

Chromosome-scale pearl millet genomes reveal *CLAMT1b* as key determinant of strigolactone pattern and *Striga* susceptibility

Received: 11 February 2024

Accepted: 1 August 2024

Published online: 12 August 2024

 Check for updates

Hendrik N. J. Kuijer^{1,2,9}, Jian You Wang^{1,2,9}, Salim Bougouffa^{1,3,9}, Michael Abrouk^{1,2,4}, Muhammad Jamil^{1,2}, Roberto Incitti^{1,3}, Intikhab Alam^{1,3}, Aparna Balakrishna^{1,2}, Derry Alvarez^{1,2}, Cristina Votta^{1,5}, Guan-Ting Erica Chen^{1,2,4}, Claudio Martinez⁶, Andrea Zuccolo^{1,2,4,7}, Lamis Berqdar^{1,2}, Salim Sioud⁸, Valentina Fiorilli⁵, Angel R. de Lera⁶, Luisa Lanfranco⁵, Takashi Gojobori³, Rod A. Wing^{1,2,4}, Simon G. Krattinger^{1,2,4}, Xin Gao^{1,3} & Salim Al-Babili^{1,2,4} ✉

The yield of pearl millet, a resilient cereal crop crucial for African food security, is severely impacted by the root parasitic weed *Striga hermonthica*, which requires host-released hormones, called strigolactones (SLs), for seed germination. Herein, we identify four SLs present in the *Striga*-susceptible line SOSAT-C88-P10 (P10) but absent in the resistant 29Aw (Aw). We generate chromosome-scale genome assemblies, including four gapless chromosomes for each line. The *Striga*-resistant Aw lacks a 0.7 Mb genome segment containing two putative *CARLACTONOIC ACID METHYLTRANSFERASE1* (*CLAMT1*) genes, which may contribute to SL biosynthesis. Functional assays show that P10*CLAMT1b* produces the SL-biosynthesis intermediate methyl carlactonoate (MeCLA) and that MeCLA is the precursor of P10-specific SLs. Screening a diverse pearl millet panel confirms the pivotal role of the *CLAMT1* section for SL diversity and *Striga* susceptibility. Our results reveal a reason for *Striga* susceptibility in pearl millet and pave the way for generating resistant lines through marker-assisted breeding or direct genetic modification.

Strigolactones (SLs) inhibit shoot branching and are released by plant roots into the rhizosphere to attract symbiotic mycorrhizal fungi (AM), particularly under phosphate (Pi) starvation^{1–3}. However, SLs also act as germination stimulants for root-parasitic weeds, such as *Orobanche* and *Striga spp*^{4,5}, posing severe agricultural problems worldwide⁶. The yield of pearl millet (*Pennisetum glaucum*; syn. *Cenchrus americanus*), the sixth most important grain globally⁷, primarily cultivated in subtropical regions, including Sub-Saharan Africa and India, is significantly affected by *Striga hermonthica*⁸, which is considered one of the seven major threats to global food security⁹.

After seed germination, *Striga* develops a haustorium that connects the emerging seedling to the host root. Host resistance to *Striga* comes in two forms. The first, post-attachment resistance, involves physical blocking, immune responses, or prevention of a vascular connection, and is triggered following haustorium attachment^{10–12}. The second is a result of the amount and pattern of released SLs (pre-attachment) as has been proposed in studies regarding sorghum, rice and maize^{13–17}.

The wild pearl millet line 29Aw (Aw) from Niger exhibits resistance to *Striga* through both pre- and post-attachment mechanisms.

A full list of affiliations appears at the end of the paper. ✉ e-mail: Salim.Babili@kaust.edu.sa

In contrast, SOSAT-C88 P10 (P10) is a susceptible millet line derived from the SOSAT variety, which originates from a cross between the landraces Sauna and Sanio and has high yields in West Africa^{11,18}.

In this work, we use two contrasting pearl millet lines Aw and P10 to determine the mechanisms of Striga susceptibility. First, we analyze the SLs produced by P10 and identify four SLs that are absent in Aw. To elucidate responsible genes, we sequence the Aw and P10 genomes and analyze their transcriptomes, resulting in chromosome-scale assemblies and the identification of candidate genes for SL biosynthesis in pearl millet. Of these, *CLAMT1b* is located on a genomic fragment that is present in P10 and absent in Aw. Transient expression in tobacco leaves shows that CLAMT1b acts in SL biosynthesis by converting carlactonic acid (CLA) into methyl carlactonoate (MeCLA). Feeding Aw with MeCLA results in the production of some of the P10-specific SLs. Analysis of a panel of pearl millet lines shows a correlation between the presence of *CLAMT1b*, the release of P10 specific SLs and

susceptibility to Striga. These results reveal a reason for Striga susceptibility in pearl millet, paving the way for generating Striga-resistant pearl millet lines.

Results and discussion

P10 releases four SLs with high Striga germination activities

To investigate the underlying resistance mechanisms, we first verified the contrasting phenotypes of the two lines under greenhouse conditions (Fig. 1a–d). Root exudates from P10 induced higher Striga germination than those from Aw (Fig. 1c), suggesting that P10 susceptibility may be related to released SLs. To test this assumption, we quantified the SLs exuded by both lines under Pi-deficient conditions using LC-MS/MS. Surprisingly, Aw exuded a significantly higher amount of the two canonical SLs orobanchol and orobanchyl acetate than P10 (Fig. 1e). However, we identified four SLs of unknown structure in P10 exudates, based on the characteristic D-ring product ion

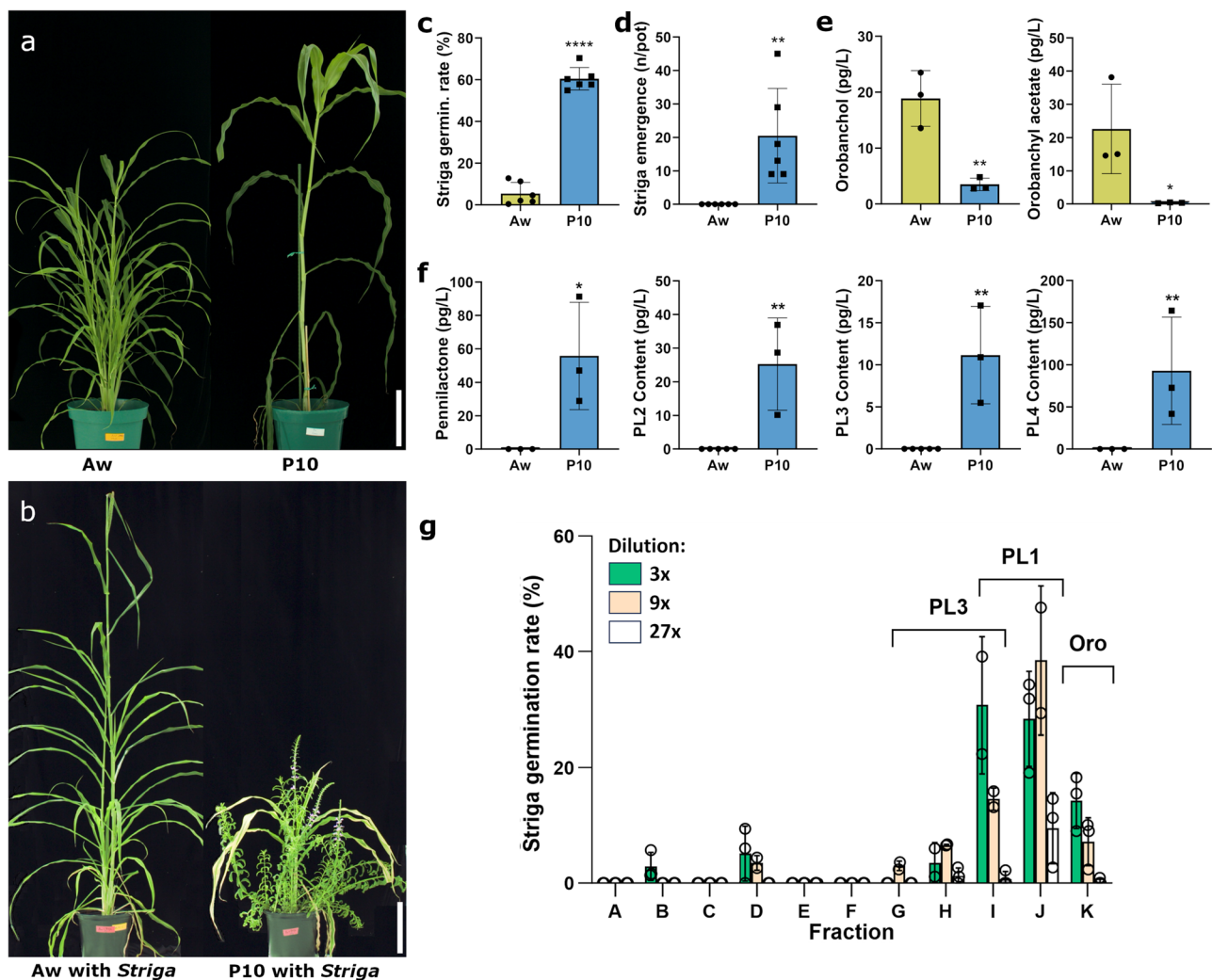


Fig. 1 | Aw and P10 are contrasting lines for Striga susceptibility and strigolactone production. **a** The wild accession, Aw, was shorter and had several tillers, whereas the domesticated P10 was a tiller monocult. **b** Aw and P10 grown in Striga-infested soil. Aw exhibited no emerging Striga and was minimally affected in growth, while P10 was heavily infested by emerging Striga, severely impacting its growth and development. **c** Less than 10% of the Striga seeds exposed to Aw's root exudate germinated, whereas exposure to P10's root exudate stimulated germination in over 70% of the Striga seeds ($n = 6$ plates; $P < 0.0001$). **d** When grown in Striga-infested soil, P10 induced an average of 20 emerging Striga plants, while Aw induced none after 70 days ($n = 6$ plants; $P = 0.0053$). **e** The root exudate of both

Aw and P10 contained orobanchol and orobanchyl acetate, with significantly lower levels in P10 ($n = 3$ seedlings; Oro, $P = 0.0066$; Oro ace, $P = 0.0459$). **f** Four new strigolactones were observed in P10's root exudate but not detectable in Aw's ($n = 3$ seedlings; except $n = 5$ for PL2 and PL3 Aw; P values for PL1 = 0.0397; PL2 = 0.0047; PL3 = 0.0038; PL4 = 0.0065). **g** Fractionation of SLs and subsequent assessment of individual fractions revealed that fractions G through J, containing PL and PL3, significantly induced Striga seed germination ($n = 3$ seedlings). Size bars indicate 20 cm. Error bars represent the mean \pm s.d. Significant differences were tested using a two-tailed t test ($*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$). Source data are provided as a Source Data file.

peak at m/z 97.028. We named these SLs that were absent in Aw exudates as follows: Pennilactone1 (PL1) with a molecular ion formula $C_{20}H_{23}O_6$ and a mass-to-charge ratio (m/z) at 359.14891 as a protonated positive ion ($[M+H]^+$), Pennilactone2 (PL2) with a molecular ion formula $C_{20}H_{25}O_7$ at m/z 377.15939 as $[M+H]^+$, PL3 with a molecular ion formula $C_{20}H_{23}O_7$; at m/z 375.14386 as $[M+H]^+$, and PL4 with molecular ion formula $C_{19}H_{32}O_{12}$ at m/z 452.19260 as $[M+H]^+$, (Fig. 1f, Supplementary Fig. 1, 2).

Next, we evaluated the seed-germinating activity of the P10 SLs after fractionation using silica gel and detected significant activity in fractions enriched with PL1 and PL3 (Fig. 1g). Confirming the SL identity of PL1, PL2, PL3 and PL4, we observed a significant decrease in their levels following zaxinone application, or treatment with its mimics, Mimic of zaxinone3 (MiZax3) and Mimic of zaxinone5 (MiZax5), which act as negative regulators of SL biosynthesis in rice^{19,20} (Supplementary Fig. 3a). Additionally, the treatment with zaxinone reduced *Striga* infestation, with a similar trend for its mimics, underscoring the role of SLs in P10 *Striga* susceptibility (Supplementary Fig. 3b, c). Our results suggest that the differing SL compositions of Aw and P10 likely account for their contrasting pre-attachment *Striga* resistance phenotypes.

Generation of chromosome-scale genome assemblies of Aw and P10

To determine the genetic differences between Aw and P10 that underlie their contrasting SL compositions, we sequenced both genomes using PacBio HiFi Sequel II, allocating three SMRT cells to each, approximately 57-fold and 51-fold coverage, respectively. The contig N50 values obtained using hifiasm corresponded to 244 Mb for P10 and 284 Mb for Aw (Fig. 2a). The assemblies were scaffolded using Omni-CTM technology. The final chromosome-scale genome assemblies both contained four gap-free chromosomes out of seven, with an order of magnitude fewer gaps than the previous best pearl millet assemblies, and total lengths of 1.915 Gb and 1.926 Gb for Aw and P10, respectively. We named and oriented the chromosomes of Aw and P10 based on the first reference genotype Tift 23D2B1-PI-P5²¹. Guided by RNA sequencing and homology with genes from related grasses, gene structure annotation identified 38,920 high-confidence genes in Aw and 40,869 in P10, with BUSCO completeness scores of 92.9% for Aw and 93.3% for P10 (Supplementary Table 1). The amounts of transposable element-related sequences were highly similar in the two accessions: 1.59 and 1.55 Gb corresponding to 82.90% and 80.50% of the total genome size for Aw and P10, respectively (Supplementary Table 2). These values, which were larger and assigned to more specific subcategories than those reported by Varshney et al.²¹ and Ramu et al.²², aligned with the expectation that they represented at least 80% of the entire genome²¹. Our genome assemblies and findings are consistent with and improve upon previous pearl millet genome assemblies^{21–24} (Supplementary Table 3, Supplementary Note 1).

No large striking structural rearrangements existed between the genomes of Aw and P10 (Fig. 2a). However, when homologs of SL biosynthesis genes were mapped to specific loci on Aw and P10 chromosomes (Fig. 2a), we found that a smaller, 0.7 Mb segment of P10 chromosome 2 was absent in the Aw genome (Fig. 2b, Supplementary Fig. 4).

A putative SL biosynthetic gene, *CLAMT1b*, is only present in P10

Genome-wide association studies on *Striga* resistance in sorghum and maize, both closely related to millets, as well as in pearl millet itself, have surprisingly not yielded many SL biosynthetic genes^{12,25–27}. Although a comparison of the candidate genes from these studies with their homologs between Aw and P10 did not yield any results for pre-attachment resistance, it did indicate potential post-attachment resistance candidates (Supplementary Note 2; Supplementary Data 1; Supplementary Fig. 5).

To identify candidate genes for SL biosynthesis in pearl millet, including the production of P10-specific SLs, we conducted a transcriptomic study on the root tissue of Aw and P10 under low Pi conditions, which stimulate SL biosynthesis, with and without treatment with MP3, which is an SL analog and therefore expected to decrease the transcript levels of SL biosynthetic genes²⁸. The total number of differentially expressed genes (DEGs) under both conditions was 5,833 in Aw and 6,890 in P10, with most being affected only by the low Pi condition (Supplementary Fig. 6a). The homologs of SL biosynthetic genes in pearl millet *PgD27* (*PgP10c0101G011332*), *PgCCD8* (*PgP10c0601G021829*), *PgMAX1-1400* (*PgP10c0601G022366*), *PgCYP706* (*PgP10c0401G051155*) and *PgCLAMT1b* (*PgP10c0201G045110*) (Supplementary Figs. 7, 8, 9) were all markedly upregulated in P10 under low Pi conditions (over 32-fold) and further by MP3 treatment. The canonical SL analog GR24 suppresses SL biosynthetic genes in rice, even under Pi deficiency²⁸. Therefore, the unexpected increase in expression observed here may be because MP3, a non-canonical SL analog, acts differently from GR24, or because SL biosynthesis is regulated differently in pearl millet compared to rice. We identified 30 and 42 genes induced by both treatments in P10 and Aw, respectively (Supplementary Fig. 6b; Supplementary Table 4). Notably, *CARLACTONOIC ACID METHYLTRANSFERASE1b* (*CLAMT1b*) was absent in the Aw genome, as were *CLAMT1c* (*PgP10c0201G045096.1*), a *CYP51* homolog and a putative acyl transferase encoded by the 0.7 Mb fragment present in P10 (Fig. 2b). CLAMT enzymes, involved in SL biosynthesis in rice and maize^{15,28}, convert carlactonoic acid (CLA) into methyl carlactonoate (MeCLA)²⁹ in the formation of non-canonical SLs. In pearl millet, we identified three *CLAMT* genes with differing responses to phosphate starvation and treatment with MP3. While the *CLAMT1a* (*PgP10c0601G022359/PgAwc0601G022076.2*) transcript levels did not respond to either treatment, *CLAMT1c* was upregulated only under low Pi, and *CLAMT1b* (Supplementary Fig. 9) transcripts increased strongly with both treatments, making it the primary candidate for the CLAMT function in pearl millet (Fig. 3a).

CLAMT1b converts CLA into MeCLA, a precursor of pennilactone1 and pennilactone2

To test this assumption, we co-expressed *PgCLAMT1b* with *OsD27*, *OsCCD7*, *OsCCD8* and *AtMAX1*, which give rise to CLA production, in tobacco leaves³⁰. Subsequent LC-MS analysis revealed the formation of MeCLA, confirmed by comparison with a MeCLA standard based on the retention time and mass fragmentation patterns (Fig. 3b; Supplementary Fig. 10). In contrast, co-expression of the CLA-forming enzymes with either *CLAMT1a* or *CLAMT1c* did not result in MeCLA formation (Supplementary Figs. 11, 12), suggesting that *CLAMT1b* is crucial for MeCLA production in pearl millet. These results also suggest that MeCLA formation might be a necessary step for the biosynthesis of P10 non-canonical SLs, which could explain their absence in Aw.

To test this hypothesis, we treated Aw seedlings grown under low Pi conditions with synthetic *rac*-MeCLA. Following feeding, we investigated the SL pattern of the root exudates and detected PL1 and PL2, as evidenced by the retention time, accurate mass, and fragmentation pattern (Fig. 3c; Supplementary Figs. 13, 14). We also observed the formation of a tentative isomer of each of the two SLs. Additionally, we proposed a putative structure for PL1 based on the mass difference and MS/MS fragmentations following MeCLA treatment (Supplementary Fig. 15). PL3 and PL4 were not detected, indicating that additional biosynthetic enzymatic activities, potentially absent in Aw, are required. This feeding experiment demonstrates CLAMT's role in determining the SL pattern in pearl millet and explains the absence of PL1 and PL2 in Aw. Based on chemical and genetic evidence, we established a model for the SL biosynthetic pathway in pearl millet (Fig. 3d). This model proposes that both Aw and P10 produce orobanchol and orobanchyl acetate; however, P10 has a unique pathway branch initiated by *CLAMT1b* that leads to MeCLA formation, which

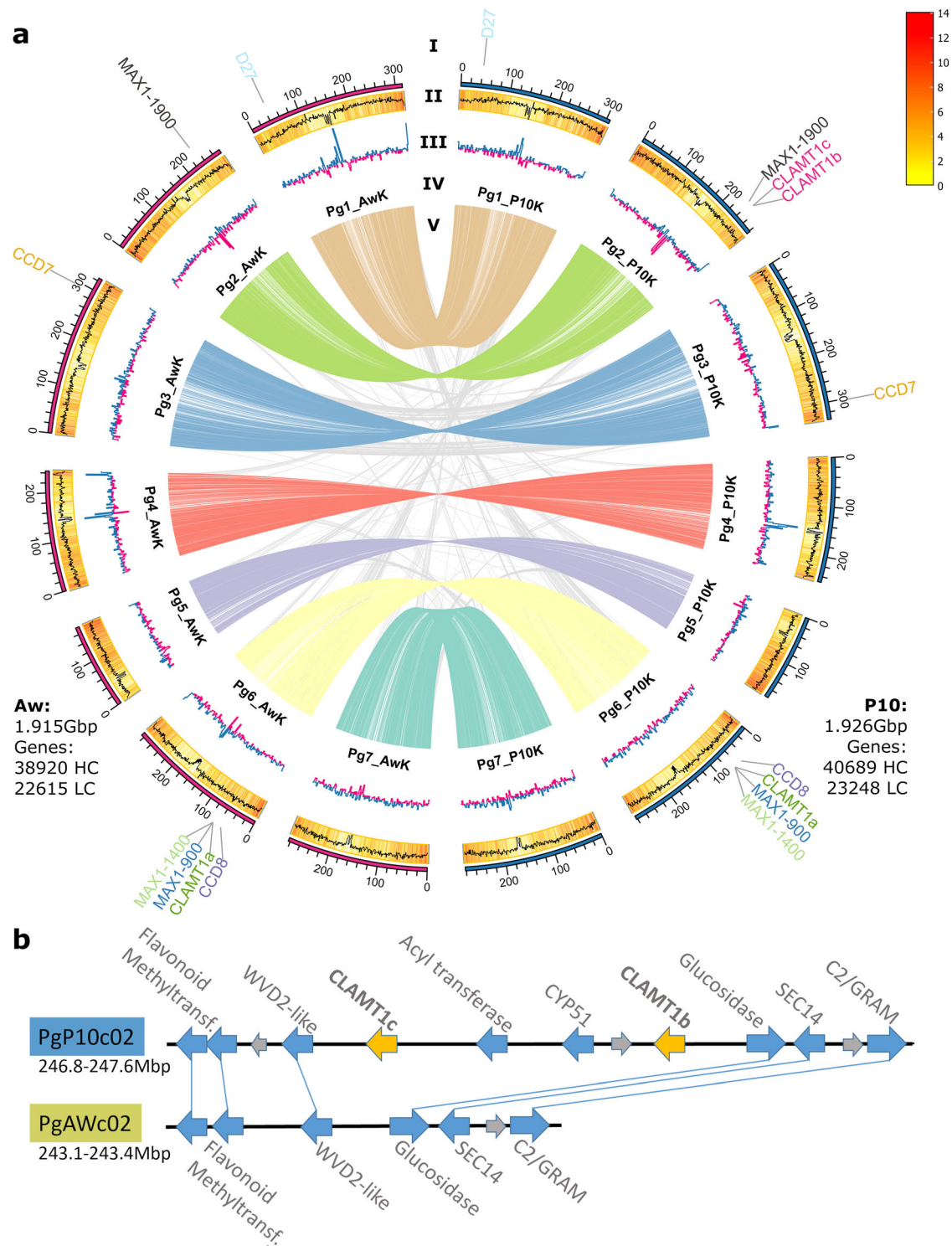


Fig. 2 | Genome assembly for Aw and P10. a Circos plot of the genomes of Aw and P10. Layers from the edge to the center are as follows: I. Location of predicted strigolactone biosynthesis genes, II. GC content (line) and gene density (bars), III. GC skew, IV. Chromosome names, and V. Synteny plot. The insets provide the total assembly size and number of high-confidence (HC) and low-confidence (LC) genes.

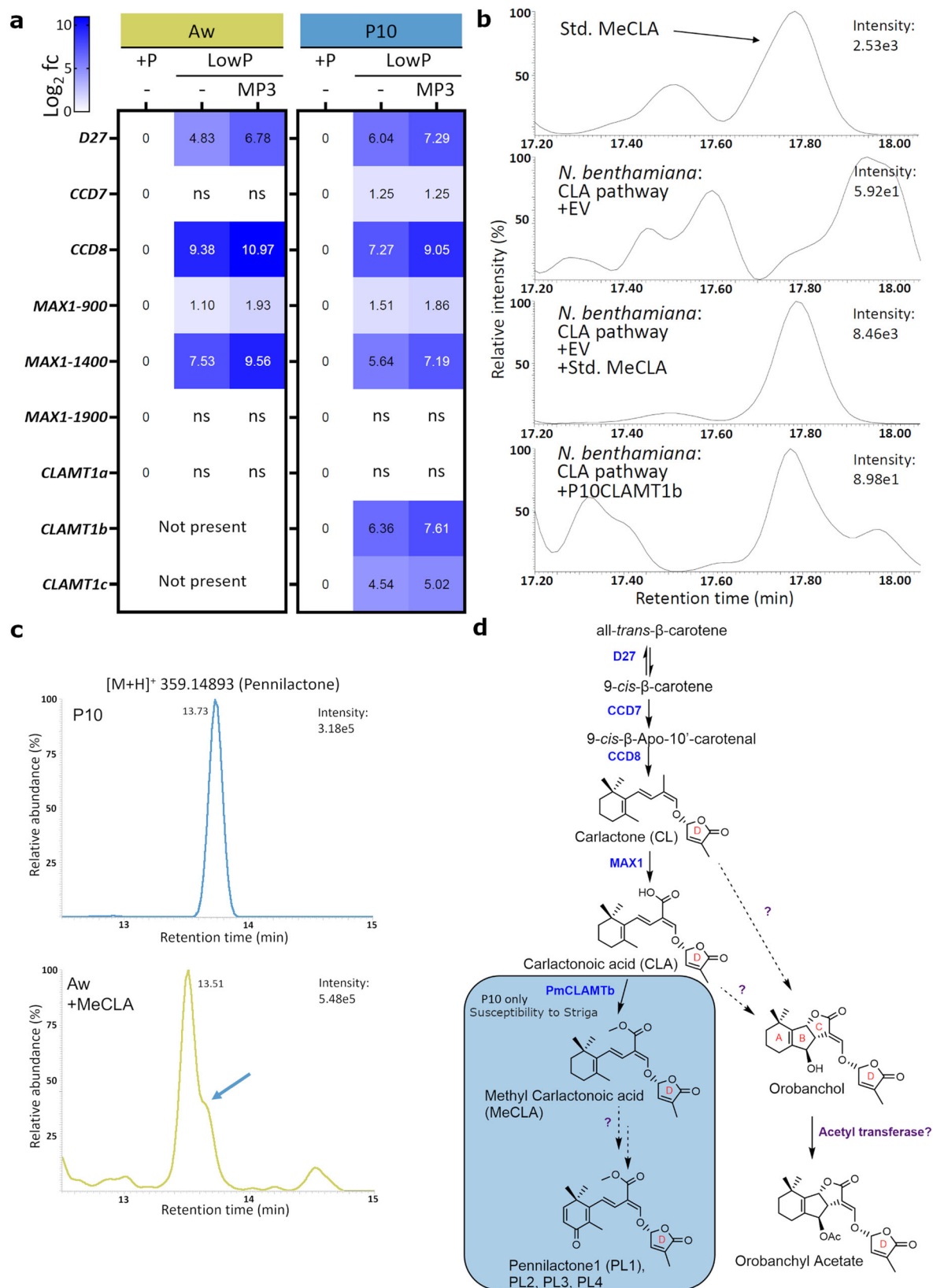
The synteny plot revealed no large-scale rearrangements between P10 and Aw.

b The 0.7 Mb section of chromosome 2, present in P10 but absent in Aw, extended from *CLAMT1c* to *CLAMT1b* and included *CYP51* and an acyl transferase. The flanking regions exhibited strong synteny, as evidenced by three HC genes on each side of the region.

acts as the precursor of PL1 and PL2. The absence of this branch in Aw may account for the higher orobanchol and orobanchyl acetate content, as the metabolic flux is not divided, and there is no competition for the CL or CLA precursor. In addition, enzymes responsible for the biosynthesis of these two SLs may be present in Aw at a higher level, compared to P10.

Presence of *CLAMT1b/c* fragment correlates with the release of pennilactones and *Striga* susceptibility

To determine whether the association between *CLAMT1b* and *Striga* susceptibility was consistent across different pearl millet varieties, we examined a diverse set of previously sequenced pearl millet accessions²³ for the presence of this gene. We discovered that *CLAMT1b*



was consistently accompanied by *CLAMT1c*, *CYP51*, and an *Acyl transferase* in the same *CLAMTb/c* fragment (Supplementary Table 5). We verified the presence of *CLAMTb/c* through genotyping (Supplementary Fig. 16) and identified it in two out of eight accessions. Subsequent SL analysis of the root exudates revealed the presence of PL1, PL3 and PL4, in both accessions containing *CLAMT1b/c* and PL2 in P1537069

only, unlike the exudates from the six accessions lacking the *CLAMT1b/c* region (Fig. 4). This suggests that the biosynthesis of PL3 and PL4 might also rely on *CLAMT*. Consistent with the findings from P10, the root exudates of P1537069 and P1583800 exhibited significantly higher Striga seed-germinating activity than those from the six accessions lacking *CLAMT1b/c* region (Fig. 4). These results demonstrate a clear

Fig. 3 | P10CLAMT1b produced methyl carlactonoate, which is a precursor of pennilactone. **a** The expression of pearl millet homologs of known SL biosynthetic genes was generally upregulated in both Aw and P10 under low Pi conditions and even more so with the addition of the SL analog MP3. *P10CLAMT1b* was strongly co-expressed with *PgD27*, *PgCCDS8*, and *PgMAX1-1400*, while *P10CLAMT1c* was induced more weakly. *PgCLAMT1a* was not significantly induced in Aw or P10 ($n = 4$ biological replicates of three plants each; differential expression analyzed using Fisher's exact test and the p -values were adjusted for multiple testing using the Benjamini–Hochberg method; P values < 0.0003). **b** Recreating the SL biosynthetic pathway up to carlactonoic acid (CLA) through transient expression in *N. benthamiana* leaves produced no peak at the same retention time as the MeCLA

standard. However, when *P10CLAMTb* was added to the experiment, a peak was observed, indicating that P10CLAMTb could convert CLA to MeCLA, as previously shown for OsCLAMT and AtCLAMT. **c** Total ion chromatograms of Aw roots fed with MeCLA produced PL1 (Extracted ion chromatogram: 359.14893; Retention time: 11.7), confirming that MeCLA is a precursor to PL1. The feeding led to two isomers with overlapping peaks, the second of which (blue arrow) corresponds to PL1. **d** The proposed SL biosynthesis pathway in pearl millet, where the boxed section is only present in P10 because of the presence of CLAMT1b. Consequently, only P10 produced downstream SLs, such as PL1, which can induce Striga seed germination when exuded from the roots, making P10 more susceptible to Striga infestation.

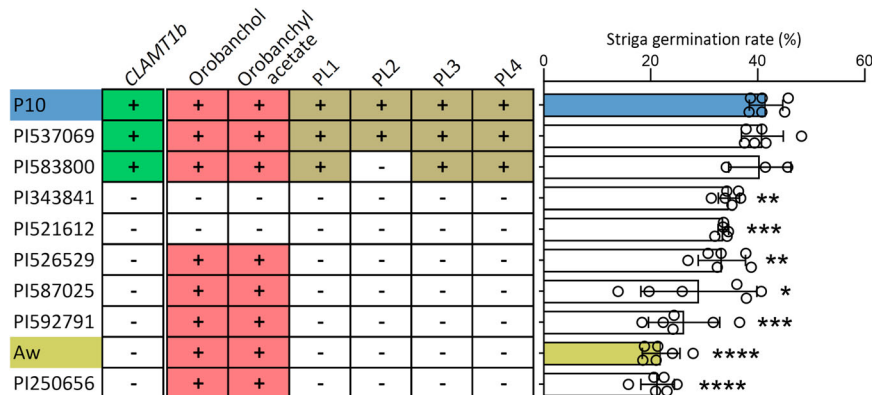


Fig. 4 | Pearl millet cultivars with the CLAMT1b gene exude PLs and induced more Striga seed germination. The root exudates from P10, Aw, and eight other sequenced pearl millet accessions were analyzed for the presence of SLs and their ability to induce Striga seed germination. Only P10, PI537069, and PI583800 contained *CLAMT1b*. The lines are arranged in descending order of Striga germination rates. Error bars represent the mean \pm s.d.; $n = 6$ biological replicates (except for

PI583800, $n = 3$; PI521612, $n = 5$). Significant differences regarding P10 were assessed using a two-tailed t -test: PI537069, $P = 0.74$ (ns); PI583800, $P = 0.70$ (ns); PI343841, $P = 0.0010$; PI521612, $P = 0.0004$; PI526529, $P = 0.0039$; PI587025, $P = 0.022$; PI592791, $P = 0.0005$; Aw, $P < 0.0001$; PI250656, $P < 0.0001$ (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$). Source data are provided as a Source Data file.

link between *CLAMT1b/c*, the SL composition, and pre-attachment Striga susceptibility in this group of accessions.

Number of tillers is affected by SLs and Striga infestation

Recent studies have shown that canonical SLs with a tricyclic lactone (ABC-ring), such as orobanchol, do not significantly contribute to the inhibition of shoot branching/tillering, a function primarily attributed to non-canonical SLs, or even specifically SLs requiring CLAMT for their biosynthesis^{14,16,31–34}. Consistent with these findings, we observed contrasting tillering phenotypes in the two lines: Aw, lacking the four non-canonical SLs (PL1, PL2, PL3, and PL4), produced an average of 10 tillers, while P10 developed only one tiller under normal growth conditions (Fig. 1a; Supplementary Fig. 17). Upon Striga pressure, the number of tillers for Aw is reduced (Supplementary Fig. 17). A possible explanation could be that while Aw is resistant to Striga, it is not immune, and a delayed and less effective attachment of Striga is nonetheless detrimental to host development^{11,35}, resulting in fewer tillers developing.

The ten pearl millet accessions panel did not share a common architecture or shoot phenotype (Supplementary Fig. 18a). We observed an inverse correlation between tiller number and Striga seed germination efficiency, with an R^2 of 0.70 for linear regression and a P -value of 0.0026. Although this correlation is significant, it is not highly predictive (Supplementary Fig. 18b, c). Tillering is a complex trait influenced by multiple factors, including SLs and others^{33,36}, which can mask the impact of the absent *CLAMT1b* gene in several accessions.

The CLAMT1b/c fragment is widely distributed among pearl millet lines

Analysis of resequencing data²¹ revealed that the *CLAMT* fragment was prevalent across a broad spectrum of cultivated pearl millet varieties

and breeding stocks from various regions, as well as in certain wild varieties, indicating that our findings have global relevance (Supplementary Fig. 19). However, the existence of pearl millet accessions lacking the *CLAMT* fragment distributed across the same categories suggests that such accessions are viable and widely utilized. Consequently, the elimination of the *CLAMT* region to impart Striga resistance is unlikely to compromise the viability of an accession.

Colonization by arbuscular mycorrhizal fungi is not reduced in P10

Mycorrhizal colonization, which is influenced by the composition of root exudate SLs, showed higher frequency and intensity values in Aw compared to P10 (Supplementary Fig. 20a, b). This finding was corroborated by molecular analyses of the expression of mycorrhiza-related marker genes (Supplementary Fig. 20c). This shows that removal of the *CLAMT* region does not reduce arbuscular mycorrhizal (AM) colonization and may even improve it, giving additional impetus to the removal. This improvement may be the result of a significantly higher release of orobanchol that was shown to be more efficient in inducing hyphal branching in AM fungi³⁷.

Altogether, we sequenced the genomes of agricultural pearl millet, SOSAT-C88 P10, susceptible to Striga, and 29Aw, a wild and resistant accession. The resulting platinum-grade, near gap-free genome assemblies, are the highest quality for pearl millet to date and can serve as a set of reference genomes for global pearl millet research. Furthermore, we demonstrated that pre-attachment resistance to Striga in pearl millet depends on the types of SLs exuded by the roots. The critical gene for the presence or absence of the newly discovered SLs is *CLAMT1b*, located on a 0.7 Mb fragment of chromosome 2. This fragment is present in susceptible accessions, such as P10, and absent in accessions with higher resistance, such as Aw. This knowledge paves

the way for the enhancement of agriculturally important pearl millet lines through marker-assisted breeding or direct genetic modification to reduce *Striga* susceptibility.

Methods

Genomic DNA extraction and sequencing

Aw and P10 seeds were germinated on half-strength Murashige and Skoog (MS) medium with agar and filter paper moistened with water respectively. The seedlings were grown in a soil/sand mixture under greenhouse conditions for 5 weeks, with the final 48 h in darkness. From each plant, multiple leaf samples exceeding 1.5 g were harvested and immediately flash-frozen in liquid nitrogen. One leaf sample per plant was ground into a fine powder with mortar and pestle while still frozen, whereas a corresponding sample from the same plant was preserved at -80°C .

High molecular weight DNA was extracted following the workflow for HMW DNA extraction for third-generation sequencing³⁸ using the Genomics-Tip 100 G kit (Qiagen), with the following modifications: For pearl millet, we used 1.5 to 2 g of ground tissue exceeding the standard recommendation of 1 g per four 100/G columns. Despite the protocol's specific warnings against shaking the 50 mL tubes during the lysis step, we found it necessary to occasionally disrupt tissue clots formation by shaking gently. The elution step extended to up to 4 h instead of the 1 h duration specified in the protocol.

Quality control of HMW DNA was performed using a FEMTO Pulse (Agilent) with the gDNA 165 kb kit (FP-1002-0275) and a separation time of 70 min. Both Aw and P10 samples contained a high amount of genomic DNA with a length above the cut-off of 50 kb. Quantification was conducted with a Qubit assay using the Qubit dsDNA BR assay kit (Thermo Fisher Scientific) yielding 288 ng/ μL for Aw and 922 ng/ μL for P10.

The HMW DNA was sheared to 20 kb and processed for PacBio HiFi sequencing by the KAUST Bioscience Core Lab. Samples were sequenced on a PacBio Sequel II using three SMRT cells each. The total throughput exceeded 30 Gb per SMRT cell for P10 and over 35 Gb per SMRT cell for Aw. The median read length was 15 kb per cell for P10 and 16 kb per cell for Aw.

Omni-CTM library construction and sequencing

The Omni-CTM library was prepared using the Dovetail[®] Omni-CTM Kit for plant tissues according to the manufacturer's protocol. Chromatin was fixed with disuccinimidyl glutarate (DSG) and formaldehyde in the nucleus from dark-treated young leaves. The cross-linked chromatin was digested in situ with DNase I. After digestion, the cells were lysed with SDS to extract chromatin fragments, which were then bound to Chromatin Capture Beads. The chromatin ends were repaired and ligated to a biotinylated bridge adapter followed by proximity ligation of adapter-containing ends. After proximity ligation, the crosslinks were reversed, the associated proteins were degraded, and the DNA was purified. The purified DNA was then converted into a sequencing library using Illumina-compatible adapters. Biotin-containing fragments were isolated with streptavidin beads prior to PCR amplification. The two libraries were sequenced on an Illumina MiSeq platform to generate >214 and >269 millions 2×150 bp read pairs for Aw and P10, respectively.

Genome assembly

PacBio HiFi reads were assembled using hifiasm³⁹ (v17.6) with default parameters (<https://github.com/chhylp123/hifiasm/>) to generate primary contig assemblies. Subsequently, construction of the pseudomolecules was performed by integration of Omni-CTM read data using Juicer⁴⁰ (v2; <https://github.com/aidenlab/juicer>) and the 3D-DNA pipeline⁴¹ (<https://github.com/aidenlab/3d-dna>). First, to generate the Hi-C contact maps for P10 and Aw genomes, Omni-CTM Illumina short reads were processed with juicer.sh (parameter: -s none --assembly).

The resulting output file "merged_nodups.txt" and the primary assembly were then used to produce an assembly with 3D-DNA3 (using run-asm-pipeline.sh with the -r 0 parameter). Juicebox⁴² (v2.14.00) was employed to visualize the Hi-C contact matrix alongside the assembly and to manually curate the assembly. The orientation and order of each pseudomolecule were defined by dot-plot comparison using chromosome⁴³ (<https://github.com/estebanpw/chromosome>) against the pearl millet genotype Tift 23D2B1-P1-P56. All the remaining contigs not anchored to the pseudomolecules were concatenated into "unanchored chromosomes". The final Hi-C contact maps and assemblies were saved using run-asm-pipeline-post-review.sh from the 3D-DNA pipeline.

RNA extraction and sequencing

Seeds of Aw and P10 descended from the sequenced individuals, were germinated as described above. The seedlings were transferred to 50 mL hydroponics tubes and grown in Hoagland solution modified in the following ways: no modifications, low phosphate (1% of normal P), low phosphate with MP3 (1.0 μM), and low phosphate with only acetone mock treatment. Treatment with MP3 and acetone mock were done only for the 6 h before harvesting. The roots of the +P and lowP plants as well as 3 day old seedlings and a flowering inflorescence were sampled for Iso-Seq.

Roots and shoot stubs were collected separately from each hydroponic growth treatment for both Aw and P10, with samples pooled from three plants at each collection and four such biological replicates obtained.

All samples were flash-frozen in liquid nitrogen and ground to a fine powder in a sterilized mortar and pestle. Then 100 mg of the samples was separated for RNA extraction. RNA extraction was performed using an RSC 48 RNA extraction robot (Maxwell) and the Maxwell RSC Plant RNA kit (Promega). Although the RNA yield for some of the 64 samples was low, ranging from 36 ng/ μL to 260 ng/ μL as measured by NanoDrop, the purity was consistently high. The RNA integrity number (RIN) scores ranged from 8 to 10 for over 92% of the samples with an average RIN of 9.0.

The extracted RNA was submitted to the KAUST Bioscience Core Lab for Iso-Seq sequencing. Samples from five different tissues for both Aw and P10 were tagged and multiplexed onto a SMRT cell each and sequencing on the PacBio Sequel II platform.

Samples for RNA-Seq were sent to Novogene (Singapore) for mRNA library preparation and 150 bp pair-end sequencing on Illumina's NovaSeq 6000 platform, targeting a throughput of 12 Gb of data per sample. The returned data were consistently high quality, with the percentage of reads scoring a Phred value over 30 (indicating a base error below 0.1%) exceeded 90% for each sample.

Transposable element identification and quantification

Transposable elements (TE) were identified searching the genome assemblies with the Extensive *denovo* TE Annotator pipeline EDTA⁴⁴ (version 2.0) run under default settings. Due to the high incidence of false positives in the prediction of helitrons, representatives of this class of TEs were removed from the final EDTA output. The TE library was then employed to mask the two genome assemblies and quantify the TE content using the tool RepeatMasker (<http://www.repeatmasker.org/>) ran under the default parameters (with the exception of the -qq option).

Genome annotation

We annotated the two genomes using the MAKER pipeline⁴⁵ v3.01.03. For a detailed breakdown of the genomes annotation, please refer to the supplementary materials (Supplementary Fig. 21) or the project's GitHub page (<https://github.com/mjfi2sb3/millet-genome-annotation>). First, we prepared the necessary transcriptomic and homology data to inform and support the prediction in the MAKER workflow. We began by preprocessing the Iso-Seq data following PacBio's recommended

workflow using SMRT tools v11.0 (<https://www.pacb.com/support/software-downloads/>). The product of this step was a set of high-quality full-length isoforms for each submitted sample. Details regarding the preprocessing of RNA-seq data can be found in another section of this manuscript. For homology evidence, we incorporated the manually curated UniProt Swiss-Prot database⁴⁶ (downloaded in Nov 2022), along with published protein annotation for *Cenchrus americanus*²¹ and *Cenchrus purpureus*⁴⁷ (elephant grass). Additionally, we included NCBI annotations for *Setaria viridis* (green millet), *Setaria italica* (foxtail millet) and *Sorghum bicolor* (sorghum).

Subsequently, we processed the repeat-masked genome assemblies through the MAKER pipeline. The workflow calls an array of tools, including NCBI BLAST tools⁴⁸ v2.2.28 +, Exonerate⁴⁹ v2.2.0, Augustus⁵⁰ v3.2.3 and tRNAscan-SE⁵¹ v2.0. We ran MAKER's workflow primarily with the default values except for `alt_splice=1` and `always_complete=1`. Iso-Seq and RNA-seq transcripts were aligned to their respective assemblies using NCBI blastn, and these alignments were subsequently refined using Exonerate.

The protein evidence was aligned and refined using NCBI blastp and Exonerate, respectively. Subsequently gene structure prediction was performed using Augustus with the species parameter set to *Zea mays*. EST and protein hints were created using alignments obtained in the previous step. MAKER was then used to assess the predicted genes, correct some of the predictions, add isoform information and calculate quality scores (AED scores).

In the final step, we divided the predicted gene models into High Confidence (HC) and Low Confidence (LC) categories using four strategies: 1) based on EST evidence (MAKER's quality index scores as well as alignment-based filtering); 2) annotating with the KEGG database⁵²; 3) annotating using InterProScan⁵³ v5; and 4) annotating against the UniProt Swiss-Prot database.

RNA-Seq data mapping onto P10 and Aw genomes

To map the RNA-Seq reads from each experiment using Spliced Transcripts Alignment to a Reference (STAR) software⁵⁴, we first created an index of each of the P10 and Aw genome assemblies. For each RNA-Seq sample, the paired-end fastq data were then mapped on to the corresponding genome assembly using STAR with the option “-outSAMstrandField intronMotif” option. Subsequently, we assembled the transcripts for each RNA-Seq sample with StringTie⁵⁵ using the BAM files generated in the previous alignment step. For each genome, we merged all transcripts from individual experiments using the StringTie merge option to produce a non-redundant set of transcripts.

Determining differential expression

In all two-way comparisons we used the R package edgeR⁵⁶ for differential expression analysis, with the default settings. We first filtered out genes having more than two replicates out of the total eight with a count per million (cpm) ≤ 0.5 . We performed the differential expression analysis using Fisher's exact test and the *p*-values were adjusted for multiple testing using the Benjamini-Hochberg method.

Gene identification and phylogeny

Homologs of known strigolactone biosynthetic pathway enzymes were identified through tblastn searches on the Aw and P10 genome assemblies using Persephone (persephonesoft.com). Phylogenetic trees of the protein families were constructed in Geneious v2023.2.1 (Biomatters) using muscle v5.1 based on the PPP alignment algorithm. The consensus trees were constructed using the neighbor-joining method, relying on the Jukes-Cantor model and was supported by 1000 bootstrap replicates.

Screening diverse pearl millet accessions collection

We acquired a panel of 10 sequenced pearl millet accessions²³ from the U.S. National Plant Germplasm System. We screened the genomes of

these accessions for the presence of the CLAMT region genes and the four flanking genes using tblastn, noting both presence and sequence similarity at the protein level (Supplementary Table 5). Seedlings of the panel were genotyped through PCR using the Phire Plant Direct kit (Thermo Fisher Scientific) directly on leaf extracts. For two accessions, P1527388 and P1186338, we detected the presence of the CLAMT fragment through genotyping, despite its absence in their genome assemblies. Consequently, we excluded these two lines from subsequent experiments.

Strigolactone collection, extraction, and measurements

Pearl millet seedlings were cultivated under controlled conditions with a day/night temperature of 28/22 °C. The seeds were surface-sterilized in a 50% sodium hypochlorite solution for 10 min and rinsed with sterile water. They were then placed in magenta boxes containing half-strength MS medium and allowed to germinate in darkness for 24 h. Following this period, they were incubated in a Percival chamber for 4 days. The germinated seedlings were then transferred into the soil for phenotyping, to sand for SL detection, or to a hydroponic system for *Striga* bioassays.

Analysis of SLs in root exudates was conducted using a previously published protocol⁵⁷. In summary, 1 L of root exudates, spiked with 20 ng GR24, was collected and applied to a C18-Fast Reversed-Phase SPE column (500 mg/3 mL; GracePure™) pre-conditioned with 3 mL of methanol and 3 mL of water. The column was then washed with 3 mL of water, and SLs were eluted with 5 mL of acetone. The SL fraction was concentrated to approximately 1 mL in an aqueous SL solution and subsequently extracted using 1 mL of ethyl acetate. Then, 750 μ L of the SL-enriched organic phase was dried under a vacuum. The residue was reconstituted in 100 μ L of acetonitrile:water (25:75, v/v) and filtered through a 0.22 μ m filter for LC-MS/MS analysis.

For SL extraction from *N. benthamiana* leaf, samples were ground to a powder in liquid nitrogen using a mortar and pestle. About 300 mg of powder was weighed out and transferred to an 8 mL brown glass vial to which cold 2 mL ethyl acetate were added. After vortexing, sonication and centrifugation at 3300 g, the supernatant was transferred into an 8 mL glass vial. The pellet was extracted once more and the supernatants combined and dried in a SpeedVac. After drying, the residue was dissolved in 50 μ L of ethyl acetate and 2 mL hexane. Further purification was performed using the Silica gel SPE column (500 mg/3 mL) preconditioned with 3 mL of ethyl acetate and 3 mL of hexane. After washing with 3 mL hexane, SLs were eluted in 3 mL ethyl acetate and evaporated to dryness under vacuum.

SL identification was performed using a UHPLC-Orbitrap ID-X Tribid Mass Spectrometer (Thermo Fisher Scientific) equipped with a heated electrospray ionization source. Chromatographic separation was achieved using Hypersil GOLD C18 Selectivity HPLC Columns (150 \times 4.6 mm; 3 μ m; Thermo Fisher Scientific). The mobile phase comprised water (A) and acetonitrile (B), each containing 0.1% formic acid. A linear gradient was applied as follows (flow rate, 0.5 mL/min): 0–15 min, 25–100% B, followed by washing with 100% B, and a 3-min equilibration with 25% B. The injection volume was 10 μ L, and the column temperature was consistently maintained at 35 °C. The MS conditions included: positive mode; spray voltage of 3500 V; sheath gas flow rate of 60 arbitrary units; auxiliary gas flow rate of 15 arbitrary units; sweep gas flow rate of 2 arbitrary units; ion transfer tube temperature of 350 °C; vaporizer temperature of 400 °C; S-lens RF level of 60; resolution of 120000 for MS; stepped HCD collision energies of 10, 20, 30, 40, and 50%; and a resolution of 30000 for MS/MS. The mass accuracy of identified compounds, with a mass tolerance of ± 5 ppm, is presented in Supplementary Table 6. All data were acquired using Xcalibur software version 4.1 (Thermo Fisher Scientific).

SLs were quantified using LC-MS/MS with a UHPLC-Triple-Stage Quadrupole Mass Spectrometer (Thermo Fisher Scientific Altis™). Chromatographic separation was achieved on a Hypersil GOLD C18

Selectivity HPLC Column (150 mm × 4.6 mm; 3 μm; Thermo Fisher Scientific), utilizing a mobile phase comprising water (A) and acetonitrile (B), each with 0.1% formic acid. The linear gradient was as follows (flow rate, 0.5 mL/min): 0–15 min, 25–100% B, followed by washing with 100% B, and a 3-min equilibration with 25% B. The injection volume was 10 μL, and the column temperature was consistently maintained at 35 °C. The MS parameters included: positive ion mode; H-ESI ion source; ion spray voltage of 5000 V; sheath gas flow rate of 40 arbitrary units; aux gas flow rate of 15 arbitrary units; sweep gas flow rate of 20 arbitrary units; ion transfer tube gas temperature of 350 °C; vaporizer temperature of 350 °C; collision energy of 17 eV; CID gas at 2 mTorr; and a Q1/Q3 mass with a full-width half maximum (FWHM) value of 0.4 Da. The characteristic Multiple Reaction Monitoring (MRM) transitions (precursor ion → product ion) were 347.14 → 97.02, 347.14 → 233.1, 347.14 → 205.1 for Oro; 389.15 → 97.02, 411.1 → 97.02, 389.15 → 233.1 for Oro Ace; 347.18 → 97.02, 347.18 → 287.1, 347.18 → 315.1, 347.18 → 329.14 for MeCLA; 299.09 → 185.06, 299.09 → 157.06, 299.09 → 97.02 for GR24; 359.14 → 97.02, 359.14 → 345.1, 359.14 → 299.1 for PL1; 377.15 → 97.02, 377.15 → 359.1, 377.15 → 249.1 for PL2; 375.14 → 97.02, 375.14 → 343.1, 375.14 → 247.1 for PL3; 452.19 → 97.02, 452.19 → 375.1, 452.19 → 315.1 for PL4.

SL collection and fractioning

Analysis of SLs in root exudates followed the protocol by Wang et al.⁵⁷. In summary, 1 L of collected root exudates was extracted using a C₁₈-Fast Reversed-Phase SPE column (500 mg/3 mL; GracePure™), which had been pre-conditioned with 3 mL of methanol and 3 mL of water. The column was then washed with 3 mL of water, and SLs were eluted with 5 mL of acetone. The SL fraction was concentrated to approximately 1 mL of aqueous solution and then extracted with 1 mL of ethyl acetate. 750 μL of SL enriched organic phase was dried under vacuum. Concentrated SL extracts of root exudates obtained from 12 replicates (-12 L) were dissolved in 1.5 mL EtOAc/ 2 mL Hexane and subjected to silica gel column chromatography (SPE column 60 g /50 mL) with a stepwise elution of Hexane/EtOAc (100:0–0:100, 10% step, 3 mL in each step) to yield 11 fractions (A–K). 1 mL of each fraction was subjected to LC-MS analysis for monitoring the potential SLs and verify the *Striga* bioassay.

Striga germination bioassays

The *Striga* germination bioassays were conducted following a previously described procedure⁵⁸. In summary, *Striga* seeds were surface-sterilized with 50% diluted commercial bleach for 5 min. Then, they were dried and uniformly spread (approximately 50–100 seeds) on 9 mm filter paper discs made of glass fiber. Subsequently, 12 seed-laden discs were placed in a 9 cm Petri dish containing a Whatman filter paper moistened with 3.0 mL of sterilized Milli-Q water. The dishes were sealed with parafilm and incubated at 30 °C for 10 days for pre-conditioning. Post-conditioning, the *Striga* seeds were treated with SLs from root exudates of various pearl millet lines and incubated again at 30 °C for 24 h. Then, germinated and total seeds were scanned and counted using SeedQuant⁵⁹, and the percentage of germination was calculated.

Striga emergence under greenhouse pot conditions

The millet lines underwent *Striga* infection testing in pots within a greenhouse setting. Approximately 2.0 L of blank soil, a mixture of sand and Stender soil, Basissubstrat, in a 1:3 ratio was placed at the base of an 8.0 L perforated plastic pot. Subsequently, approximately 40,000 *Striga* seeds, equating to roughly 100 mg, were evenly distributed within a 5.0 L soil mixture and layered atop the blank soil in the pot. The *Striga* seeds within each pot underwent a pre-conditioning period of 10 days at 30 °C with light irrigation maintained under greenhouse conditions. Following this, a single 10-day-old seedling was planted centrally in each pot. The millet plants were

cultivated under standard growth conditions, with a temperature of 30 °C and 65% RH. *Striga* emergence was monitored and recorded for each pot at 70 days post-millet sowing.

Transient expression in *Nicotiana benthamiana* leaf

The correct transcripts for the three *CLAMT* genes in pearl millet were identified from the annotation for *CLAMT1a* and *CLAMT1b*, while the ambiguity in *CLAMT1c* was addressed by picking the most likely predicted transcript (*CLAMT1c-Iso1*) and generating an alternative by FGENESH+ (Berrysoft) (*CLAMT1c-Iso2*). To generate pearl millet *CLAMT* plasmids for transient expression in *Nicotiana benthamiana*, the full-length cDNA of *CLAMT1b*, *CLAMT1a*, *CLAMT1c-iso1*, *CLAMT1c-Iso2* (Supplementary Table 7) were amplified by Phusion polymerase (New England Biolabs) from cDNA (*CLAMT1b*) or synthesized fragments (*CLAMT1a* and *CLAMT1c*; Azenta Life Sciences) using primers indicated in Supplementary Table 8. The PCR products were purified and sequenced. Following Sanger sequencing, the gene sequences were amplified by using primers with suitable restriction enzyme sites. The resulting fragments were digested and ligated into the linearized entry vector pVIA_2.1 which includes the CaMV35S promoter (www.pri.wur.nl/UK/products/ImpactVector/).

After sequence confirmation of the pVIA_2.1 entry clones, Gateway LR clonase II enzyme mix (Invitrogen) reactions were performed to transfer the fragments into the pBinPlus binary vector⁶⁰, generating p35S:PBIN-*CLAMT1b*, p35S:PBIN-*CLAMT1a*, p35S:PBIN-*CLAMT1c-iso1* and p35S:PBIN-*CLAMT1c-iso2*. Additionally, we cloned the Arabidopsis *Atmax1* and *Atclamt* cDNAs in the same binary vector; pBinPlus for transient expression in *N. benthamiana*.

The binary vector harboring various genes was introduced into *Agrobacterium tumefaciens* strain AGL0 via electroporation. Positive clones were cultured at 28 °C at 220 rpm for 2 days in LB medium supplemented with 50 mg/L Kanamycin and 35 mg/L Rifampicin. Cells were collected by centrifugation for 15 min at 3300 g and room temperature. They were then resuspended in 10 mM MES-KOH buffer (pH 5.7) with 10 mM MgCl₂ and 100 mM acetosyringone (49-hydroxy-3',5'-dimethoxyacetophenone; Sigma) to achieve a final OD₆₀₀ of 0.5. The suspension was incubated with gentle rolling at 22 °C for 2–4 h. For various gene combinations, equal concentrations of *Agrobacterium* strains carrying different constructs were mixed, using strains with empty vectors to compensate for gene dosage in each combination. Additionally, an *Agrobacterium* strain containing a gene for the TBSV P19 protein was included to enhance protein production by inhibiting gene silencing. *N. benthamiana* plants were cultivated in soil pots in a greenhouse under a 14 h light/10 h dark cycle at 25 °C and 22 °C, respectively. Combinations of constructs in *Agrobacterium* were infiltrated into the leaves of 5-week-old *N. benthamiana* plants using a 1-mL syringe. Leaves at the same developmental stage were selected to reduce variability. For each gene combination, two to three leaves per plant were infiltrated, with three plants serving as individual biological replicates. The bacterial suspension was gently injected into the abaxial side of the leaf to ensure distribution throughout the entire leaf area. Six days post-infiltration, the leaves were collected for subsequent analysis.

Analysis of resequencing data

We downloaded 1036 sequence read archives (SRA) from NCBI SRA study SRP063925 and converted them to fastq files using sratools⁶¹ v3.0.7. We mapped the paired-end reads to our reference P10 assembly using the bwa-mem2 v2.2.1 mem subcommand with default parameters⁶². We then extracted read mappings that fall within the region of interest (ROI). We estimated the mean coverage for the ROI and that of the harboring chromosome using the samtools⁶³ v1.16.1 subcommand coverage with a minimum MAPQ score of 15. We calculated the coverage ratio (mean coverage of chromosome / mean coverage of ROI) as a proxy for the presence or absence of

the ROI and plotted these ratios using Python. A detailed breakdown of the command workflow (Supplementary Fig. 22) is available on our GitHub page (<https://github.com/mjfi2sb3/millet-genome-annotation>).

Mycorrhizal colonization of P10 and Aw

P10 and Aw were cultivated in sand and inoculated with approximately 1,000 sterile spores of *Rhizophagus irregularis* (DAOM 197198, Agromnutrition, Labège, France). They received watering twice weekly, alternating between with tap water and a modified Long-Ashton (LA) solution containing 3.2 μM $\text{Na}_2\text{HPO}_4 \cdot 12 \text{H}_2\text{O}$.

All the plants were sampled at 45-days post-inoculation (dpi), corresponding to the late stage of mycorrhization. To evaluate the level of mycorrhization, we performed a morphological analysis according to Trouvelot et al.⁶⁴. Moreover, we conducted qRT-PCR assays to assess the expression level of the fungal housekeeping gene (RiEF) and a fungal gene preferentially expressed in the intraradical structures (RiPEIP1)⁶⁵, and two plant AM marker phosphate transporter genes (PtHL9; EcPt4)^{66,67}. We used alpha-tubulin (TUA)⁶⁸ as the reference plant gene.

Statistical analysis

Data are represented as mean and their variations as SD. The statistical significance was determined by the two-tailed unpaired Student's *t* test or one-way ANOVA and Tukey's multiple comparison test, using a probability level of $P < 0.05$. All statistical elaborations were performed using Prism 9 (GraphPad).

Reporting summary

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

Data availability

The Genome assemblies are available on the European Nucleotide Archive (ENA) under bioproject [PRJEB71762](https://www.ebi.ac.uk/ena/record/PRJEB71762). The same ENA bioproject also contains the genomic HiFi data, Omni-C™ data, RNA-Seq, Iso-Seq data, the assembled RNA-seq transcripts, and full-length Iso-Seq transcripts. We have also hosted the above data on the DRYAD digital repository [<https://doi.org/10.5061/dryad.nk98sf80k>] along with the gene model predictions, repeat annotations and other useful information. The Aw and P10 genomes were also uploaded to our platform at <https://millet.kaust.edu.sa/> and to the independently hosted platform at <https://web.persephonesoft.com/?data=genomes/P10> and https://web.persephonesoft.com/?data=genomes/Pg_Aw. With data tracks for gene models, RNA-seq and Iso-Seq as well as BLAST search and synteny analysis functions. The public resequencing data, SRA study SRP063925, were downloaded from NCBI. Source data are provided with this paper.

References

- Al-Babili, S. & Bouwmeester, H. J. Strigolactones, a novel carotenoid-derived plant hormone. *Annu. Rev. Plant Biol.* **66**, 161–186 (2015).
- Fiorilli, V., Wang, J. Y., Bonfante, P. & Lanfranco, L. Apocarotenoids: old and new mediators of the arbuscular mycorrhizal symbiosis. *Front. Plant Sci.* **10**, 484961 (2019).
- Wang, J. Y., Fiorilli, V., Lanfranco, L. & Asami, T. Editorial: Specialized metabolites manipulating organismal behaviors and rhizospheric communications. *Front. Plant Sci.* **14**, 1197058 (2023).
- Yoneyama, K., Awad, A. A., Xie, X., Yoneyama, K. & Takeuchi, Y. Strigolactones as germination stimulants for root parasitic plants. *Plant Cell Physiol.* **51**, 1095–1103 (2010).
- Nomura, S. et al. Structural requirements of strigolactones for germination induction and inhibition of *Striga gesnerioides* seeds. *Plant Cell Rep.* **32**, 829–838 (2013).
- Parker, C. Parasitic weeds: a world challenge. *Weed Sci.* **60**, 269–276 (2012).
- Satyavathi, C. T., Ambawat, S., Khandelwal, V. & Srivastava, R. K. Pearl millet: a climate-resilient nutricereal for mitigating hidden hunger and provide nutritional security. *Front. Plant Sci.* **12**, 659938 (2021).
- Runo, S. & Kuria, E. K. Habits of a highly successful cereal killer, *Striga*. *PLOS Pathog.* **14**, e1006731 (2018).
- Pennisi, E. Armed and dangerous. *Science* <https://doi.org/10.1126/science.327.5967.804> (2010).
- Fishman, M. R. & Shirasu, K. How to resist parasitic plants: pre- and post-attachment strategies. *Curr. Opin. Plant Biol.* **62**, 102004 (2021).
- Dayou, O. et al. Two-tier witchweed (*Striga hermonthica*) resistance in wild pearl millet (*Pennisetum glaucum*) 29Aw. *Weed Sci.* **69**, 300–306 (2021).
- Kavuluko, J. et al. GWAS provides biological insights into mechanisms of the parasitic plant (*Striga*) resistance in sorghum. *BMC Plant Biol.* **21**, 392 (2021).
- Gobena, D. et al. Mutation in sorghum LOW GERMINATION STIMULANT 1 alters strigolactones and causes *Striga* resistance. *Proc. Natl Acad. Sci.* **114**, 4471–4476 (2017).
- Ito, S. et al. Canonical strigolactones are not the major determinant of tillering but important rhizospheric signals in rice. *Sci. Adv.* <https://doi.org/add1278> (2022).
- Li, C. et al. Maize resistance to witchweed through changes in strigolactone biosynthesis. *Science* <https://doi.org/abq4775> (2023).
- Chen, G. et al. Disruption of the rice 4-deoxyorobanchol hydroxylase unravels specific functions of canonical strigolactones. *Proc. Natl Acad. Sci.* **120**, e2306263120 (2023).
- Wang, J. Y., Chen, G. E., Braguy, J., & Al-Babili, S. Distinguishing the functions of canonical strigolactones as rhizospheric signals. *Trends Plant Sci.* <https://doi.org/10.1016/j.tplants.2024.02.013> (2024).
- Omanya, G. O. et al. Participatory varietal selection with improved pearl millet in West Africa. *Exp. Agric.* **43**, 5–19 (2007).
- Wang, J. Y. et al. The apocarotenoid metabolite zaxinone regulates growth and strigolactone biosynthesis in rice. *Nat. Commun.* **10**, 1–9 (2019).
- Wang, J. Y. et al. Efficient mimics for elucidating zaxinone biology and promoting agricultural applications. *Mol. Plant* **13**, 1654–1661 (2020).
- Varshney, R. K. et al. Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* **35**, 969–976 (2017).
- Ramu, P. et al. Improved pearl millet genomes representing the global heterotic pool offer a framework for molecular breeding applications. *Commun. Biol.* **6**, 1–11 (2023).
- Yan, H. et al. Pangenomic analysis identifies structural variation associated with heat tolerance in pearl millet. *Nat. Genet.* **55**, 507–518 (2023).
- Salson, M. et al. An improved assembly of the pearl millet reference genome using Oxford Nanopore long reads and optical mapping. *G3 Genes/Genomes/Genet.* **13**, jkad051 (2023).
- Mallu, T. S. et al. Mechanisms of pre-attachment *Striga* resistance in sorghum through genome-wide association studies. *Mol. Genet. Genomics* **297**, 751–762 (2022).
- Badu-Apraku, B. et al. Identification of QTLs for grain yield and other traits in tropical maize under *Striga* infestation. *PLoS ONE* **15**, e0239205 (2020).
- Rouamba, A. et al. Genome-wide association analyses of agronomic traits and *Striga hermonthica* resistance in pearl millet. *Sci. Rep.* **13**, 1–12 (2023).
- Haider, I. et al. Transcriptome analysis of the phosphate starvation response sheds light on strigolactone biosynthesis in rice. *Plant J.* **114**, 355–370 (2023).

29. Mashiguchi, K. et al. A carlactonoic acid methyltransferase that contributes to the inhibition of shoot branching in Arabidopsis. *Proc. Natl Acad. Sci.* **119**, e2111565119 (2022).
30. Zhang, Y. et al. Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis. *Nat. Chem. Biol.* **10**, 1028–1033 (2014).
31. Wakabayashi, T. et al. Direct conversion of carlactonoic acid to orobanchol by cytochrome P450 CYP722C in strigolactone biosynthesis. *Sci. Adv.* **5**, eaax9067 (2019).
32. Cui, J. et al. Fertilization controls tiller numbers via transcriptional regulation of a MAX1-like gene in rice cultivation. *Nat. Commun.* **14**, 1–13 (2023).
33. Dun, E. A., Brewer, P. B., Gillam, E. M. J. & Beveridge, C. A. Strigolactones and shoot branching: what is the real hormone and how does it work? *Plant Cell Physiol.* **64**, 967–983 (2023).
34. Wang, J. Y. et al. Disruption of the cytochrome CYP711A5 gene reveals MAX1 redundancy in rice strigolactone biosynthesis. *J. Plant Physiol.* **287**, 154057 (2023).
35. Van Ast, A., Bastiaans, L. & Kropff, M. J. A comparative study on *Striga hermonthica* interaction with a sensitive and a tolerant sorghum cultivar. *Weed Res.* **40**, 479–493 (2000).
36. McSteen, P. Hormonal regulation of branching in grasses. *Plant Physiol.* **149**, 46–55 (2009).
37. Mori, N., Nishiuma, K., Sugiyama, T., Hayashi, H. & Akiyama, K. Carlactone-type strigolactones and their synthetic analogues as inducers of hyphal branching in arbuscular mycorrhizal fungi. *Phytochemistry* **130**, 90–98 (2016).
38. Driguez, P., Carty, K., Putra, A., & Ermini, L. Workflow for generating HMW plant DNA for third generation sequencing with high N50 and high accuracy. protocols.io. <https://doi.org/10.17504/protocols.io.bafmibk6> (2021).
39. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
40. Durand, N. C. et al. JuiceR provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
41. Dudchenko, O. et al. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* <https://doi.org/aal3327> (2017).
42. Durand, N. C. et al. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* **3**, 99–101 (2016).
43. Pérez-Wohlfeil, E., Diaz-del-Pino, S. & Trelles, O. Ultra-fast genome comparison for large-scale genomic experiments. *Sci. Rep.* **9**, 1–10 (2019).
44. Ou, S. et al. Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **20**, 275 (2019).
45. Holt, C. & Yandell, M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinform.* **12**, 491 (2011).
46. Bairoch, A. & Apweiler, R. The SWISS-PROT protein sequence database: its relevance to human molecular medical research. *J. Mol. Med. (Berl.)* **75**, 312–316 (1997).
47. Yan, Q. et al. The elephant grass (*Cenchrus purpureus*) genome provides insights into anthocyanidin accumulation and fast growth. *Mol. Ecol. Resour.* **21**, 526–542 (2021).
48. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
49. Slater, G. S. C. & Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinform.* **6**, 31 (2005).
50. Stanke, M. & Waack, S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**, ii215–ii225 (2003).
51. Lowe, T. M. & Eddy, S. R. TRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
52. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (1999).
53. Jones, P. et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).
54. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2012).
55. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
56. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. EdgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2009).
57. Wang, J. Y. et al. Protocol for characterizing strigolactones released by plant roots. *STAR Protoc.* **3**, 101352 (2022).
58. Jamil, M. et al. Cytokinins as an alternative suicidal *Striga* germination compound. *Weed Res.* <https://doi.org/10.1111/wre.12585> (2023).
59. Braguy, J. et al. SeedQuant: a deep learning-based tool for assessing stimulant and inhibitor activity on root parasitic seeds. *Plant Physiol.* **186**, 1632–1644 (2021).
60. van Engelen, F. A. et al. pBINPLUS: an improved plant transformation vector based on pBIN19. *Transgenic Res.* **4**, 288–290 (1995).
61. Sayers, E. W., O’Sullivan, C. & Karsch-Mizrachi, I. In *Plant Bioinformatics. Methods in Molecular Biology* (ed. Edwards, D.) **2423** https://doi.org/10.1007/978-1-0716-2067-0_1 (Humana, 2022).
62. Vasimuddin, M., Misra, S., Li, H. & Aluru, S. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. IEEE International Parallel and Distributed Processing Symposium (IPDPS), Rio de Janeiro, Brazil, 2019 pp. 314–324 (2019).
63. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience*, **10**. <https://doi.org/10.1093/gigascience/giab008> (2021).
64. Trouvelot, A., et al. in *Mycorrhizae. Physiology and Genetics* (eds Gianinazzi-Pearson, V. & Gianinazzi, S) 217–221 (INRA Press, 1986).
65. Fiorilli, V. et al. RiPEIP1, a gene from the arbuscular mycorrhizal fungus *Rhizophagus irregularis*, is preferentially expressed in planta and may be involved in root colonization. *Mycorrhiza* **26**, 609–621 (2016).
66. Ceasar, S. A., Hodge, A., Baker, A. & Baldwin, S. A. Phosphate concentration and arbuscular mycorrhizal colonisation influence the growth, yield and expression of twelve PHT1 family phosphate transporters in foxtail millet (*Setaria italica*). *PLoS ONE* **9**, e108459 (2014).
67. Pudake, R. N. et al. Expression of four phosphate transporter genes from Finger millet (*Eleusine coracana* L.) in response to mycorrhizal colonization and Pi stress. *3 Biotech* **7**, 17 (2017).
68. Saha, P. & Blumwald, E. Assessing reference genes for accurate transcript normalization using quantitative real-time PCR in pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *PLoS ONE* **9**, e106308 (2014).

Acknowledgements

The authors would like to thank Prof. Steven Runo, Kenyatta University, Kenya for the P10 and Aw lines of great purity. Further thanks to Alexander Putra from the KAUST BioScience Core Lab for the primary sequencing, Vijayalakshmi Ponnakanti for greenhouse support, Luis River-Serna for technical support, and the other members of The BioActives Lab for miscellaneous support. We also thank the KAUST supercomputing laboratory (KSL) for providing resources and support; Dr. Helene Berges, Dr. Sandrine Arribat, Dr. William Marande and Dr. Stephane Cauet of INRAE, CNRGV French Plant Genomic Resource Center. We also thank the U.S. National Plant Germplasm System for providing the 10 pearl millet panel. We thank Dr. Prakash Gangashetty and Dr. Rakesh Srivastava of ICRISAT-India for their support and for valuable discussions. We thank Prof. Harro Bouwmeester, Plant Hormones Biology, University of Amsterdam for valuable discussions and

for providing plasmids for transient expression in *N. benthamiana* (pBIN-OsD27, pBIN-OsCCD7 and pBIN-OsCCD8), harboring the indicated cDNAs under the control of the CaMV 35S promoter. This work was supported by baseline funding from King Abdullah University of Science and Technology and by the Bill & Melinda Gates Foundation (grant number OPP1136424) given to S. A.-B.

Author contributions

These authors contributed equally: Hendrik NJ Kuijer, Jian You Wang, Salim Bougouffa. The study was conceived by S.A.-B. Research designed and concept proposed by J.Y.W. and S.A.-B. Genomic DNA extraction by H.N.J.K., genome assemblies by M.A., genome annotations by S.B., repeat analysis by A.Z. Conceptualization of genome sequencing: M.A., S.G.K., R.A.W., X.G., T.G. RNA-Seq by H.N.J.K. with L.B., mapping by I.A., data processing to DEG tables by R.I. CLAMT fragment analysis in resequencing data by S.B. Interpretational bioinformatics by H.N.J.K. Discovery of new SLs and fractioning by J.Y.W. MeCLA synthesis by C.M and A.R.d.L. Chemical analysis, treatments and feeding by J.Y.W with G.T.C. and S.S. Striga bioassay and pearl millet phenotyping by M.J., J.Y.W., L.B., and H.N.J.K. Transient expression by A.B. and D.A. with H.N.J.K. Mycorrhiza colonization analysis by C.V., V.F. and L.L. Manuscript written by H.N.J.K., J.Y.W., and S.A.-B. with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-51189-w>.

Correspondence and requests for materials should be addressed to Salim Al-Babili.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024

¹The BioActives Lab, Biological and Environmental Sciences and Engineering (BESE), King Abdullah University of Science and Technology (KAUST), 23955-6900 Thuwal, Kingdom of Saudi Arabia. ²Center for Desert Agriculture, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. ³Computational Bioscience Research Center, Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), 23955-6900 Thuwal, Kingdom of Saudi Arabia. ⁴Plant Science Program, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. ⁵Department of Life Sciences and Systems Biology, University of Torino; Viale Mattioli 25, Torino 10125, Italy. ⁶Universidad de Vigo, Facultad de Química and CINBIO, 36310 Vigo, Spain. ⁷Institute of Crop Science, Sant'Anna School of Advanced Studies, Pisa 56127, Italy. ⁸Analytical Chemistry Core Lab, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. ⁹These authors contributed equally: Hendrik N. J. Kuijer, Jian You Wang, Salim Bougouffa. ✉ e-mail: Salim.Babili@kaust.edu.sa