, Soyars C, Demesa-Arevalo E, Man J, Liu L, Lemmon ZH, Jones DS, Van Eck J, et al (2019) Evolution of buffering in a genetic circuit controlling plant stem cell proliferation. Nat Genet **51**: 786–792

Ron M, Kajala K, Pauluzzi G, Wang D, Reynoso MA, Zumstein K, Garcha J, Winte S, Masson H, Inagaki S, et al (2014) Hairy root transformation using Agrobacterium rhizogenes as a tool for exploring cell type-specific gene expression and function using tomato as a model. Plant Physiol **166**: 455–469

Sánchez C, Vielba JM, Ferro E, Covelo G, Solé A, Abarca D, de Mier BS, Díaz-Sala C (2007) Two SCARECROW-LIKE genes are induced in response to exogenous auxin in rooting-competent cuttings of distantly related forest species. Tree Physiol **27**: 1459–1470

Savadel SD, Hartwig T, Turpin ZM, Vera DL, Lung P-Y, Sui X, Blank M, Frommer WB, Dennis JH, Zhang J, et al (2021) The native cistrome and sequence motif families of the maize ear. PLoS Genet **17**: e1009689

Schones DE, Smith AD, Zhang MQ (2007) Statistical significance of cis-regulatory modules. BMC Bioinformatics **8**: 19

Schumacher K, Schmitt T, Rossberg M, Schmitz G, Theres K (1999) The *Lateral suppressor (Ls)* gene of tomato encodes a new member of the VHIID protein family. Proc Natl Acad Sci U S A **96**: 290–295

Shojaee A, Huang S-SC (2023) Robust discovery of gene regulatory networks from single-cell gene expression data by Causal Inference Using Composition of Transactions. Brief Bioinform. doi: 10.1093/bib/bbad370

Silverstone AL, Ciampaglio CN, Sun T (1998) The Arabidopsis RGA gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway. Plant Cell **10**: 155–169

Singh A, Ram H, Abbas N, Chattopadhyay S (2012) Molecular Interactions of GBF1 with HY5 and HYH Proteins during Light-mediated Seedling Development in Arabidopsis thaliana*. J Biol Chem **287**: 25995–26009

Smith AD, Sumazin P, Xuan Z, Zhang MQ (2006) DNA motifs in human and mouse proximal promoters predict tissue-specific expression. Proc Natl Acad Sci U S A **103**: 6275–6280

Stuurman J, Jäggi F, Kuhlemeier C (2002) Shoot meristem maintenance is controlled by a GRAS-gene mediated signal from differentiating cells. Genes Dev **16**: 2213–2218

Synek L, Pleskot R, Sekereš J, Serrano N, Vukašinović N, Ortmannová J, Klejchová M, Pejchar P, Batystová K, Gutkowska M, et al (2021) Plasma membrane phospholipid signature recruits the plant exocyst complex via the EXO70A1 subunit. Proc Natl Acad Sci U S A. doi: 10.1073/pnas.2105287118

Tello-Ruiz MK, Jaiswal P, Ware D (2022) Gramene: A Resource for Comparative Analysis of Plants Genomes and Pathways. Methods Mol Biol **2443**: 101–131

Tong H, Jin Y, Liu W, Li F, Fang J, Yin Y, Qian Q, Zhu L, Chu C (2009) DWARF AND LOW-TILLERING, a new member of the GRAS family, plays positive roles in brassinosteroid signaling in rice. Plant J **58**: 803–816

Tong H, Liu L, Jin Y, Du L, Yin Y, Qian Q, Zhu L, Chu C (2012) DWARF AND LOW-TILLERING acts as a direct downstream target of a GSK3/SHAGGY-like kinase to mediate brassinosteroid responses in rice. Plant Cell **24**: 2562–2577

Triozzi PM, Schmidt HW, Dervinis C, Kirst M, Conde D (2021) Simple, efficient and open-source CRISPR/Cas9 strategy for multi-site genome editing in Populus tremula × alba. Tree Physiol **41**: 2216–2227

Tu X, Mejía-Guerra MK, Valdes Franco JA, Tzeng D, Chu P-Y, Shen W, Wei Y, Dai X, Li P, Buckler ES, et al (2020) Reconstructing the maize leaf regulatory network using ChIP-seq data of 104 transcription factors. Nat Commun **11**: 5089

Wang Q, Li M, Wu T, Zhan L, Li L, Chen M, Xie W, Xie Z, Hu E, Xu S, et al (2022) Exploring Epigenomic Datasets by ChIPseeker. Curr Protoc **2**: e585

Weirauch MT, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I, Cook K, et al (2014) Determination and inference of eukaryotic transcription factor sequence specificity. Cell **158**: 1431–1443

Welch D, Hassan H, Blilou I, Immink R, Heidstra R, Scheres B (2007) Arabidopsis JACKDAW and MAGPIE zinc finger proteins delimit asymmetric cell division and stabilize tissue boundaries by restricting SHORT-ROOT action. Genes Dev **21**: 2196–2204

Wild M, Davière J-M, Cheminant S, Regnault T, Baumberger N, Heintz D, Baltz R, Genschik P, Achard P (2012) The Arabidopsis DELLA RGA-LIKE3 is a direct target of MYC2 and modulates jasmonate signaling responses. Plant Cell **24**: 3307–3319

Xue L, Cui H, Buer B, Vijayakumar V, Delaux P-M, Junkermann S, Bucher M (2015) Network of GRAS transcription factors involved in the control of arbuscule development in Lotus japonicus. Plant Physiol **167**: 854–871

Yang M, Yang Q, Fu T, Zhou Y (2011) Overexpression of the Brassica napus BnLAS gene in Arabidopsis affects plant development and increases drought tolerance. Plant Cell Rep **30**: 373–388

Yang T, Ali M, Lin L, Li P, He H, Zhu Q, Sun C, Wu N, Zhang X, Huang T, et al (2023) Recoloring tomato fruit by CRISPR/Cas9-mediated multiplex gene editing. Hortic Res **10**: uhac214

Yoon EK, Dhar S, Lee M-H, Song JH, Lee SA, Kim G, Jang S, Choi JW, Choe J-E, Kim JH, et al (2016) Conservation and Diversification of the SHR-SCR-SCL23 Regulatory Network in the Development of the Functional Endodermis in Arabidopsis Shoots. Mol Plant **9**: 1197–1209

Yuan Y, Fang L, Karungo SK, Zhang L, Gao Y, Li S, Xin H (2016) Overexpression of VaPAT1, a GRAS transcription factor from Vitis amurensis, confers abiotic stress tolerance in Arabidopsis. Plant Cell Rep **35**: 655–666

Yu H, Bai F, Ji C, Fan Z, Luo J, Ouyang B, Deng X, Xiao S, Bisseling T, Limpens E, et al (2023) Plant lysin motif extracellular proteins are required for arbuscular mycorrhizal symbiosis. Proc Natl Acad Sci U S A **120**: e2301884120

Zhang A, Wei Y, Shi Y, Deng X, Gao J, Feng Y, Zheng D, Cheng X, Li Z, Wang T, et al (2021) Profiling of H3K4me3 and H3K27me3 and Their Roles in Gene Subfunctionalization in Allotetraploid Cotton. Front Plant Sci **12**: 761059

Zhu W, Miao X, Qian J, Chen S, Jin Q, Li M, Han L, Zhong W, Xie D, Shang X, et al (2023) A translatome-transcriptome multi-omics gene regulatory network reveals the complicated functional landscape of maize. Genome Biol **24**: 60

Figure Captions

Figure 1. GRAS family DAP-seq results in *Sorghum bicolor.* A) Immunoblot of Sorghum GST-tagged SHR protein eluted of affinity beads. Proteins were induced, isolated, and bound to affinity beads either in the absence or presence (0.10 M) of D-Mannitol. B) All genes with GRAS peaks in their promoter region called from the three DAP-seq pulldowns for Sorghum SHR, SCL23, and SCL3 mapped to the BTx623 v3 genome. C). Distribution of all significant peaks for SHR, SCL23, and SCL3 relative to gene transcriptional start sites. D) Upset plots of the share of genomic features where SHR, SCL23, and SLC3 are binding within the BTx623 genome.

Figure 2. Sorghum GRAS family binding profiles in maize B73 and oryza Nipponbare. A) Distribution of all significant peaks for SbSHR, SbSCL23, and SbSCL3 relative to gene transcriptional start sites in the maize B73 genome. B) All maize B73 genes with GRAS peaks in their promoters. Called from the three DAP-seq pulldowns for Sorghum SbSHR, SbSCL23, and SbSCL3 using maize B73 DNA as the template. C) Uncharacterized DNA motifs that were detected in the SbSCL3 pulldowns in B73. D) Distribution of all significant peaks for SbSHR, SbSCL23, and SbSCL3 relative to gene transcriptional start sites in the Nipponbare genome. E) All rice Nipponbare genes with GRAS peaks in their promoters. Called from the three DAP-seq pulldowns for Sorghum SbSHR, SbSCL23, and SbSCL3 using Nipponbare DNA as the template. F) Uncharacterized DNA motifs that were detected in the SbSCL23 and SHR pulldowns in Nipponbare.

Figure 3. DAP-seq PWM projections in sorghum, maize, and oryza. A) The SHR motifs and SCL23 uncharacterized motifs identified in the BTx623 DAP-seq analysis. B) are some examples of common TF motifs projected across the genomes (taken from JASPAR).

Figure 4. Expression and network analysis of the 240 genes whose promoters have shared binding peaks from SHR, SCL23 and SLC3. A) Log2-fold expression heatmap of genes that have peaks in their promoter region from all three SHR, SCL23, and SCL3. Tissue expression data is from the root apex, elongation zone, and lateral root regions (data from Gladman, *et al.* 2022). B) Log2-fold expression of (A) with colorized indication if those genes also were identified from the PWM projection analysis. C). Integrated Genome Viewer display of a LysM gene (SORBI_3002G222500) that has strong overlap between SHR, SCL23, and SCL3 and the localization of all three peaks corresponds with H3Kme3 pileup taken from whole root samples grown in hydroponics under normal and limiting phosphorus conditions (data from Gladman, *et al.* 2022). D) Cytoscape network display of available protein-protein interaction and co-expression data from the STRING-DB resource. The larger network on the lower right is enriched basic molecular ontologies (purine biosynthesis and translation).

Supplemental figure 1. DNA binding motifs in GRAS transcription factor peaks. These are the DNA recognition motifs that were enriched under the DAP-seq peaks within gene promoters for A) SHR, B) SCL23, and C) SLC3

Supplemental Figure 2. Combining GRAS DAP-seq peaks and root H3K4me3 marks. Integrated Genome Viewer images of examples of short genes (protein polypeptide product < 50 amino acids) that have overlapping GRAS family DAP-seq peaks and H3Kme3 pileups in the promoter region. (Histone methylation data taken from Gladman *et al.*, 2022).

Figure 1

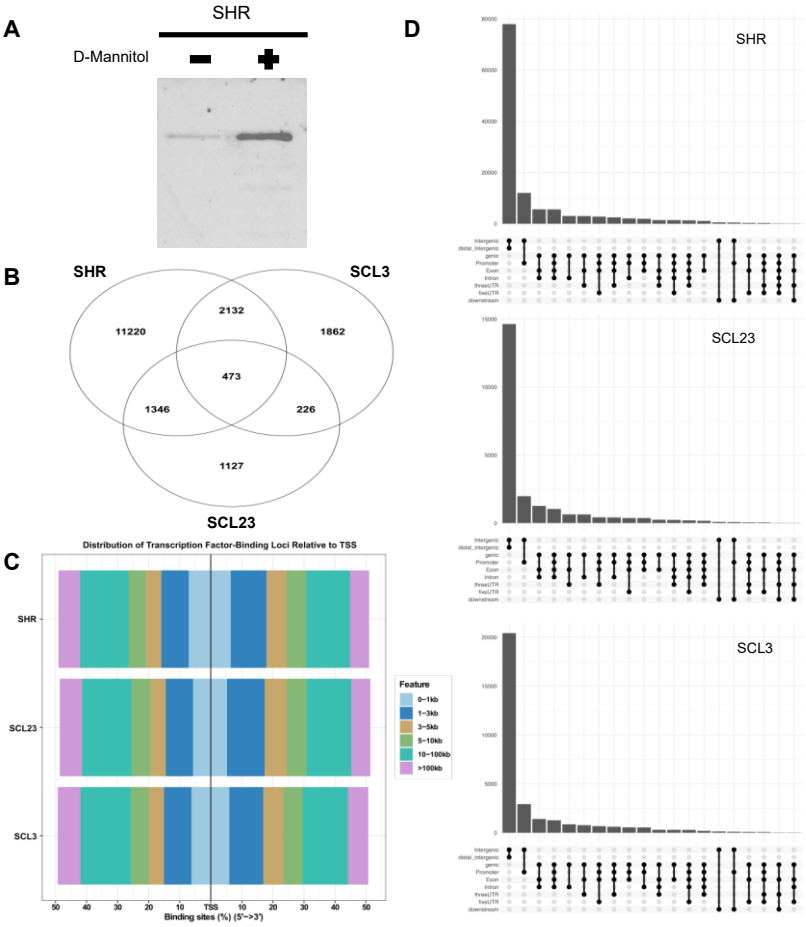


Figure 2

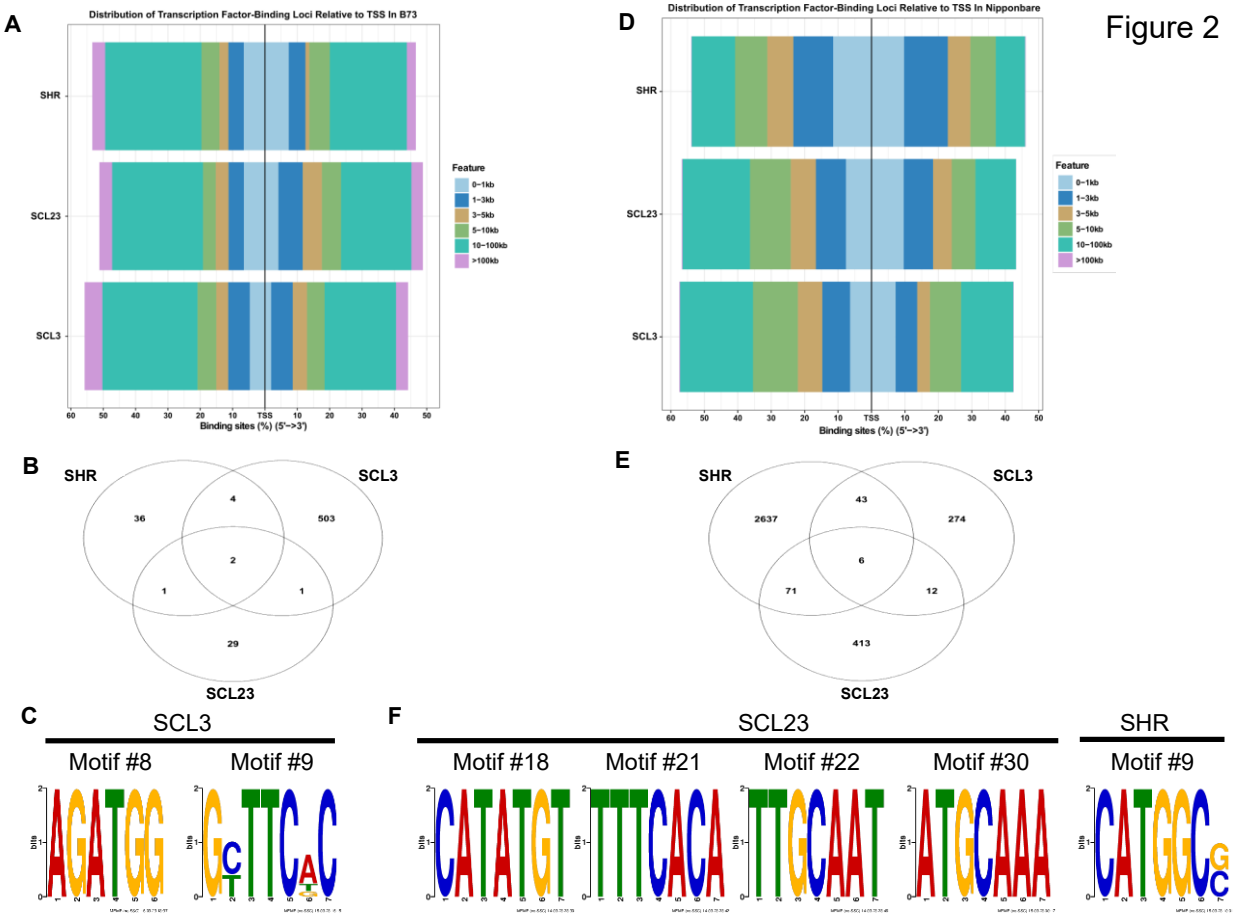
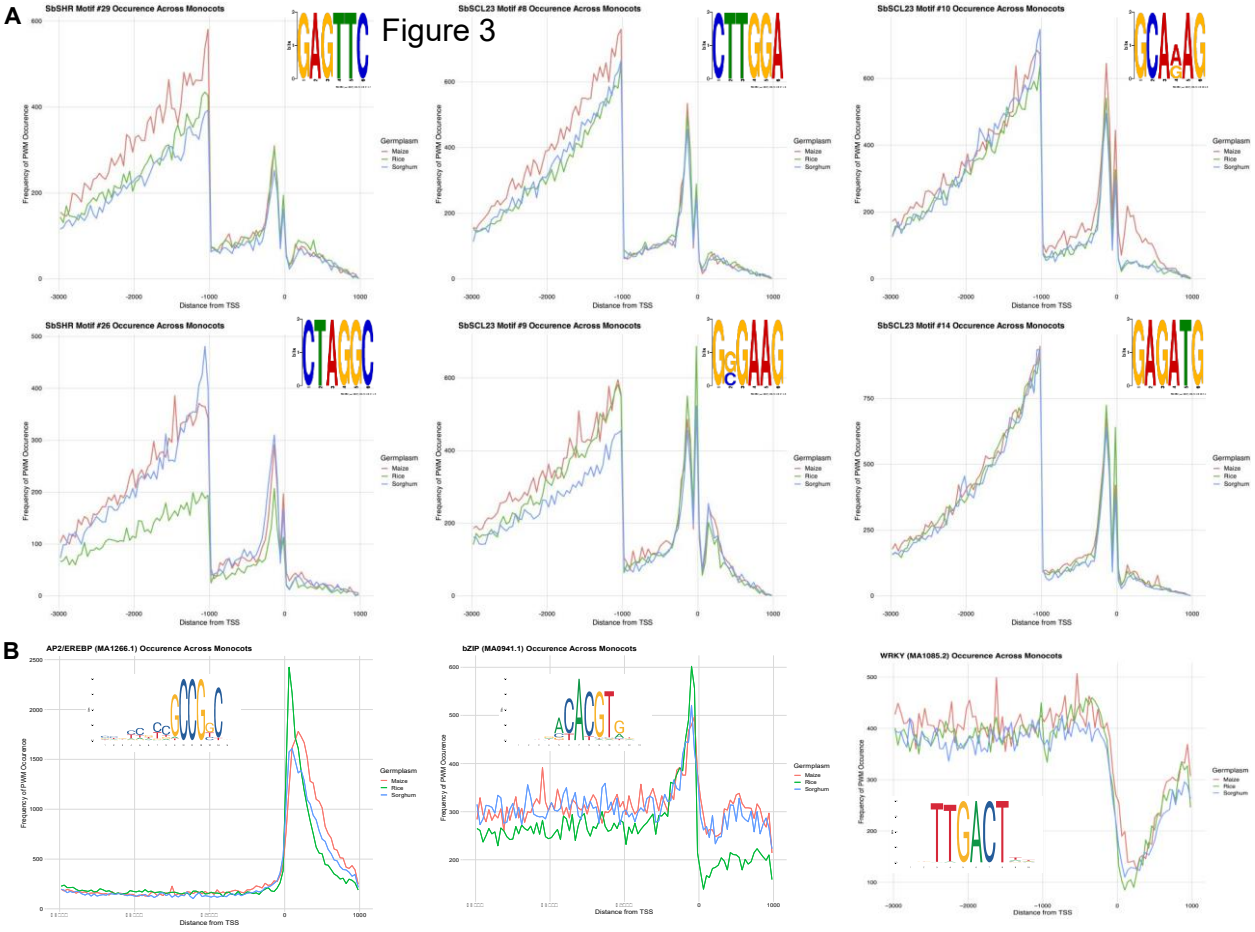
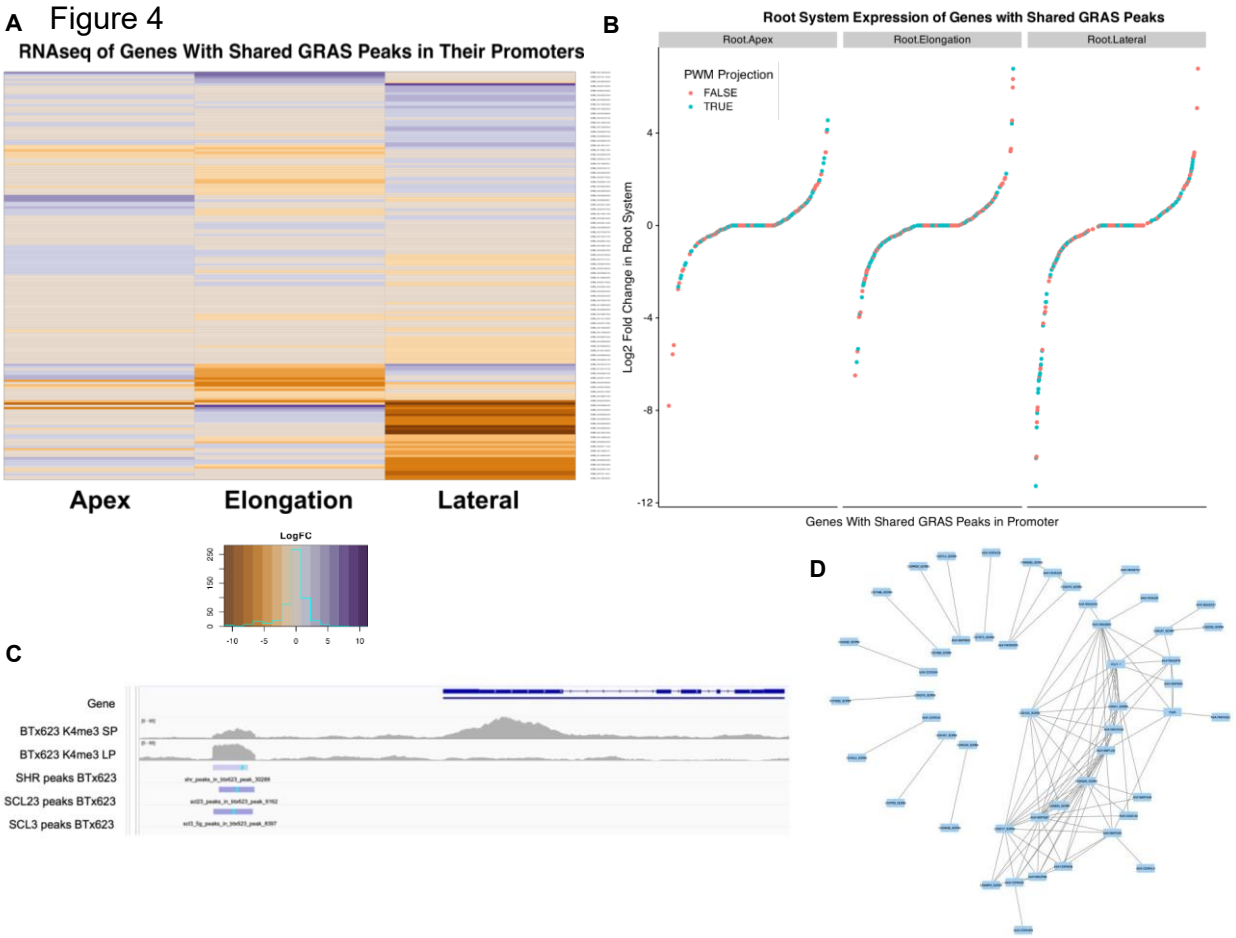


Figure 3



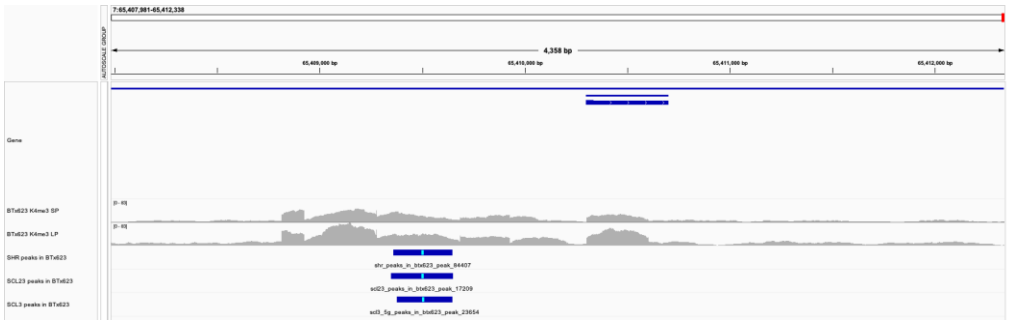
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

Figure 4

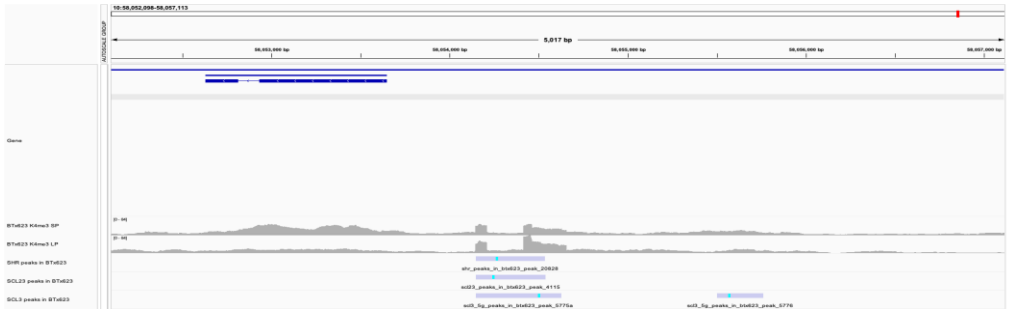


Supplemental Figure 2

SORBI_3007G226900



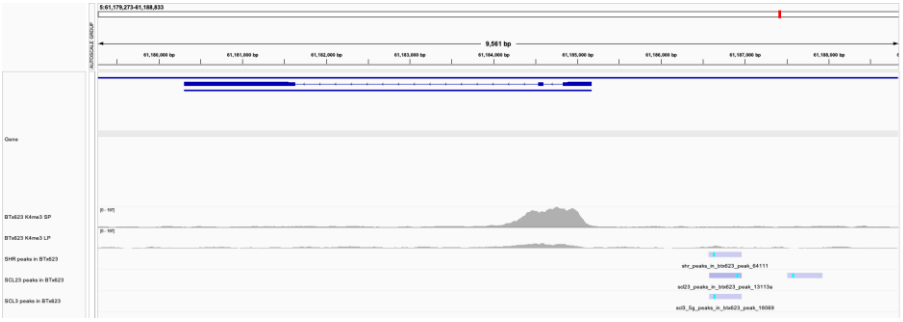
SORBI_3010G238300



856
857

Supplemental Figure 2 (continued)

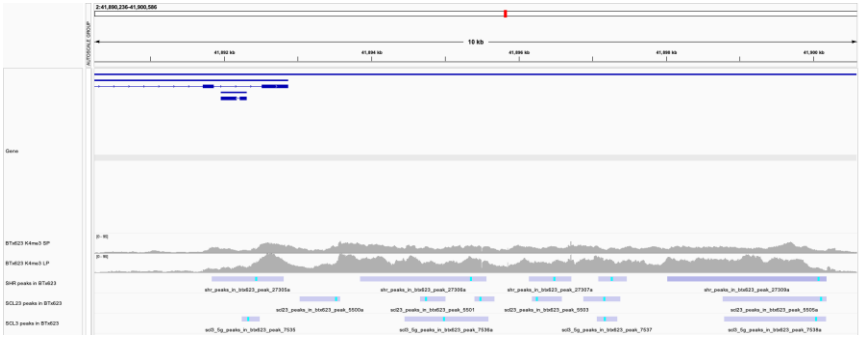
SORBI_3005G145800



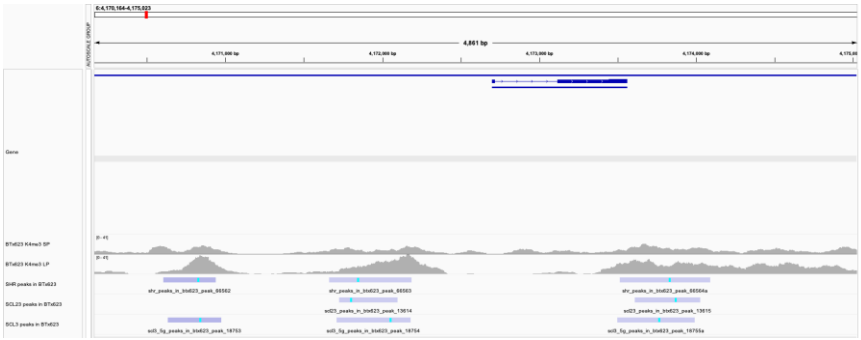
SORBI_3010G201332

Supplemental Figure 2 (continued)

SORBI_3002G149600



SORBI_3006G024450



865