Short title: Specialized Metabolite Diversity

Title: Switchgrass metabolomics reveals striking genotypic and developmental differences in saponins

Author for contact:
Robert Last, Department of Biochemistry and Molecular Biology, Michigan State University, 603 Wilson RD, East Lansing, MI 48823 USA, email: lastr@msu.edu, telephone: (517) 432-3278, ORCID ID: 0000-0001-6974-9587

Xingxing Li[a, b], A. Daniel Jones[a, b] and Robert L. Last[a, b, c] *

[a]Department of Biochemistry and Molecular Biology, Michigan State University, MI 48824;
[b] DOE Great Lakes Bioenergy Research Center, Michigan State University, MI 48824.
[c]Department of Plant Biology, Michigan State University, MI 48824

ORCID ID: 0000-0002-7725-0329 (X. L.); 0000-0002-7408-6690 (A.D.J.); 0000-0001-6974-9587 (R.L.L.)

One sentence summary: Switchgrass structurally diverse steroidal saponins and phenolics vary in abundance and structures in a tissue- and ecotype-specific manner.

*Corresponding author: lastr@msu.edu

List of author contributions:
X.L. Conceived and performed research, created figures and tables and wrote manuscript
A.D.J. Conceived of experimental approaches, provided technical input, reviewed data, edited manuscript
R.L.L Conceived of experimental approaches, wrote and edited manuscript

Funding information:

This work was supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0018409.

Email address of Author for Contact: lastr@msu.edu

Abstract

Switchgrass (*Panicum virgatum* L.) is a bioenergy crop that grows productively on low-fertility lands not suitable for food production. We hypothesize that traits such as low soil nitrogen demand, tolerance to water limitation and resistance to insect pests and microbial pathogens are influenced by low molecular weight compounds known as specialized metabolites. We leveraged untargeted liquid chromatography-mass spectrometry (LC-MS) and quantitative gas chromatography-mass spectrometry (GC-MS) to identify differences in above- and below-ground metabolomes of three northern upland and three southern lowland switchgrass cultivars. This analysis documented abundant steroidal saponins and terpenoid glycosides as well as varied phenolic compounds in switchgrass extracts. We identified many metabolite 'features' (annotated as retention time/mass-to-charge ratio pairs), which differentially accumulated between upland and lowland ecotypes. These include saponins built on at least five different steroidal sapogenin cores. The total saponin concentrations were statistically different between roots of the two switchgrass ecotypes. In contrast, flavonoids such as quercetin mainly exhibited a tissue-specific accumulation pattern and predominantly accumulated in shoots. These results set the stage for testing the impacts of differentially accumulating metabolites on biotic and abiotic stress tolerance and inform development of low-input bioenergy crops.

Introduction

Environmentally sustainable and economical production of transportation fuels and industrial feedstocks using plant biomass is a highly desirable goal for the bioeconomy (Sands et al., 2017). Dedicated energy crops that are productive with low or no chemical fertilizers and pesticides on land that is unsuitable for food and fiber crops have received much attention (Sands et al., 2017). This requires development of plants with a suite of 'ideal' traits (Jiao et al., 2010), including perennial life cycle, rapid growth under low soil fertility and water content as well as resilience to pests and pathogens.

Plants are master biochemists, producing a wide variety of general and specialized metabolites adaptive to their ecological niches (Pichersky & Lewinsohn, 2011). The structurally diverse tissue- and clade-specific specialized metabolites play varied roles in plants coping with biotic and abiotic stresses, both by deterring and promoting interactions. For instance, glucosinolates produced by crucifers such as mustard, cabbage and horseradish, mediate interactions with insect herbivores (Hartmann, 2007). Terpenoids – a diverse class of natural products – can function as antimicrobial phytoalexins (Ahuja et al., 2012). Specialized metabolites may also promote beneficial interactions, ranging from pollinator attraction to signaling that leads to beneficial microbe-plant interactions. As a classic example, flavonoids inducing the rhizobial lipochitooligosaccharides ('Nod factors') initiate the rhizobium-legume nitrogen fixation symbiosis (Poole et al., 2018). Strigolactones are signaling molecules involved in plant symbioses with Arbuscular Mycorrhizal fungi, which in turn enables efficient plant phosphate uptake (Massalha et al., 2017). For these reasons, modifying plant specialized metabolism is an attractive target for bioengineering or trait breeding to create low-input bioenergy crops that can thrive on 'marginal' lands.

While hundreds of thousands of specialized metabolites are estimated to be produced by plants (Pichersky & Lewinsohn, 2011), there are reasons why this number is almost certainly an underestimate. First, these metabolites are taxonomically restricted, often showing interspecies or even intraspecies variation (Pichersky & Lewinsohn, 2011); thus any species, ecotype or cultivar sampled will underrepresent phenotypic diversity. Second, specialized metabolites tend to be produced in a subset of cell- or tissue-types in any plant species analyzed; thus cataloging

the metabolic potential of even a single species requires extraction of multiple tissues over the plant's development. Third, accumulation of these metabolites can be impacted by growth conditions and induced by abiotic or biotic stress (Tissier et al., 2014). Finally, identification and structural characterization of newly discovered metabolites require specialized capabilities, typically a combination of mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy analysis (Last et al., 2007). These methodologies traditionally are labor intensive and inherently low throughput. As a result, plant specialized metabolite structures and biosynthetic pathways remain underexplored.

Developments in the field of liquid chromatography mass spectrometry (LC-MS) based untargeted metabolomics (Schrimpe-Rutledge et al., 2016) enable characterization of plant metabolomes in an increasingly high-throughput manner. Rather than identifying the structure of each metabolite in a sample, metabolomics seeks to obtain a snapshot of global patterns based on the mass spectra of metabolites that are chromatographically separated based on their physical properties (Last et al., 2007). This approach can be used as a powerful tool to aid crop breeding (Turner et al., 2016) and assessing biotech crops (Christ et al., 2018) in agriculture.

The North American native perennial switchgrass (*Panicum virgatum* L.) has the potential to be cultivated as a low-input bioenergy crop for growing on nonagricultural land (Sanderson et al., 2006). The two principal ecotypes of switchgrass are phenotypically divergent, including in flowering time, plant size, physiology and disease resistance. Upland ecotypes exhibit robust freezing tolerance but produce relatively low biomass yield in part due to early flowering (Nielsen, 1947; Kiniry et al., 2013; Casler et al., 2015; Sage et al., 2015). Plants of the lowland ecotype typically found in riparian areas produce large amounts of biomass and are more flooding- and heat-tolerant, pathogen-resistant and nutrient-use-efficient than the upland ecotype (Aspinwall et al., 2013; Uppalapati et al., 2013; Casler et al., 2015). However, these lowland ecotypes do not perform well in northern areas, largely due to lack of cold tolerance. Although the characterization of the switchgrass microbiome is still in its infancy, there is evidence that the microbiomes of these two ecotypes differ and change throughout the growing season (Grady et al., 2019; Singer et al., 2019).

Microbiome-associated traits have received significant attention in recent years since they can influence plant health and productivity (Berendsen et al., 2012; Compant et al., 2019; Wei et al.,

2019). Therefore, new switchgrass varieties with enhanced microbiome traits will be highly desirable. C_{27} core steroidal saponins ubiquitous in monocot plants (Moses et al., 2014) were documented in switchgrass in large amounts (Lee et al., 2009). Saponins containing either steroidal or triterpene (C_{30}) cores were hypothesized to selectively modulate the growth of plant root microbiota (Moses et al., 2014; Huang et al., 2019). Beyond saponins, quercetin-derived flavonoids (Uppugundla et al., 2009) and biotic/abiotic stress-elicited $C_{10} - C_{20}$ terpenes (Pelot et al., 2018; Muchlinski et al., 2019) were also detected in switchgrass.

This study developed and deployed approaches to analyze switchgrass metabolomes by untargeted LC-MS and targeted gas chromatography (GC)-MS. We documented specialized metabolite (including terpenoid glycosides and polyphenols) differences in the tiller, rhizome and root tissue of three upland and three lowland switchgrass cultivars at three developmental stages. These included 157 LC-MS features showing >1000-fold accumulation differences between the upland and lowland ecotypes. Steroidal saponins were especially abundant. Multiple forms of aglycones were identified, and their total root contents differed between the two ecotypes. Our study provides a comprehensive analysis of the specialized metabolites produced by different switchgrass cultivars, and sets the stage for developing dedicated bioenergy crops with varied plant and microbiome traits.

Results

Switchgrass untargeted metabolite profiling

To characterize the metabolome of varied switchgrass genotypes at different developmental stages, LC-MS was used to profile metabolites extracted with 80% methanol across a sample panel containing three upland (Dacotah, Summer and Cave-in-Rock) and three lowland (Alamo, Kanlow and BoMaster) cultivars grown from seed in a controlled environment (**Materials and Methods**). Collection of the tiller (shoot), rhizome and root tissues was performed on plants at three developmental stages – vegetative, transition (between vegetative and reproductive) and early reproductive **(Materials and Methods, Fig S1 A**). In total, 6,240 distinct analytical signals (including multiple adducts from single analytes) were detected in positive mode from sample groups included in this study (**Fig S1 B**). This set was reduced to 4,668 features through a process of combining different adducts derived from the same analytes (see Methods). Downstream data analysis selected the 2586 of 4,668 with threshold maximum abundance ≥ 500 (**Supplementary Dataset**). This is referred to as the 'total detected features' throughout this study.

We employed two complementary approaches to obtain evidence for classes of the features. First, Relative Mass Defect (RMD) filtering (Ekanayaka et al., 2015) was used to assign all metabolite signals to putative chemical classes **(Materials and Methods**). As a result, 42% and 15% of the 2586 features in the dataset were annotated as terpenoid glycosides and polyphenol derived metabolites, respectively (**Fig 1A and Supplementary Dataset**). Then, we searched the MS^1 and MS^E data in available online mass-spectral databases (**Materials and Methods**) and found strong matches to 169 previously characterized metabolites (**Supplementary Dataset**).

Striking metabolome differences between tissue types, genotypes and developmental stages

The untargeted metabolome data provide a high-level view of specialized metabolite variation among the samples, and broad patterns of variation were revealed using hierarchical clustering analysis (HCA, **Fig 1B)**. The aerial (tiller) and subterranean tissue (root and rhizome) metabolites differed noticeably, consistent with the hypothesis that there are fundamental dissimilarities between the above- and below-ground tissue metabolomes. In contrast, the rhizome and root tissues were more similar to each other. Differences in metabolomes of the

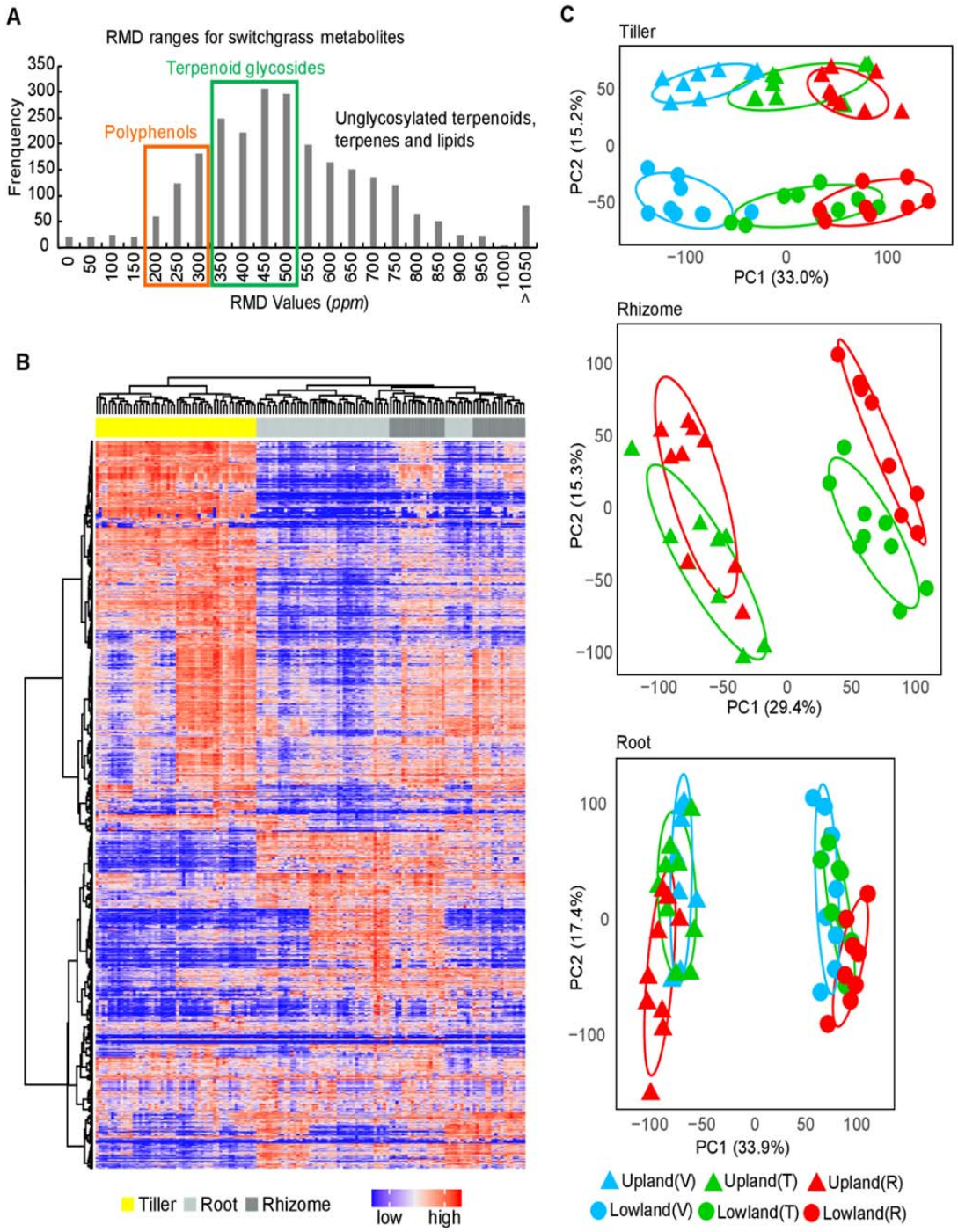


Figure 1. Untargeted metabolome profiling for switchgrass. (A) Histogram of RMD values for the total 2586 features detected in this study by LC-MS in positive ion mode. The green and orange rectangles highlight regions corresponding to the ranges of the RMD values anticipated for terpenoid glycosides and polyphenols, respectively. (B) The metabolome of the six switchgrass cultivars, three tissue types and three developmental stages shown by a heatmap with HCA. The row and column clusters symbolize the 2586 features and 48 sample groups (containing 139 individual samples), respectively. The values representing the metabolite abundances that were used to make the heatmap were scaled to a range from 0 (the lowest abundance) to 1 (the highest abundance). (C) PCA-score plots for the switchgrass tiller, rhizome and root metabolite profiles (n=8 for 'Upland Vegetative' and 'Lowland Reproductive'; n=9 for all the other groups). The percentage of explained variation is shown on the x- and y-axes. V, vegetative phase; T, transition phase; R, reproductive phase.

173 three tissue types by switchgrass cultivars (genotypes) were investigated using principal

component analysis (PCA). Profiles of the tillers clustered into distinct groups in the PCA scores plot, corresponding with the upland (triangles) and lowland (circles) switchgrass ecotypes (**Fig 1C, top panel**). Separation of the metabolite profiles was especially clear for different developmental stages (developmental stages are differentiated based on color in **Fig 1C, top panel**). The PCAs also showed clear-cut differences in metabolite profiles between the upland and lowland genotypes in both the rhizomes and roots (**Fig 1C, middle and bottom panels, respectively**). The developmental stage-associated variance in metabolite profiles of these two subterranean tissues was less apparent compared to that in the aerial tissue. Taken together, the PCA revealed large metabolite differences between the upland and lowland switchgrass cultivars of the three tissues across all three different developmental stages sampled.

Upland and lowland ecotypes have strikingly distinct metabolomes

We next focused on each *developmental stage x tissue type* combination and identified the metabolite features that differentially accumulated in either upland or lowland ecotypes. Surprisingly, 25% (256 of 1035, **Table 1 and Supplementary Dataset**) of the features detected in extracts of the *vegetative-stage tillers* predominantly accumulated in one or the other switchgrass ecotype. Such features were termed as ecotype 'differentially accumulated features' (DAFs). Specifically, there are 126 upland enriched and 130 lowland enriched DAFs in extracts of the *vegetative-stage tillers* (**Fig 2A and Table 1**). In comparison, analysis of *vegetative-stage roots* revealed a total of 879 features with 35% (310, **Table 1 and Supplementary Dataset**) meeting the ecotype DAF statistical threshold (**Fig 2B**). Of these 310 DAFs, approximately equal numbers of features were found to be either upland- (149) or lowland- (161) enriched (**Table 1**). DAFs were also identified for the other *developmental stage x tissue type* combinations (**Fig S2 A-F, Table 1 and Supplementary Dataset)**. Altogether, 1416 unique ecotype DAFs were identified for the eight *developmental stage x tissue type* combinations included in this study (**Fig 2C and Supplementary Dataset)**, accounting for approximately half of the features in the full dataset. These include 770 DAFs predominantly accumulated in upland and 646 predominantly accumulated lowland DAFs (**Fig 2C inset barplot**). Based on RMD filtering 46% and 13% of the DAFs were predicted to be terpenoid glycosides (**Fig 2 A-B and Fig S2 A-F, green dots**) and polyphenol-derived metabolites (**Fig 2 A-B and Fig S2 A-F, orange dots**) respectively.

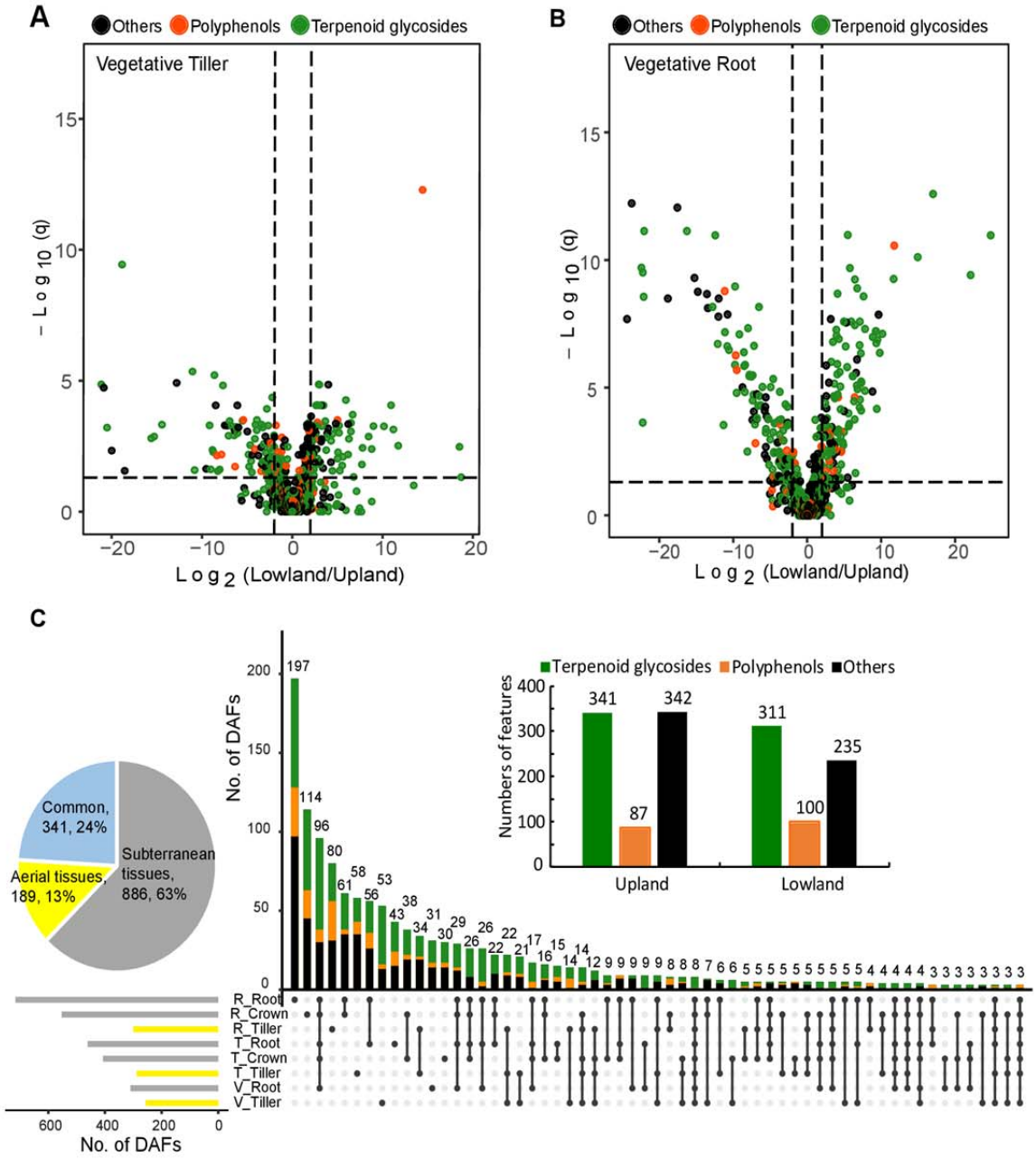


Figure 2. Differentially accumulated features (DAFs) were identified between the upland and lowland ecotypes. Significance analysis (cutoff threshold: adjusted P-value ≤ 0.05; fold changes ≥ 2) was performed to screen for the DAFs between the upland and lowland switchgrass ecotypes (n = 8 or 9) in various developmental stage x tissue type samples. Results of the analyses for (A) vegetative-stage tillers and (B) vegetative-stage roots are shown here using volcano plots. Putative terpenoid glycosides, polyphenols and metabolites from the other categories were classified using RMD filtering and the results color coded. (C) In total, 1416 unique (non-overlapping) ecotype DAFs were identified for the eight developmental stage x tissue type combinations. The inserted barplot shows that upland and lowland ecotypes accumulated similar numbers of the predominant DAFs likely terpenoid glycosides (green) and polyphenols (orange). The inserted pie chart indicates percentages of the DAFs contributed by aerial (tiller) vs. subterranean (root/rhizome) tissues.

203 A striking result from this analysis is that 11% (157) of the 1416 DAFs showed >1000-fold

accumulation difference between the two ecotypes in at least one of the eight *developmental stage x tissue type* combinations, including a few organic acids, sesquiterpenoids and phenolics identified by database searching (**Supplementary Dataset)**. Noticeably, 82% (129/157) of these highly DAFs were unique to subterranean tissues (rhizome and/or root). A similar, but less striking, result was seen for the total 1416 DAFs: 63% (886) were only found in subterranean tissues while only 13% (189) were unique to the aerial tissues. The remaining 24% were detected in both above- and below-ground tissues (**Fig 2C inset pie chart)**. These results suggest that switchgrass subterranean tissues are the major sources of the specialized metabolic genetic diversity.

Examples of large ecotype and tissue differences in switchgrass specialized metabolites

A striking feature of our large dataset is the observation that some metabolites show >1000-fold differences between ecotypes or tissues. This is especially noticeable for features annotated as saponins based on our four criteria (**Materials and Methods**), where we observed 21 unknown DAFs with >1000-fold accumulation difference between ecotypes. As an example of the saponin feature annotation, **Fig 3A** shows a portion of chromatogram that is rich in terpenoid glycosides and polyphenols. The HPLC peak eluting at 6.09 min was the most abundant saponin-like feature in diverse switchgrass samples included in this study. The MS^1 spectrum of this feature showed $[M + NH_4]^+$ at *m/z* 1212 (**Fig 3A**), $[M - H_2O]^+$ at *m/z* 1177 (**Fig 3B right panel**) and $[M + H]^+$ at *m/z* 1195 (not shown). The (MS^E) fragment mass spectra (**Fig 3B right panel**) revealed ions consistent with two sequential losses of 162 Da (each interpreted as glucose minus H_2O) and three neutral losses of 146 Da (interpreted as losses of three anhydrorhamnose units). A fragment ion ($C_{27}H_{43}O_3^+$) at *m/z* 415 with an RMD value of 795 *ppm* was also detected (**Fig 3B right panel**), matching the molecular formula of protonated diosgenin ($C_{27}H_{42}O_3$, 414 Da) as described by Lee et al. (2009). Based on the interpretation of the MS data, we proposed a putative structure of this saponin (**Fig 3B right panel**).

In total, we identified 176 distinct saponin-like features (**Supplementary Dataset**) with four different core fragment ion masses (**Table S1**). To compare abundances of these saponin-like features across the switchgrass samples, ion current intensities (peak areas) of the 176 features were combined and termed as the 'total saponin'. The right panel in **Fig 3C** shows that roots of the lowland ecotype accumulated more total saponin than the upland roots at all three

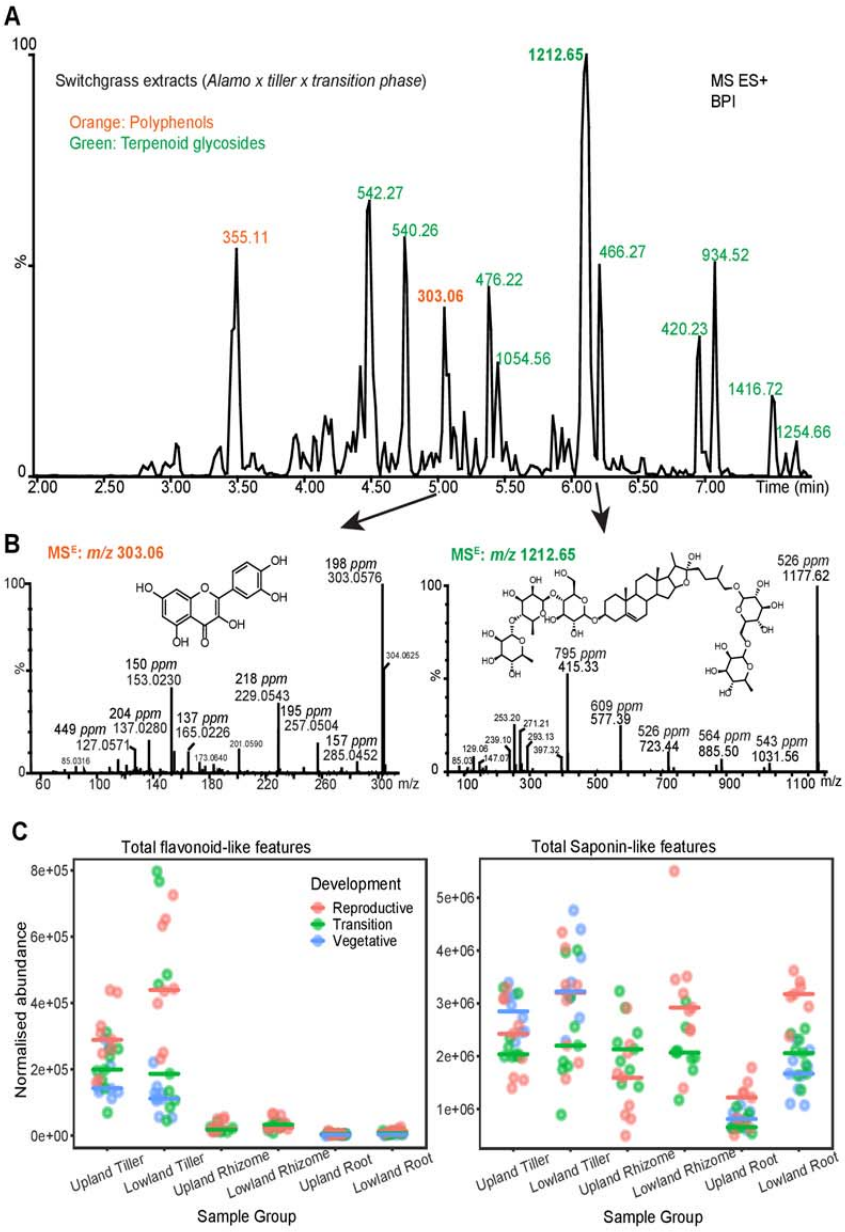


Figure 3. Saponins, terpenoid glycosides and polyphenols (e.g. flavonoids) dominate switchgrass extracts (A) A portion of LC-MS base peak intensity (BPI) chromatogram rich in saponins, terpenoid glycosides and phenolics. This BPI chromatogram was generated from the extracts of the transition-stage tillers of lowland switchgrass, Alamo. The numbers on tops of the peaks are m/z values of the most abundant features in the peaks (Green text, terpenoid glycosides; Orange text, polyphenols). (B) High energy MSE fragmentation pattern of the m/z 1212 saponin was shown in the right panel. The fragment ions (spectra) that are thought to be derived from the fully glycosylated saponin are labeled with their RMD values (upper numbers in ppm). A putative structure of this saponin is shown in the inset. The MSE fragmentation pattern of the m/z 303 quercetin was shown in the left panel. The m/z 127, 153 and 165 were derived from the A ring. The m/z 137 was derived from the B ring. The m/z 229, 257 and 285 correspond to [M+H-H2O-2CO]+, [M+H-H2O-CO]+ and [M+H-H2O]+ respectively. (C) Relative quantification for the sum of saponin-like (right) and flavonoid-like (left) features in each genotype x developmental stage x tissue type class (n = 8 or 9). Colored horizontal bars represent median values. Normalized abundances (y-axis) were calculated as (ion intensity of the feature / ion intensity of the internal standard) x 1000.

234 developmental stages. In contrast, rhizomes of the lowland ecotypes accumulated more total

saponin than the upland rhizomes only at the reproductive stage (**Fig 3C right panel**). Ecotype-specific differences were not observed in aerial tissue: tillers of the two ecotypes accumulated comparable amounts of total saponins across development stages, (**Fig 3C right panel**).

Switchgrass extracts are also rich in polyphenol derived metabolites, such as quercetin-type flavonoids (Uppugundla et al., 2009). Twenty-three flavonoid-like features (**Supplementary Dataset**) were identified in our dataset, including abundant quercetin aglycone (eluting at 5.05 min with molecular ion of *m/z* 303.06; **Fig 3B left panel**). As was done for saponins features, we analyzed differences in total flavonoids across *ecotype x developmental stage x tissue type* combinations. In contrast to total saponins, both switchgrass ecotypes have higher total flavonoids in tillers than rhizomes and roots (**Fig 3C, left panel**). Flavonoids displayed weaker ecotype-specific differences and only in the *reproductive-stage tillers* as a result of a few outlier lowland samples. These accumulated unusually high levels of total flavonoids (**Fig 3C left panel**).

Evidence for multiple switchgrass saponin cores

The LC-MS results revealed that the saponins are highly differentially accumulated between the two switchgrass ecotypes. To more precisely determine saponin concentrations in different tissues and switchgrass cultivars, a GC-MS based quantification method was developed to quantify the sugar-free sapogenin aglycones in a panel of switchgrass samples after hydrolytic removal of sugars (**Materials and Methods**). Because saponins based on the diosgenin core were reported in switchgrass extracts (Lee et al., 2009), commercial diosgenin was used as an external quantification standard. This method detected three distinct diosgenin-derived peaks (D1, D2 & D3, **Fig 4A top trace, Fig S4A top trace and Table 2**) from the commercial standard with D2 as the major peak. We hypothesize that D3 is a less abundant diosgenin structural isomer and D1 a dehydrated diosgenin that formed during sample treatment. Identities of these peaks were inferred from their putative pseudomolecular ions from chemical ionization (CI) MS (**Fig 4B**) and fragment ions from electron ionization (EI) MS (**Fig S4B**). A search of the EI mass spectra against National Institute of Standards and Technology (NIST) NIST17 GC-MS library (www.chemdata.nist.gov) confirmed these compounds as diosgenin (**Fig S4C**).

Hydrolysis and derivatization of switchgrass samples yielded six potential sapogenin peaks in GC-MS chromatograms (S1 – S6, **Fig 4A** and **Fig S4A**). Based on their putative

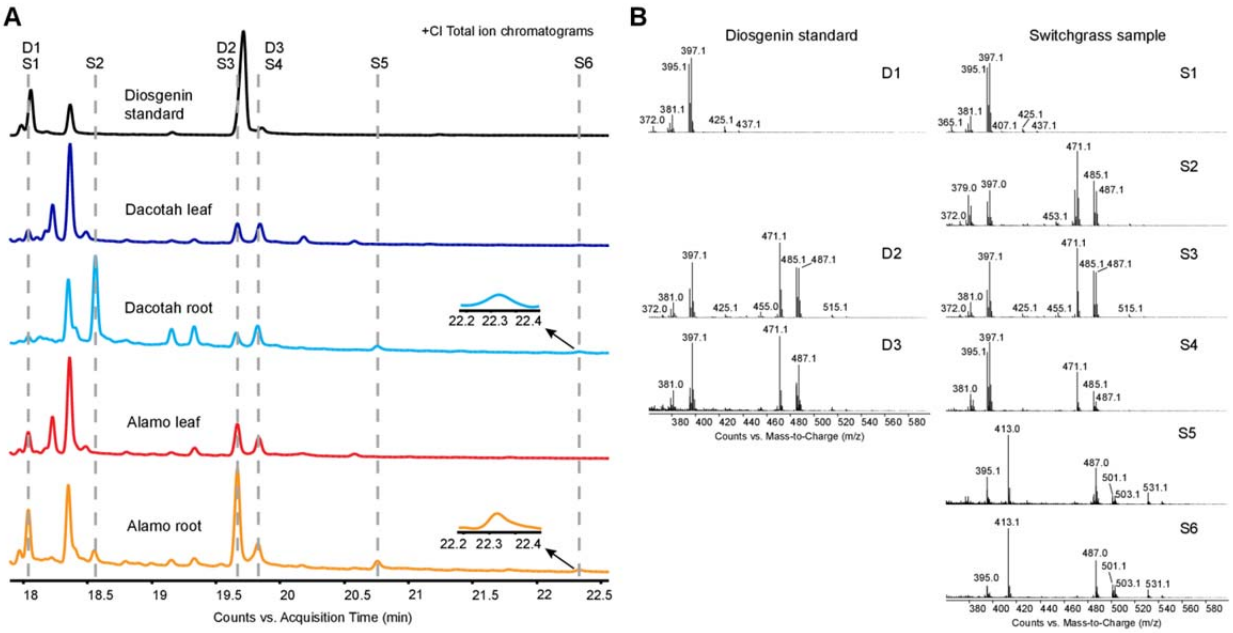


Figure 4. Potential sapogenin peaks were identified in switchgrass extracts by GC-MS. (A) +(CI) GC-MS total ion chromatograms (TICs) of diosgenin standard (black), Dacotah leaf (dark blue), Dacotah root (blue), Alamo leaf (red) and Alamo root (orange). The identified potential sapogenin peaks in the switchgrass and standard samples are indicated and aligned by the dashed lines. Note that the S1, S3 and S4 peaks in all switchgrass samples have slightly different RTs from the D1, D2 and D3 peaks correspondingly in the diosgenin standard sample (black). Zoomed-in views for the peak S6 in Dacotah root and Alamo root are indicated by the arrows. (B) The (CI) GC-MS spectral patterns of the sapogenin peaks detected in the standard and switchgrass samples.

pseudomolecular ions (**Fig 4B**) and fragmentation patterns (**Fig S4B**), we annotated S2, S3 and S4 as diosgenin structural isomers, while S5 and S6 were annotated as oxydiosgenin, which contains one more oxygen than diosgenin. As with D1 in the standard, S1 appears to be a dehydrated diosgenin that formed during sample treatment (**Table 2**). Among these six switchgrass sapogenins, S3 and S4 elute close to each other and have a slightly different retention index compared with D2 and D3, respectively. In contrast, diosgenin isomer S2 has a retention index distinctly shorter than S3 and S4 (**Fig 4A**, **Fig S4A and Table 2**). The two oxydiosgenin isomers, S5 and S6, were below the limit of detection in the diosgenin standard sample. Altogether, these results provide evidence that the substantial chemical diversity in >100 saponins observed by LC-MS is the result of variable glycosylation of at least five different switchgrass sapogenin cores.

Confirmation of differential accumulation of sapogenins in roots of the two ecotypes

Results of GC-MS quantification revealed ecotype-specific sapogenin profiles in subterranean tissues (**Table S2**). This is easily visualized in the GC-MS HCA heatmap in **Fig 5A**, where all

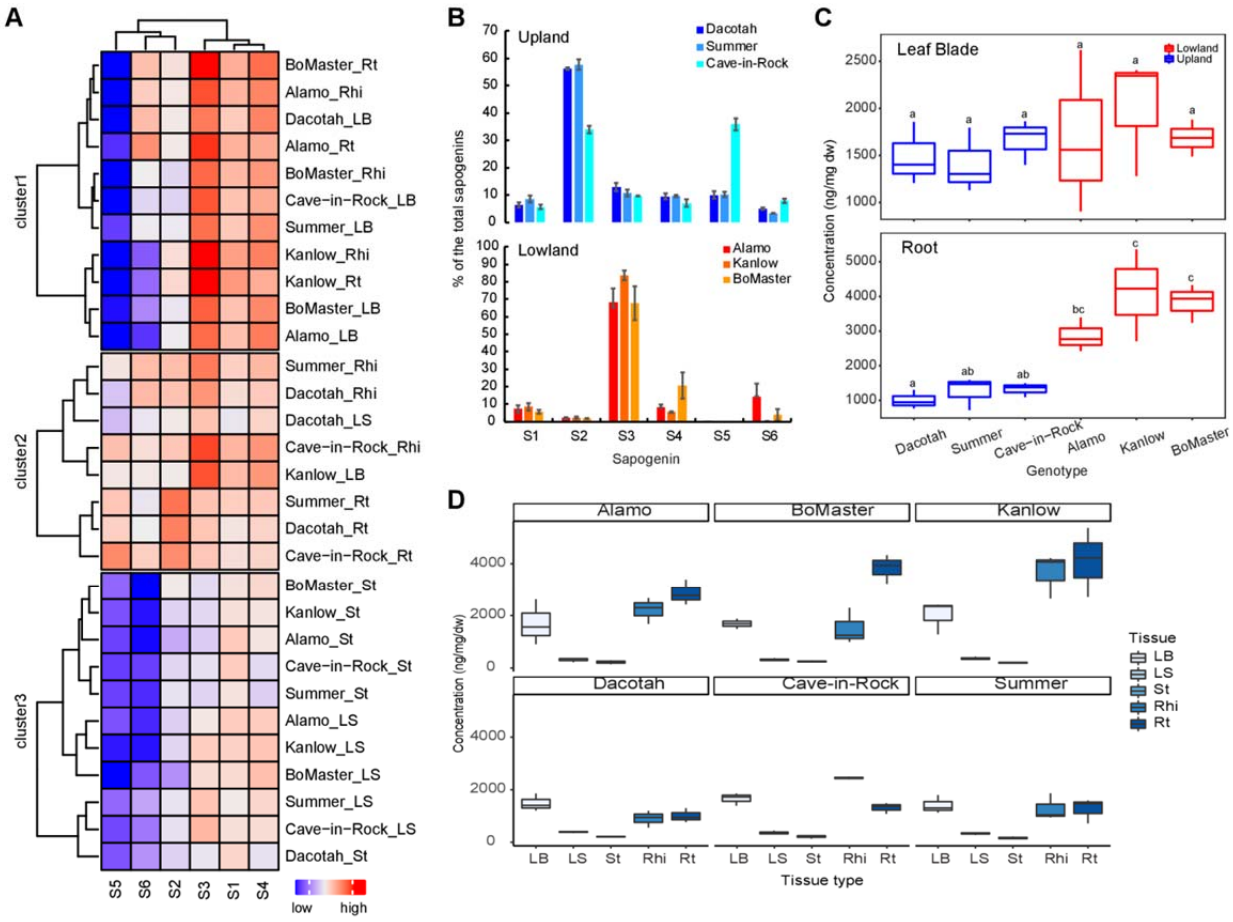


Figure 5. Variations in sapogenin concentrations among the switchgrass genotypes, tissue types and genotype x tissue type. (A) A heatmap with HCA generated using data from Table S2 showing a profile of concentrations of the individual sapogenins in the different genotype x tissue type combinations. The concentration values were log10 scaled to a range between 0 (lowest) and 4 (highest). (B) The ratio of the individual sapogenins in roots of the three upland and three lowland switchgrass cultivars. Heights of the bars reflect the means of the three replicates for individual cultivars; error bars show the standard error of the mean. (C) Comparison of the total sapogenin among the six switchgrass cultivars in leaf blade (Kruskal-Wallis test: P = 0.766) and root (Kruskal-Wallis test, P = 0.016). Different lower-case letters on top of the boxes designate statistically different means (Post-Hoc test: Dunn's test). (D) Comparison of the total sapogenin concentrations among the five tissue types for each switchgrass cultivar. LB, leaf blade; LS, leaf sheath, St, stem; Rhi, rhizome; Rt, root.

subterranean lowland ecotype samples clustered in clade 1, while those of the upland ecotype fall into clade 2. S3 was dominant in roots of the lowland cultivars, representing three-quarters of the total sapogenin, while S2 accounted for half of sapogenins in upland cultivars roots (**Fig 5B**). In contrast, aerial samples displayed tissue-specific relationships across ecotypes: samples from leaf blade (except Kanlow) in clade 1 and sheath and stem (except Dacotah) in clade 3 (**Fig 5A**). Taken together, these results support the LC-MS metabolomics data indicating that the ecotype-specific metabolite differences are greatest in subterranean tissues (**Fig 2 and Fig S2**).

This conclusion was reinforced by analysis of total sapogenin concentrations in tissues of upland vs. lowland switchgrass (**Table S2**). **Fig 5C** (lower panel) shows that the total sapogenins in lowland roots are uniformly higher than those in roots of the three upland cultivars, with Kanlow and BoMaster showing statistical significance ($p<0.05$, Kruskal-Wallis test). However, leaf blade total sapogenins revealed no ecotype-related statistical difference (**Fig 5C upper panel**). Total sapogenin concentrations were also compared across tissue types for each switchgrass cultivar (**Fig 5D**). Comparable total sapogenins were found in leaf blades and roots of the three upland cultivars - Dacotah, Summer and Cave-in-Rock (**Fig 5D lower panel**). By contrast, total root sapogenins were higher than those found in leaf blades of all three lowland cultivars (**Fig 5D upper panel)**. Taken together, quantitative analysis of sapogenin cores supports the results of LC-MS untargeted metabolomics profiling showing strong genetic difference in root saponin abundance.

Discussion

Switchgrass is a compelling low-fertilizer and low-pesticide cellulosic bioenergy crop candidate (Sanderson et al., 2006), which has strong tolerance to marginal conditions, such as low soil fertility and drought. An especially compelling characteristic is the characterization of two principal ecotypes with populations that vary in phenotypes such as biomass production, flowering time and cold tolerance between the two principal ecotypes (Casler et al., 2015). Our results show a strong divergence in metabolite profiles of accessions in the upland and lowland ecotypes, which sets the stage for identifying metabolites involved in local adaptation of switchgrass. As an example, we found that steroidal saponins – a class of glycosylated plant specialized metabolites – are abundant and differentially accumulated in roots of upland and lowland ecotypes. Plant specialized metabolites are documented to play important roles in recruiting beneficial microorganisms and combating harmful microbes (Massalha et al., 2017). Our study suggests an opportunity to test the bioactivities of the saponins and potentially other differentially accumulated metabolites in interactions between switchgrass and their microbiomes, which would inform the development of locally adapted switchgrass varieties.

Switchgrass extracts contain abundant terpenoid glycosides and polyphenols

We used a combination of LC-MS based untargeted metabolite fingerprinting and *de novo* feature-annotation (RMD filtering) to investigate switchgrass metabolites from three upland and three lowland cultivars. The samples were collected in different *developmental stages x tissue type* combinations for a total of 48 contrasts. Approximately 42% and 15% of the total detected features in the dataset are annotated as terpenoid glycosides and polyphenol-derived metabolites including flavonoids, respectively. The remaining features fell into other categories such as sugars and lipids. Before this survey, there were relatively few published studies about switchgrass metabolites. Leaf steroidal saponins were identified in switchgrass and other *Panicum* species in the 1990's as a cause of hepatogenous photosensitization in farm animals (Patamalai et al., 1990; Holland et al., 1991; Puoli et al., 1992; Munday et al., 1993). Switchgrass, along with other monocot crops including wheat, rice and maize, are also known to produce diterpenoid-derived antimicrobial phytoalexins (Murphy & Zerbe, 2020). In addition to these terpenoids, polyphenols such as quercitrin (flavonoids), were also documented in

switchgrass (Uppugundla et al., 2009). Our metabolomics analysis confirms and extends these published results.

Above- vs. below-ground switchgrass metabolomes

The switchgrass above- and below-ground tissues were shown to have distinct metabolite profiles (**Fig 1B**) by untargeted metabolomics. The analysis also revealed that metabolomes of the above- and below-ground switchgrass tissues are differently influenced by developmental stages. There are clear differences in metabolite profiles of tillers across the three developmental stages (**Fig 1C, top rectangle**). In comparison, metabolite profiles of rhizomes and roots are similar to each other across developmental stages (**Fig 1C, middle and bottom rectangles**). Metabolic changes in perennial grass belowground tissues are especially important during the start and end of the growing season. For example, in the spring nutrients are mobilized to the shoot to promote early growth, and the reverse occurs going into dormancy (Palmer et al., 2017). As there are documented genetic differences in winter survival between upland and lowland varieties, it is of interest to investigate genotype-specific metabolome differences during these physiologically dynamic periods.

Divergence in metabolomes of upland and lowland ecotypes

We identified 1416 DAFs between the upland and lowland ecotypes, where >60% are exclusively from subterranean tissues (**Fig 2C**). These ecotype DAFs make up more than half of the total detected features. Strikingly, 157 of the 1416 DAFs show > 1,000-fold accumulation difference between the two ecotypes. In contrast, results of a metabolomics analysis for maize showed that metabolite profiles of the six genetically defined populations from a GWAS panel failed to separate in PCA, even when the metabolite data were independently analyzed within the same tissue type (Zhou et al., 2019). The strong divergence in metabolomes between the two switchgrass ecotypes provides abundant genetic diversity that could be deployed to breed locally adapted varieties (Casler et al., 2015; Lowry et al., 2019).

Ecotype divergence between upland and lowland switchgrass was previously documented for many plant traits including flowering time, plant size, physiological processes and disease resistance (Milano et al., 2016). The genetic architectures of these divergent traits are being analyzed using quantitative trait locus (QTL) mapping and whole genome association studies

(GWASs) (Milano et al., 2016; Grabowski et al., 2017; Lowry et al., 2019). Among these, disease resistance is related to plant specialized metabolism in other well-studied systems. For example, avenacins (triterpenoid saponins) protect oats (*Avena spp*.) from the soilborne fungal pathogen, *Gaeumannomyces graminis* var. *tritici,* which causes the 'take-all' disease (Osbourn et al., 1994). It is known that the lowland ecotype is more resistant to rust infections than the upland ecotype (Uppalapati et al., 2013; VanWallendael et al., 2020). Given the divergence between upland and lowland switchgrass metabolomes, the metabolome-based QTL or GWAS might be useful for the discovery of specialized metabolites and corresponding biosynthetic genes/pathways involved in rust resistance of the lowland switchgrass.

Chemical diversity and ecotype-specific accumulation of switchgrass saponins

Saponins stand out as a class of highly abundant switchgrass metabolites. The GC-MS analysis identified six different steroidal sapogenins, including an anhydrodiosgenin (S1), three diosgenin isomers (S2, S3 and S4) as well as two oxydiosgenin isomers (S5 and S6) that were not previously reported in the switchgrass literature (**Table 2)**. Diosgenin and oxydiosgenin explained two of the four sapogenin core masses detected by LC-MS analysis (*m/z* 415 and 431, respectively, **Table S1**). Despite the abundance of glycosylated triterpene and larger steroidal saponins in LC-MS analysis, we did not obtain evidence for corresponding sapogenin core masses (*m/z* 455 and 457, **Table S1**) in our GC-MS analysis. This is probably due to the lack of standards for these sapogenins.

Glycosylation strongly contributes to the substantial chemical diversity of switchgrass saponins. Differences in glycosylation of only about five sapogenin cores lead to more than 100 distinct saponin-like features ranging from 800 to 1400 Da. This glycosylation diversity comes from the composition, quantity and position of oligosaccharide chains. The most common monosaccharides in saponin sugar chains are glucose, galactose, arabinose, rhamnose, xylose and glucuronic acid (Thimmappa et al., 2014). Neutral losses corresponding to anhydrous glucose/galactose and rhamnose that associate with switchgrass saponins (under the high dissociation-energy MS^E mode) were observed in this study. Glycosylation increases polarity of saponins and is often associated with bioactivity (Thimmappa et al., 2014).

Total saponins show an ecotype-specific accumulation in switchgrass roots (**Fig 3C right panel and Fig 5C lower panel**). The average saponin concentrations in roots of three upland cultivars,

Dacotah, Summer and Cave-in-Rock, are 1.0, 1.3 and 1.3 mg/mg of dw, respectively. The average saponins in roots of three lowland cultivars, Alamo, Kanlow and BoMaster, are statistically higher, at 2.9, 4.1 and 3.8 mg/mg of dw respectively (**Table S2**). These root saponin contents are intermediate between previously studied plants. On the one extreme, cereals and grasses are generally deficient in saponins, with some exceptions such as oat, which accumulates both triterpenoid and steroidal saponins (Osbourn, 2003). Legumes fall at the other extreme: *Medicago truncatula* (Huhman et al., 2005) and two *Medicago sativa* cultivars, Radius (Bialy et al., 1999) and Kleszczewska (Oleszek et al., 1990), contain 5.9, 5.0 and 9.3 µg/mg dw saponins in their roots, respectively. The availability of cultivars with different root saponins makes switchgrass an attractive target for understanding the bioactivity of these specialized metabolites. A comparative metatranscriptomics study (Turner, et al., 2013) revealed strong rhizosphere microbiome differences among crops: the cereal wheat did not accumulate detectable saponin, the cereal oat produced anti-fungal avenacin saponins , the legume pea produced large amounts of triterpenoid saponins, as well as in an oat mutant deficient in avenacin production. Thus, differential saponin accumulation in switchgrass roots may be indicative of microbiome-modulator activities of these compounds.

Conclusion

Our metabolomics survey of switchgrass revealed differences across tissue types, developmental stages and cultivars. The results provide evidence for especially large divergence in metabolomes of the upland and lowland switchgrass ecotypes. Particularly, steroidal saponins differentially accumulated in roots of upland and lowland ecotypes, consistent with potential microbiome-modulator activities of these compounds. Altogether, our analysis offers a starting place to identify targets for plant breeding or metabolic engineering that may select for adaptive microbiomes and will be useful for development of locally adapted switchgrass varieties with increased biomass yield and decreased economically- and environmentally-costly inputs.

Materials and Methods

Plant materials

The xix switchgrass cultivars used in this study were the upland ecotypes Dacotah, Summer and Cave-in-Rock and lowland ecotypes Alamo, Kanlow and BoMaster. The seeds were ordered from Native Connections (http://nativeconnections.net, Three Rivers, MI). The plants were grown under controlled growth conditions: temperature set at 27 °C with 16 h light (500 μE $m^{-2}s^{-1}$) per day and relative humidity set to 53%. Seeds were sown directly in a 1:1 mixture of sand and vermiculite, watered twice a week with deionized water and fertilized once every two weeks using half-strength Hoagland's solution (Hoagland & Arnon, 1950).

For untargeted LC-MS analysis, plant tissues were harvested at one-, two- and three-months after imbibition, corresponding to the vegetative, transition and early reproductive developmental stages, respectively. Roots, rhizomes and tillers (a portion of stem with several leaves, **Fig S1 A**) were collected separately for all the switchgrass cultivars. For example, one sample represented a specific *cultivar (genotype) x developmental stage x tissue type* combination (**Fig S1 B**). There were three biological replicates from three independent plants with two exceptions: two samples each from two independent plants for vegetative phase Cave-in-Rock samples and early reproductive phase Alamo samples. For the GC-MS quantification of sapogenins, samples were only collected from the 3-month-old (early reproductive phase) plants. All samples were immediately frozen in liquid nitrogen and stored at -80 °C until extraction.

Metabolite extraction

All chemicals were obtained from Sigma Aldrich (St. Louis, MO) unless otherwise specified. The samples were frozen in liquid nitrogen and powdered using 15-mL polycarbonate grind vial sets (OPS Diagnostics, Lebanon, NJ) on a Mini G high throughput homogenizer (SPEX SamplePrep, Metuchen, NJ). 500 mg of each sample was extracted at 4 °C overnight (14 – 16 hours) in 5 mL of 80% methanol containing 1 μM telmisartan internal standard. Extracts were centrifuged at 4000 g for 20 min at room temperature to remove solids. Supernatant from each sample was transferred to an HPLC vial and stored at -80 °C prior to LC-MS analysis. For GC-MS, 50 mg of lyophilized sample was extracted in 1 mL of 80% methanol following the workflow described for LC-MS sample preparation above, as described by Tzin et al. (2019).

UPLC-ESI-QToF-MS analysis

Reversed-phase Ultra Performance Liquid Chromatography – Positive Mode Electrospray Ionization – Quadrupole Time-of-Flight MS (UPLC-(+)ESI-QToF-MS) analyses were performed with a Waters Acquity UPLC system coupled to a Waters Xevo G2-XS quadrupole time-of-flight (QToF) mass spectrometer (Waters, Milford, MA). The chromatographic separations were performed using a reversed-phase, UPLC BEH 1.7 μm C18, 2.1 mm x 150 mm column (Waters) with a flow rate of 0.4 mL/min. The mobile phase consisted of solvent A (10 mM ammonium formate/water) and solvent B (100% acetonitrile). The column oven was maintained at 40 °C. Separations were achieved utilizing a 20-min method, injecting 10 μL of extract and using the following method (%A/%B): 0-1.0 min hold (99/1), linear gradient to 15 min (1/99), hold (1/99) until 18 min, returning at 18.01 min (99/1) and holding until 20 min. The Xevo G2-XS QToF was operated using the following static instrument parameters: desolvation temperature of 350 °C; desolvation gas flow rate at 600 L/h; capillary voltage of 3.0 kV; cone voltage of 30 V. Mass spectra were acquired in continuum mode over *m/z* 50 to 1500 using data-independent acquisition (DIA) and MS^E, with collision potential scanned between 20 – 80 V for the higher-energy function. The MS system was calibrated using sodium formate, and leucine enkephalin was used as the lock mass compound but automated mass correction was not applied during data acquisition. QC and reference samples were analyzed every 20 injections to evaluate the stability of the LC-MS system.

Data processing and metabolite mining for the untargeted metabolomics analysis

Acquired raw MS data were processed using the Progenesis QI software package (v.3.0, Waters, Milford, MA) using retention time (RT) alignment, lock mass correction, peak detection, adduct grouping and deconvolution. The identified compounds were defined by the RT and *m/z* information and we also refer to these as *features*. The parameters used with Progenesis processing were as follows: sensitivity for peak picking, default; minimum chromatographic peak width, 0.15 min and RT range, 0.3 to 15.5 min. Intensities (ion abundances) of all the detected features were normalized to the internal standard, telmisartan, before downstream statistical analyses. Online databases – including KEGG, MassBank, PubChem and MetaboLights – were used to annotate features based on 10 *ppm* precursor tolerance, 95% isotope similarity and 10 *ppm* theoretical fragmentation pattern matching with fragment tolerance.

The complementary method Relative Mass Defect (RMD) filtering (Ekanayaka et al., 2015) was used to assist annotation of the unknown features. Briefly, an RMD value of each feature was calculated in *ppm* as (mass defect/measured monoisotopic mass) x 10^6. This value reflects the fractional hydrogen content of a feature and provides an estimate of the relative reduced states of carbons in the metabolite precursor of that feature. Such information is useful for predicting chemical categories of the features detected in untargeted metabolomics; in this study helping to define whether a feature is a terpenoid glycoside (RMD of 350-550) or phenolic (RMD of 200-350). Features with RMD > 1200 are likely contaminants (e.g. inorganic salts) in the MS system.

The criteria used for predicting saponin-like features are: 1) precursor ion mass range of 800 – 1500 Da; 2) 350 – 550 *ppm* pseudomolecular ion RMD; 3) one or multiple neutral losses of the mass matching a monosaccharide (or monosaccharide – H_2O) detected at elevated collision energy; 4) a fragment ion in the range of *m/z* 400 – 500 with RMD >700 *ppm* consistent with an aglycone, only detected at elevated collision energy. The criteria used for predicting flavonoid-like features are: 1) precursor ion *m/z* > 271 (corresponding to $[M+H]^+$ of the flavone, apigenin); 2) precursor ion RMD range of 180 – 350 *ppm*; 3) detection of fragment ions with *m/z* 127, 137, 153 or 165 at elevated collision energy. These are common characteristic fragment ions derived from the A-ring for commonly observed flavones and flavonols (Ma et al., 1997).

Analysis of switchgrass sapogenins by acid hydrolysis, derivatization and GC-MS

To analyze sapogenins, acid hydrolysis was carried out according to the protocol of Kiełbasa et al. (2019). In brief, 300 µL of switchgrass extract, 200 µL of distilled water and 100 µL of 12 M hydrochloric acid (MilliporeSigma, Burlington, MA) were mixed in a polypropylene microcentrifuge tube and incubated at 85°C for 2 h. The samples were cooled and evaporated to dryness under vacuum with the temperature ≤ 40 °C. The resultant pellet was dissolved in 500 µL distilled water and extracted with 500 µL ethyl acetate for phase partition. After this, 300 µL of the ethyl acetate layer was transferred to a new microcentrifuge tube and evaporated to dryness under vacuum at room temperature. The dry residue was dissolved in 100 µL *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA), derivatized overnight at 60 °C and analyzed using a 30 m VF5 column (Agilent Technologies, Santa Clara, CA; 0.25 mm ID, 0.25 µm film thickness) coupled to an Agilent 5975 single quadrupole MS (Agilent Technologies, Santa Clara, CA) operated using 70 eV electron ionization (EI) or chemical ionization (CI). The MS scanning

range was *m/z* 80 – 800. Splitless sample injection was used, with helium as carrier gas at constant flow of 1 mL/min and the inlet and transfer line held at 280 °C. The GC temperature program was as follows: held at 50 °C for 1 min; ramped at 30 °C /min to 200 °C; ramped at 10 °C /min to 320 °C and held for 10 min.

Relative quantification of switchgrass sapogenins was performed using commercially available diosgenin (~95%, Sigma Aldrich) as an external standard. Serially diluted standards (6 - 192 µg/mL dissolved in 80% ethanol) were pre-treated in the same way as the other samples: after hydrochloric acid hydrolysis, four target peaks were identified that are derived from the commercial diosgenin. They were termed as "diosgenin 1 – 4" (D1 – D4). D4, eluting at 21.3 minutes, could be detected by EI GC-MS at high standard concentrations (**Fig S3A**). All target peaks were combined when plotted against the standard's concentrations to generate six-point response curve. Duplicate technical replicate analyses were done for each standard sample used to generate the standard response curve, which was linear ($r^2 > 0.97$, **Fig S3B**), and was used to calculate relative concentrations for sapogenins detected in switchgrass extracts. The relative quantifications were based on the peak area calculated from total ion chromatograms for standards and the unknown switchgrass sapogenins with chemical structures similar to diosgenin. Six individual plants were harvested for each switchgrass genome type and pooled into three groups of two individual plants. Pooling permitted collection of enough tissue to perform separate analysis of leaf blade, leaf sheath, stem, rhizome and root to overcome the issue of limited amount of plant tissue.

Statistical analysis

To visualize the metabolomic variation in tissue types, developmental stages and genome types of switchgrass, we used Hierarchical Clustering Analysis (HCA) and Principal Component Analysis (PCA). Signals were normalized to internal standard area and tissue mass, and log-transformed prior to these analyses. To assess the relationship among samples and among features, hierarchal clustering with Euclidean distance as the similarity measure and Ward.D2 as the clustering algorithm was used. The relationship results were visualized in the form of dendrograms on the heatmap. Significance analyses were carried out using the Progenesis QI software (Waters) to identify the differentially accumulated features (DAFs) between the upland and lowland ecotypes. The cutoff threshold of the significance analyses was adjusted Student's t-

test *p*-value ≤ 0.05 and fold-change ≥ 2. Results of the analyses were visualized by volcano plots. To examine statistical differences in the total sapogenin concentrations among samples Kruskal-Wallis test and Post-Hoc Dunn's tests were performed in R (v. 3.5.1). *P*-value ≤ 0.05 was considered statistically significant.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. The method for sample collection and sample panel involved in this study.

Supplemental Figure S2. Significance analysis for the differentially accumulated features (DAFs) between the upland and lowland switchgrass ecotypes for each given "tissue type x development stage" combination.

Supplemental Figure S3. The six-point instrument response curve over a range of 6 – 192 mg/ml by serial dilution of the reference standard, diosgenin.

Supplemental Figure S4. Potential sapogenin peaks identified in switchgrass extracts by (EI) GC-MS.

Supplemental Table S1. The top 20 most abundant (according to the "Average Abundance") saponin-like features detected in switchgrass.

Supplemental Table S2. Relative sapogenin concentrations (ng/mg dw) in different tissues of three upland switchgrass cultivars (Dacotah, Summer and Cave-in-Rock) and three lowland switchgrass cultivars (Alamo, Kanlow and BoMaster).

Supplemental Data S1. All LC-MS features reported in this study

Supplemental Data S2. Differentially accumulated features (vegetative tiller)

Supplemental Data S3. Differentially accumulated features (vegetative root)

Supplemental Data S4. Differentially accumulated features (transition tiller)

Supplemental Data S5. Differentially accumulated features (transition rhizome)

560 **Supplemental Data S6.** Differentially accumulated features (transition root)

561 **Supplemental Data S7.** Differentially accumulated features (reproductive tiller)

562 **Supplemental Data S8.** Differentially accumulated features (reproductive rhizome)

563 **Supplemental Data S9.** Differentially accumulated features (reproductive root)

564 **Supplemental Data S10.** Higly differentially accumulated features

565

566 **ACKNOWLEDGMENTS**

567 We thank Dr. Gregory Bonito (Michigan State University) for providing the switchgrass
568 seeds used in this study and helpful advice on the manuscript; Dr. Anne-Sophie Bohrer
569 (Michigan State University) for helping with the manuscript editing. This material is based upon
570 work supported by the Great Lakes Bioenergy Research Center, U.S. Department of Energy,
571 Office of Science, Office of Biological and Environmental Research under Award Number DE-
572 SC0018409.

573

574

Table 1. Numbers of the total detected features and differentially accumulated features (DAFs) in extracts of the tissues from each of the eight *developmental stage x tissue type* combinations.

Developmental stage x Tissue type	**No. of total detected features** [a]	**No. of DAFs**		
		Upland enriched	**Lowland enriched**	**Total** [b]
Vegetative-stage x tiller	1035	126	130	256
Vegetative-stage x root	879	149	161	310
Transition-stage x tiller	1037	155	133	288
Transition-stage x rhizome	1114	268	139	407
Transition-stage x root	965	226	235	461
Reproductive-stage x tiller	1612	100	200	300
Reproductive-stage x rhizome	1378	324	229	553
Reproductive-stage x root	1327	408	309	717

[a] The features are overlapping. Thus, the sum of the numbers in this column is larger than 2586, which is the total detected features in this study.

[b] The DAFs are overlapping. Thus, the sum of the numbers in this column is larger than 1416, which is the total detected DAFs in this study.

583 **Table 2. Different forms of the sapogenins identified in switchgrass extracts by GC-MS**
584 **analysis.** Noticeably, the annotations for TMS-derivatized diosgenin (414 [aglycone] + 72 [TMS]
585 = 486) and oxydiosgenin (430 [aglycone] + 72 [TMS] = 502) are supported by the cross-
586 reference between GC-MS and LC-MS data. The LC-MS analysis detected the positively
587 charged aglycone fragment ions, *m/z* 415 and 431 (**Table S1**), corresponding to the protonated
588 diosgenin and oxydiosgenin respectively. TMS, trimethylsilyl group [$Si(CH_3)_3$].

Peak	Retention index (RI) [a]	Key ion annotations (*m/z*)[b]	Abundant fragment ion *m/z* [c]	Formula of the neutral molecule	Mass of the neutral molecule (Da)	Annotation of the neutral molecule [d]
D1	3116	397 $[M+H]^+$	139, 282	$C_{27}H_{40}O_2$	396	Anhydro-diosgenin
D2	3341	397 $[M+H\text{-}TMS\text{-}H_2O]^+$, 471 $[M\text{-}CH_3]^+$, 485 $[M\text{-}H]^+$, 487 $[M+H]^+$	139, 187, 282	$C_{30}H_{50}O_3Si$	486	Diosgenin (TMS derivatized)
D3	3363	397 $[M+H\text{-}TMS\text{-}H_2O]^+$, 471 $[M\text{-}CH_3]^+$, 485 $[M\text{-}H]^+$, 487 $[M+H]^+$	139, 187, 282	$C_{30}H_{50}O_3Si$	486	Diosgenin (TMS derivatized)
S1	3114	397 $[M+H]^+$	139, 282	$C_{27}H_{40}O_2$	396	Anhydro-diosgenin
S2	3194	397 $[M+H\text{-}TMS\text{-}H_2O]^+$, 471 $[M\text{-}CH_3]^+$, 485 $[M\text{-}H]^+$, 487 $[M+H]^+$	139, 282	$C_{30}H_{50}O_3Si$	486	Diosgenin (TMS derivatized)
S3	3339	397 $[M+H\text{-}TMS\text{-}H_2O]^+$, 471 $[M\text{-}CH_3]^+$, 485 $[M\text{-}H]^+$, 487 $[M+H]^+$	139, 187, 282	$C_{30}H_{50}O_3Si$	486	Diosgenin (TMS derivatized)
S4	3361	397 $[M+H\text{-}TMS\text{-}H_2O]^+$, 471 $[M\text{-}CH_3]^+$, 485 $[M\text{-}H]^+$, 487 $[M+H]^+$	139, 187, 282	$C_{30}H_{50}O_3Si$	486	Diosgenin (TMS derivatized)
S5	3462	395 $[M+H\text{-}TMS\text{-}2H_2O]^+$, 413$[M+H\text{-}TMS\text{-}H_2O]^+$, 487 $[M\text{-}CH_3]^+$, 501 $[M\text{-}H]^+$, 503 [M+H]+	139, 187, 298	$C_{30}H_{50}O_4Si$	502	Oxydiosgenin (TMS derivatized)
S6	3620	395 $[M+H\text{-}TMS\text{-}2H_2O]^+$, 413$[M+H\text{-}TMS\text{-}H_2O]^+$, 487 $[M\text{-}CH_3]^+$, 501 $[M\text{-}H]^+$, 503 [M+H]+	139, 187, 298	$C_{30}H_{50}O_4Si$	502	Oxydiosgenin (TMS derivatized)

589
590 [a] The RIs were calculated using a homologous series of n-alkane standards with the same GC-
591 MS method as the samples (**Fig S4A**).
592 [b] The *m/z* information for the key ion annotations was obtained by (CI) GC-MS (**Fig 4B**).
593 [c] The *m/z* information for the fragment ions was obtained by (EI) GC-MS (**Fig S4B**).
594 [d] The annotations were made according to the pseudomolecular and fragment ion information
595 (the names were given based upon the predicted formulas). The structures of the molecules are
596 unknown.

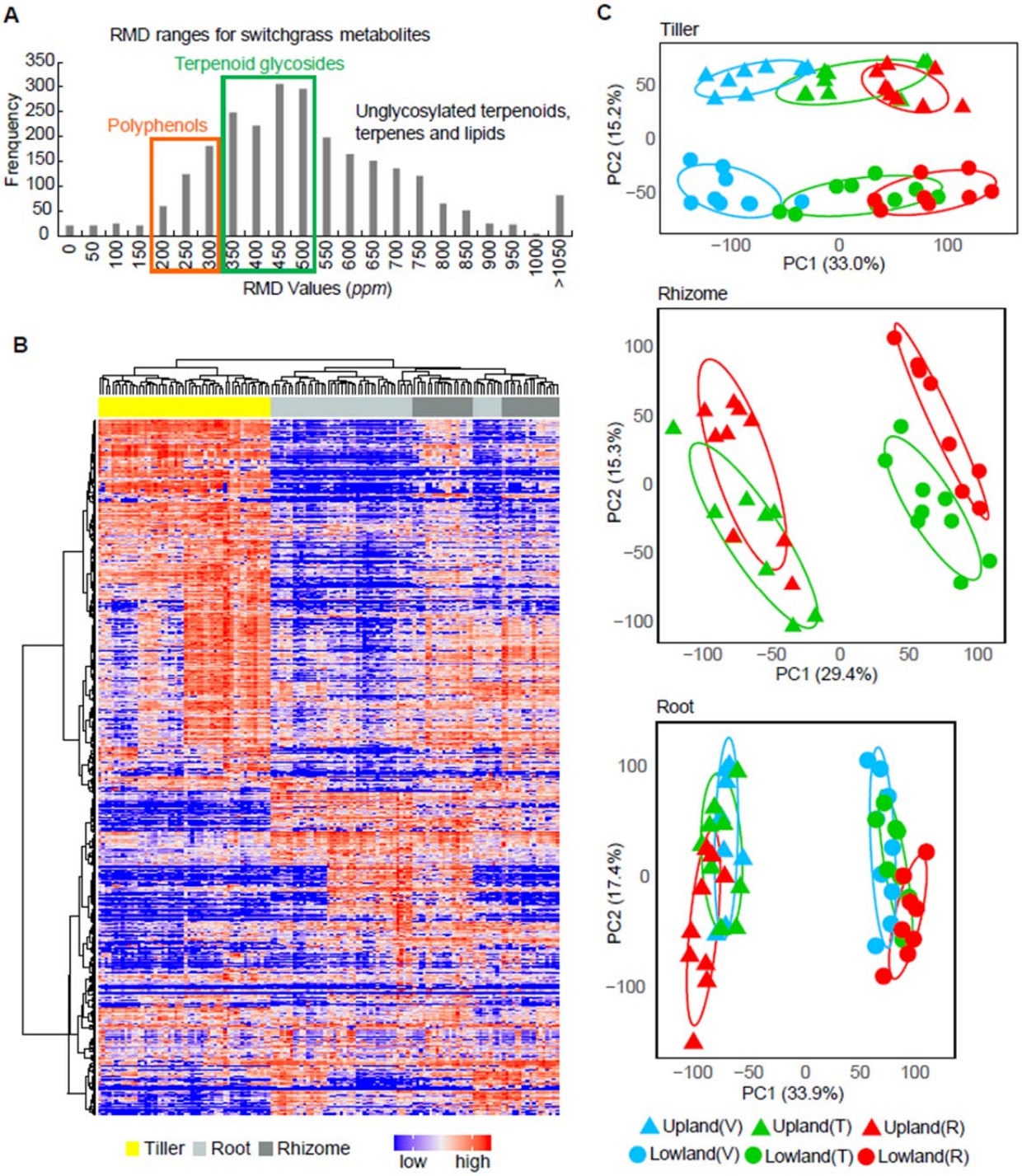


Figure 1. Untargeted metabolome profiling for switchgrass. (A) Histogram of RMD values for the total 2586 features detected in this study by LC-MS in positive ion mode. The green and orange rectangles highlight regions corresponding to the ranges of the RMD values anticipated for terpenoid glycosides and polyphenols, respectively. (B) The metabolome of the six switchgrass cultivars, three tissue types and three developmental stages shown by a heatmap with HCA. The row and column clusters symbolize the 2586 features and 48 sample groups (containing 139 individual samples), respectively. The values representing the metabolite abundances that were used to make the heatmap were scaled to a range from 0 (the lowest abundance) to 1 (the highest abundance). (C) PCA-score plots for the switchgrass tiller, rhizome and root metabolite profiles (n=8 for 'Upland Vegetative' and 'Lowland Reproductive'; n=9 for all the other groups). The percentage of explained variation is shown on the x- and y-axes. V, vegetative phase; T, transition phase; R, reproductive phase.

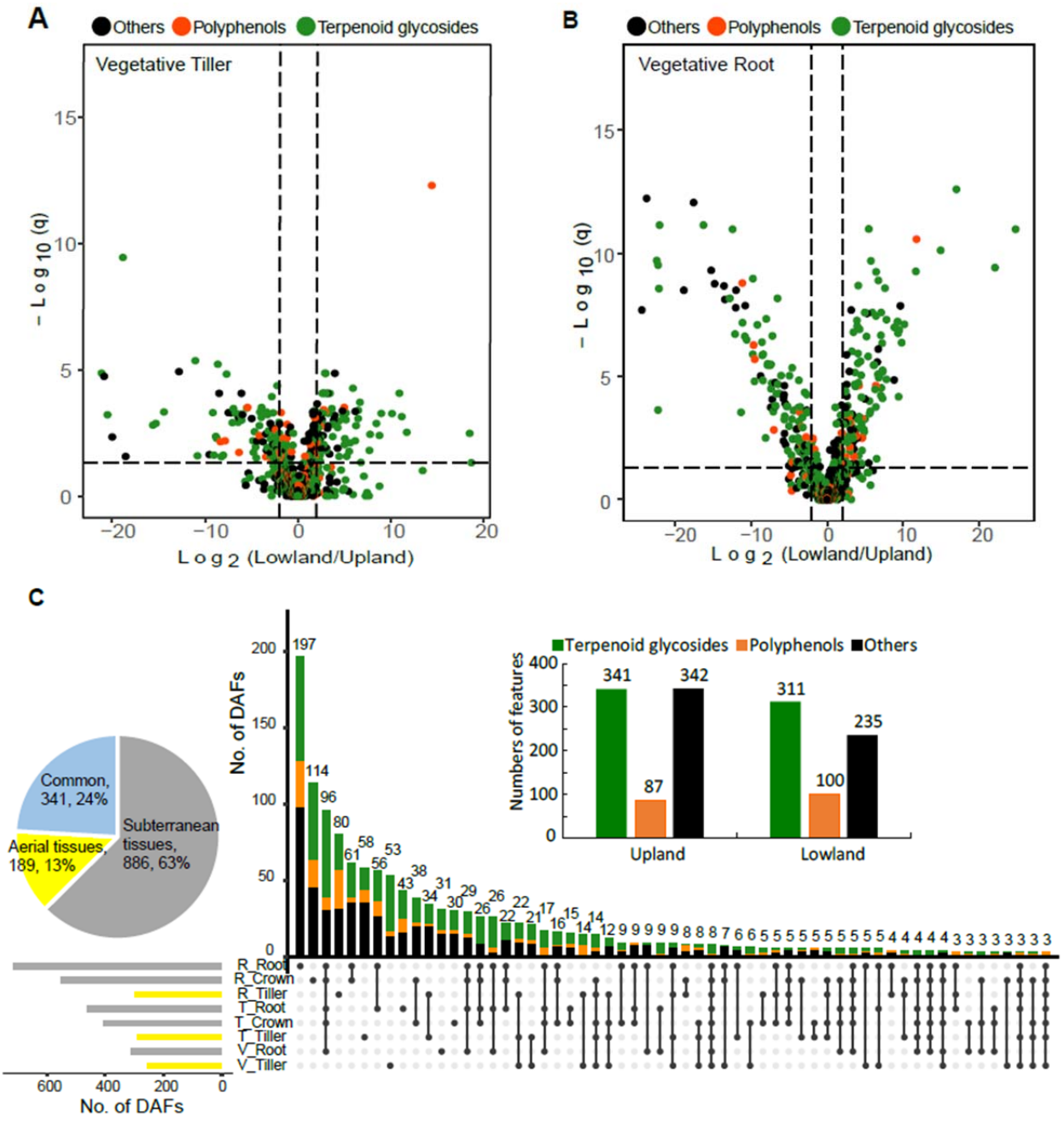
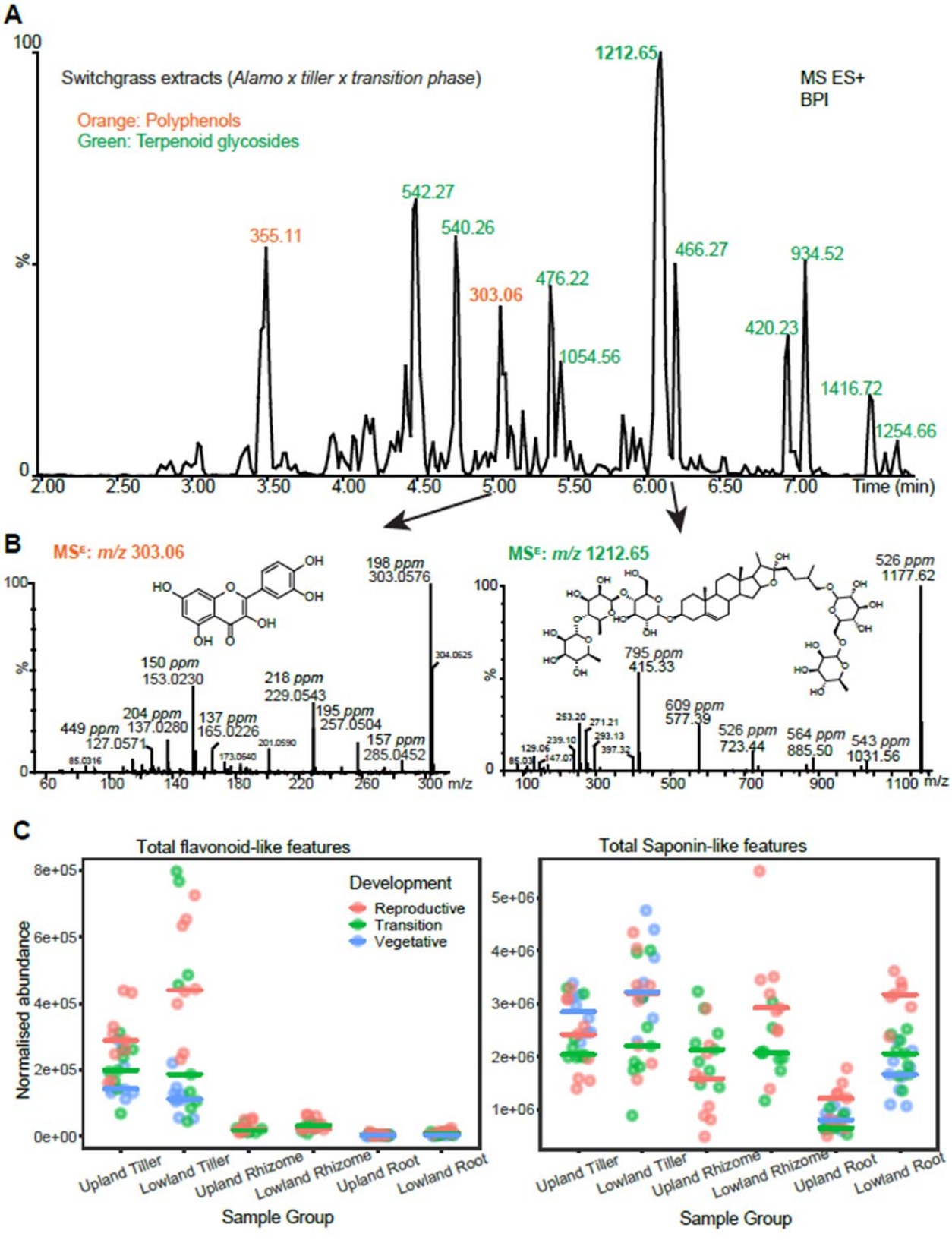


Figure 2. Differentially accumulated features (DAFs) were identified between the upland and lowland ecotypes. Significance analysis (cutoff threshold: adjusted *P*-value ≤ 0.05; fold changes ≥ 2) was performed to screen for the DAFs between the upland and lowland switchgrass ecotypes (n = 8 or 9) in various *developmental stage x tissue type* samples. Results of the analyses for (A) *vegetative-stage tillers* and (B) *vegetative-stage roots* are shown here using volcano plots. Putative terpenoid glycosides, polyphenols and metabolites from the other categories were classified using RMD filtering and the results color coded. (C) In total, 1416 unique (non-overlapping) ecotype DAFs were identified for the eight *developmental stage x tissue type* combinations. The inserted barplot shows that upland and lowland ecotypes accumulated similar numbers of the predominant DAFs likely terpenoid glycosides (green) and polyphenols (orange). The inserted pie chart indicates percentages of the DAFs contributed by aerial (tiller) vs. subterranean (root/rhizome) tissues.



655

Figure 3. Saponins, terpenoid glycosides and polyphenols (e.g. flavonoids) dominate switchgrass extracts (A) A portion of LC-MS base peak intensity (BPI) chromatogram rich in saponins, terpenoid glycosides and phenolics. This BPI chromatogram was generated from the extracts of the *transition-stage tillers* of lowland switchgrass, Alamo. The numbers on tops of the peaks are *m/z* values of the most abundant features in the peaks (Green text, terpenoid glycosides; Orange text, polyphenols). (B) High energy MS^E fragmentation pattern of the *m/z* 1212 saponin was shown in the right panel. The fragment ions (spectra) that are thought to be derived from the fully glycosylated saponin are labeled with their RMD values (upper numbers in *ppm*). A putative structure of this saponin is shown in the inset. The MS^E fragmentation pattern of the *m/z* 303 quercetin was shown in the left panel. The *m/z* 127, 153 and 165 were derived from the A ring. The *m/z* 137 was derived from the B ring. The *m/z* 229, 257 and 285 correspond to $[M+H-H_2O-2CO]^+$, $[M+H-H_2O-CO]^+$ and $[M+H-H_2O]^+$ respectively. (C) Relative quantification for the sum of saponin-like (right) and flavonoid-like (left) features in each *genotype x developmental stage x tissue type* class (n = 8 or 9). Colored horizontal bars represent median values. Normalized abundances (y-axis) were calculated as *(ion intensity of the feature / ion intensity of the internal standard) x 1000*.

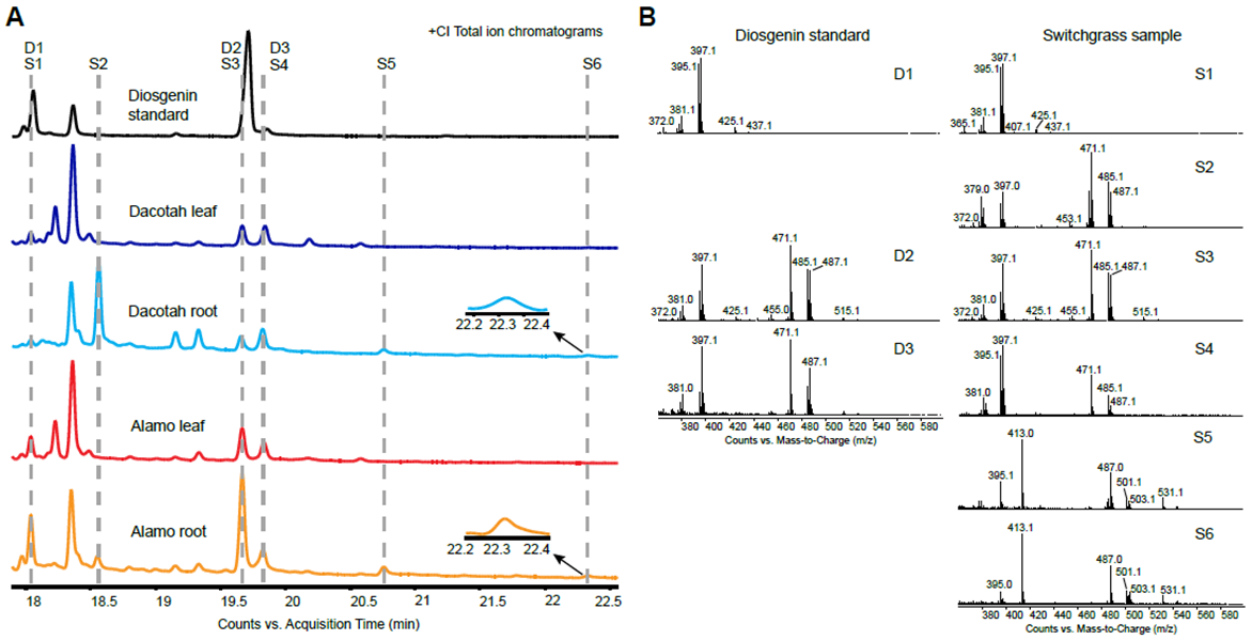


Figure 4. Potential sapogenin peaks were identified in switchgrass extracts by GC-MS. (A) +(CI) GC-MS total ion chromatograms (TICs) of diosgenin standard (black), Dacotah leaf (dark blue), Dacotah root (blue), Alamo leaf (red) and Alamo root (orange). The identified potential sapogenin peaks in the switchgrass and standard samples are indicated and aligned by the dashed lines. Note that the S1, S3 and S4 peaks in all switchgrass samples have slightly different RTs from the D1, D2 and D3 peaks correspondingly in the diosgenin standard sample (black). Zoomed-in views for the peak S6 in Dacotah root and Alamo root are indicated by the arrows. (B) The (CI) GC-MS spectral patterns of the sapogenin peaks detected in the standard and switchgrass samples.

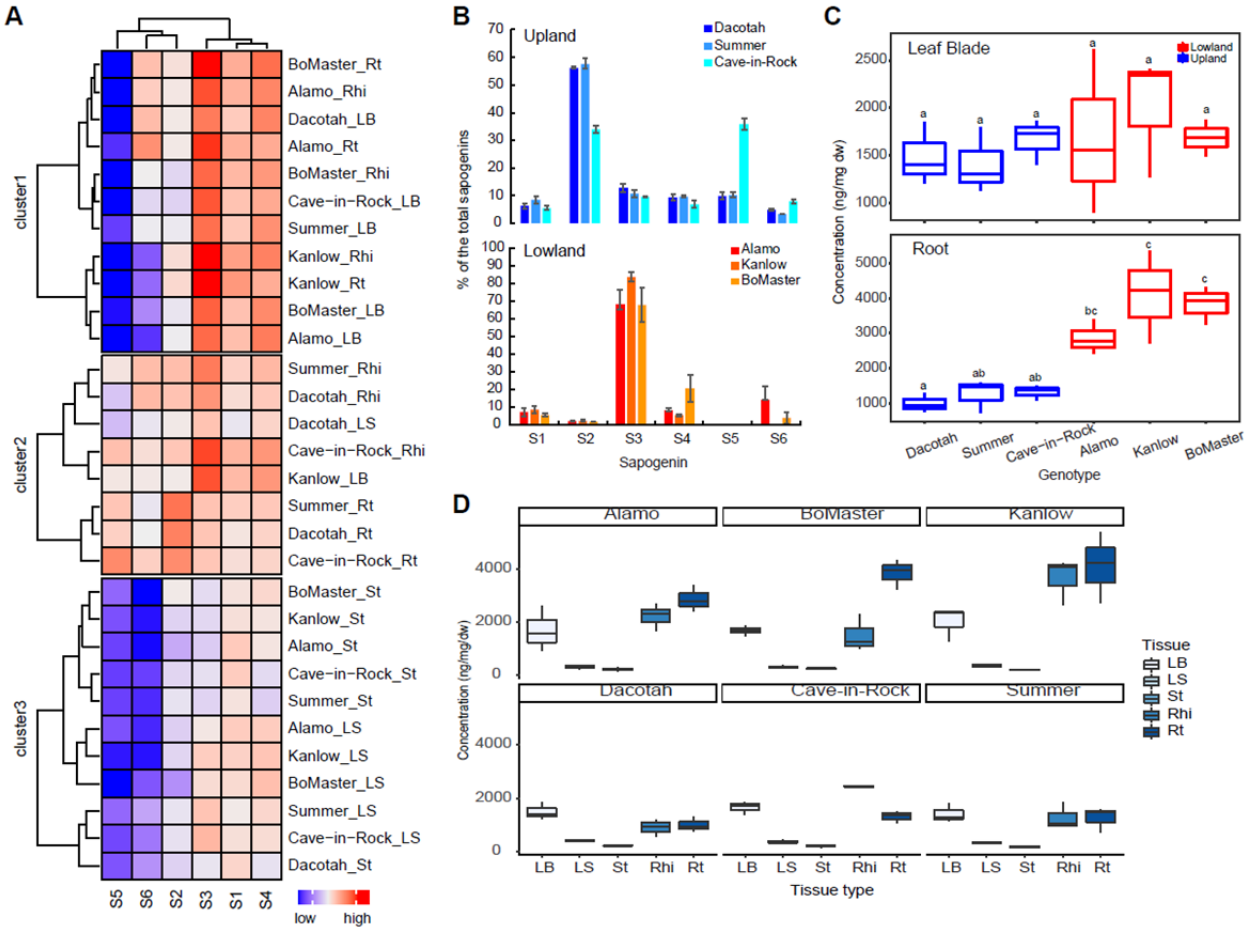


Figure 5. Variations in sapogenin concentrations among the switchgrass genotypes, tissue types and *genotype x tissue type*. (A) A heatmap with HCA generated using data from **Table S2** showing a profile of concentrations of the individual sapogenins in the different *genotype x tissue type* combinations. The concentration values were log10 scaled to a range between 0 (lowest) and 4 (highest). (B) The ratio of the individual sapogenins in roots of the three upland and three lowland switchgrass cultivars. Heights of the bars reflect the means of the three replicates for individual cultivars; error bars show the standard error of the mean. (C) Comparison of the total sapogenin among the six switchgrass cultivars in leaf blade (Kruskal-Wallis test: $P = 0.766$) and root (Kruskal-Wallis test, $P = 0.016$). Different lower-case letters on top of the boxes designate statistically different means (Post-Hoc test: Dunn's test). (D) Comparison of the total sapogenin concentrations among the five tissue types for each switchgrass cultivar. LB, leaf blade; LS, leaf sheath, St, stem; Rhi, rhizome; Rt, root.

717

Parsed Citations

Ahuja, I., Kissen, R., & Bones, A. M. (2012). Phytoalexins in defense against pathogens. Trends in Plant Science, 17(2), 73–90.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Aspinwall, M. J., Lowry, D. B., Taylor, S. H., Juenger, T. E., Hawkes, C. V., Johnson, M.-V. V., Kiniry, J. R., & Fay, P. A. (2013). Genotypic variation in traits linked to climate and aboveground productivity in a widespread C₄ grass: Evidence for a functional trait syndrome. The New Phytologist, 199(4), 966–980.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Berendsen, R. L., Pieterse, C. M. J., & Bakker, P. A. H. M. (2012). The rhizosphere microbiome and plant health. Trends in Plant Science, 17(8), 478–486.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bialy, Z., Jurzysta, M., Oleszek, W., Piacente, S., & Pizza, C. (1999). Saponins in alfalfa (Medicago sativa L.) root and their structural elucidation. Journal of Agricultural and Food Chemistry, 47(8), 3185–3192.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Casler, M. D., Vogel, K. P., & Harrison, M. (2015). Switchgrass Germplasm Resources.

Christ, B., Pluskal, T., Aubry, S., & Weng, J.-K. (2018). Contribution of Untargeted Metabolomics for Future Assessment of Biotech Crops. Trends in Plant Science, 23(12), 1047–1056.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Compant, S., Samad, A., Faist, H., & Sessitsch, A. (2019). A review on the plant microbiome: Ecology, functions, and emerging trends in microbial application. Journal of Advanced Research, 19, 29–37.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ekanayaka, E. A. P., Celiz, M. D., & Jones, A. D. (2015). Relative Mass Defect Filtering of Mass Spectra: A Path to Discovery of Plant Specialized Metabolites. Plant Physiology, 167(4), 1221.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Grabowski, P. P., Evans, J., Daum, C., Deshpande, S., Barry, K. W., Kennedy, M., Ramstein, G., Kaeppler, S. M., Buell, C. R., Jiang, Y., & Casler, M. D. (2017). Genome-wide associations with flowering time in switchgrass using exome-capture sequencing data. New Phytologist, 213(1), 154–169.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Grady, K. L., Sorensen, J. W., Stopnisek, N., Guittar, J., & Shade, A. (2019). Assembly and seasonality of core phyllosphere microbiota on perennial biofuel crops. Nature Communications, 10(1), 1–10.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hartmann, T. (2007). From waste products to ecochemicals: Fifty years research of plant secondary metabolism. Phytochemistry, 68(22), 2831–2846.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hoagland, D. R. (Dennis R., & Arnon, D. I. (Daniel I. (1950). The water-culture method for growing plants without soil. Berkeley, Calif.: College of Agriculture, University of California.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Holland, P. T., Miles, C. O., Mortimer, P. H., Wilkins, A. L., Hawkes, A. D., & Smith, B. L. (1991). Isolation of the steroidal sapogenin epismilagenin from the bile of sheep affected by Panicum dichotomiflorum toxicosis. Journal of Agricultural and Food Chemistry, 39(11), 1963–1965.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Huang, A. C., Jiang, T., Liu, Y.-X., Bai, Y.-C., Reed, J., Qu, B., Goossens, A., Nützmann, H.-W., Bai, Y., & Osbourn, A. (2019). A specialized metabolic network selectively modulates Arabidopsis root microbiota. Science (New York, N.Y.), 364(6440), eaau6389.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Huhman, D. V., Berhow, M. A., & Sumner, L. W. (2005). Quantification of Saponins in Aerial and Subterranean Tissues of Medicago truncatula. Journal of Agricultural and Food Chemistry, 53(6), 1914–1920.
Pubmed: Author and Title

Google Scholar: Author Only Title Only Author and Title

Jiao, Y., Wang, Y., Xue, D., Wang, J., Yan, M., Liu, G., Dong, G., Zeng, D., Lu, Z., Zhu, X., Qian, Q., & Li, J. (2010). Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. Nature Genetics, 42(6), 541–544.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kiełbasa, A., Krakowska, A., Rafińska, K., & Buszewski, B. (2019). Isolation and determination of saponin hydrolysis products from Medicago sativa using supercritical fluid extraction, solid-phase extraction and liquid chromatography with evaporative light scattering detection. Journal of Separation Science, 42(2), 465–474.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kiniry, J. R., Anderson, L. C., Johnson, M.-V. V., Behrman, K. D., Brakie, M., Burner, D., Cordsiemon, R. L., Fay, P. A., Fritschi, F. B., Houx, J. H., Hawkes, C., Juenger, T., Kaiser, J., Keitt, T. H., Lloyd-Reilley, J., Maher, S., Raper, R., Scott, A., Shadow, A., ... Zibilske, L. (2013). Perennial Biomass Grasses and the Mason–Dixon Line: Comparative Productivity across Latitudes in the Southern Great Plains. BioEnergy Research, 6(1), 276–291.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Last, R. L., Jones, A. D., & Shachar-Hill, Y. (2007). Towards the plant metabolome and beyond. Nature Reviews Molecular Cell Biology, 8(2), 167–174.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lee, Stephen T., Mitchell, R. B., Wang, Z., Heiss, C., Gardner, D. R., & Azadi, P. (2009). Isolation, Characterization, and Quantification of Steroidal Saponins in Switchgrass (Panicum virgatum L.). Journal of Agricultural and Food Chemistry, 57(6), 2599–2604.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Lowry, D. B., Lovell, J. T., Zhang, L., Bonnette, J., Fay, P. A., Mitchell, R. B., Lloyd-Reilley, J., Boe, A. R., Wu, Y., Rouquette, F. M., Wynia, R. L., Weng, X., Behrman, K. D., Healey, A., Barry, K., Lipzen, A., Bauer, D., Sharma, A., Jenkins, J., ... Juenger, T. E. (2019). QTL × environment interactions underlie adaptive divergence in switchgrass across a large latitudinal gradient. Proceedings of the National Academy of Sciences, 116(26), 12933.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ma, Y. L., Li, Q. M., Heuvel, H. V. den, & Claeys, M. (1997). Characterization of flavone and flavonol aglycones by collision-induced dissociation tandem mass spectrometry. Rapid Communications in Mass Spectrometry, 11(12), 1357–1364.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Massalha, H., Korenblum, E., Tholl, D., & Aharoni, A. (2017). Small molecules below-ground: The role of specialized metabolites in the rhizosphere. The Plant Journal, 90(4), 788–807.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Milano, E. R., Lowry, D. B., & Juenger, T. E. (2016). The Genetic Basis of Upland/Lowland Ecotype Divergence in Switchgrass (Panicum virgatum). G3.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Moses, T., Papadopoulou, K. K., & Osbourn, A. (2014). Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives. Critical Reviews in Biochemistry and Molecular Biology, 49(6), 439–462.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Muchlinski, A., Chen, X., Lovell, J. T., Köllner, T. G., Pelot, K. A., Zerbe, P., Ruggiero, M., Callaway, L., 3rd, Laliberte, S., Chen, F., & Tholl, D. (2019). Biosynthesis and Emission of Stress-Induced Volatile Terpenes in Roots and Leaves of Switchgrass (Panicum virgatum L.). Frontiers in Plant Science, 10, 1144–1144.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Munday, S. C., Wilkins, A. L., Miles, C. O., & Holland, P. T. (1993). Isolation and structure elucidation of dichotomin, a furostanol saponin implicated in hepatogenous photosensitization of sheep grazing Panicum dichotomiflorum. Journal of Agricultural and Food Chemistry, 41(2), 267–271.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Murphy, K. M., & Zerbe, P. (2020). Specialized diterpenoid metabolism in monocot crops: Biosynthesis and chemical diversity. Phytochemistry, 172, 112289.

Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Nielsen, E. L. (1947). Polyploidy and Winter Survival in Panicum virgatum L.1.

Oleszek, W., Jurzysta, M., Price, K. R., & Fenwick, G. R. (1990). High-performance liquid chromatography of alfalfa root saponins. Journal of Chromatography A, 519(1), 109–116.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Osbourn, A. E., Clarke, B. R., Lunness, P., Scott, P. R., & Daniels, M. J. (1994). An oat species lacking avenacin is susceptible to infection by Gaeumannomyces graminis var. Tritici. Physiological and Molecular Plant Pathology, 45(6), 457–467.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Osbourn, Anne E. (2003). Saponins in cereals. Phytochemistry, 62(1), 1–4.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Palmer, N. A., Saathoff, A. J., Scully, E. D., Tobias, C. M., Twigg, P., Madhavan, S., Schmer, M., Cahoon, R., Sattler, S. E., Edmé, S. J., Mitchell, R. B., & Sarath, G. (2017). Seasonal below-ground metabolism in switchgrass. The Plant Journal, 92(6), 1059–1075.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Patamalai, B., Hejtmancik, E., Bridges, C. H., Hill, D. W., & Camp, B. J. (1990). The isolation and identification of steroidal sapogenins in Kleingrass. Veterinary and Human Toxicology, 32(4), 314–318.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Pelot, K. A., Chen, R., Hagelthorn, D. M., Young, C. A., Addison, J. B., Muchlinski, A., Tholl, D., & Zerbe, P. (2018). Functional Diversity of Diterpene Synthases in the Biofuel Crop Switchgrass. Plant Physiology, 178(1), 54–71.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Pichersky, E., & Lewinsohn, E. (2011). Convergent Evolution in Plant Specialized Metabolism. Annual Review of Plant Biology, 62(1), 549–566.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Poole, P., Ramachandran, V., & Terpolilli, J. (2018). Rhizobia: From saprophytes to endosymbionts. Nature Reviews Microbiology, 16(5), 291–303.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Puoli, J. R., Reid, R. L., & Belesky, D. P. (1992). Photosensitization in Lambs Grazing Switchgrass. Agronomy Journal, 84(6), 1077–1080.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sage, R. F., de Melo Peixoto, M., Friesen, P., & Deen, B. (2015). C 4 bioenergy crops for cool climates, with special emphasis on perennial C 4 grasses. Journal of Experimental Botany, 66(14), 4195–4212.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sanderson, M. A., Adler, P. R., Boateng, A. A., Casler, M. D., & Sarath, G. (2006). Switchgrass as a biofuels feedstock in the USA. Canadian Journal of Plant Science, 86(Special Issue), 1315–1325.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sands, R. D., Malcolm, S. A., Suttles, S. A., & Marshall, E. (2017). Dedicated Energy Crops and Competition for Agricultural Land (No. 1477-2016–121188). AgEcon Search.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schrimpe-Rutledge, A. C., Codreanu, S. G., Sherrod, S. D., & McLean, J. A. (2016). Untargeted Metabolomics Strategies-Challenges and Emerging Directions. Journal of the American Society for Mass Spectrometry, 27(12), 1897–1905.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Singer, E., Bonnette, J., Kenaley, S. C., Woyke, T., & Juenger, T. E. (2019). Plant compartment and genetic variation drive microbiome composition in switchgrass roots. Environmental Microbiology Reports, 11(2), 185–195.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Thimmappa, R., Geisler, K., Louveau, T., O'Maille, P., & Osbourn, A. (2014). Triterpene Biosynthesis in Plants. Annual Review of Plant Biology, 65(1), 225–257.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Tissier, A., Ziegler, J., & Vogt, T. (2014). Specialized Plant Metabolites: Diversity and Biosynthesis. In G.-J. Krauss & D. H. Nies (Eds.), Ecological Biochemistry (pp. 14–37).
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Turner, M. F., Heuberger, A. L., Kirkwood, J. S., Collins, C. C., Wolfrum, E. J., Broeckling, C. D., Prenni, J. E., & Jahn, C. E. (2016). Non-targeted Metabolomics in Diverse Sorghum Breeding Lines Indicates Primary and Secondary Metabolite Profiles Are Associated with Plant Biomass Accumulation and Photosynthesis. Frontiers in Plant Science, 7.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Turner, T. R., Ramakrishnan, K., Walshaw, J., Heavens, D., Alston, M., Swarbreck, D., Osbourn, A., Grant, A., & Poole, P. S. (2013). Comparative metatranscriptomics reveals kingdom level changes in the rhizosphere microbiome of plants. The ISME Journal, 7(12), 2248–2258.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Tzin, V., Snyder, J. H., Yang, D. S., Huhman, D. V., Watson, B. S., Allen, S. N., Tang, Y., Miettinen, K., Arendt, P., Pollier, J., Goossens, A., & Sumner, L. W. (2019). Integrated metabolomics identifies CYP72A67 and CYP72A68 oxidases in the biosynthesis of Medicago truncatula oleanate sapogenins. Metabolomics, 15(6), 85.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Uppalapati, S. R., Serba, D. D., Ishiga, Y., Szabo, L. J., Mittal, S., Bhandari, H. S., Bouton, J. H., Mysore, K. S., & Saha, M. C. (2013). Characterization of the Rust Fungus, Puccinia emaculata, and Evaluation of Genetic Variability for Rust Resistance in Switchgrass Populations. BioEnergy Research, 6(2), 458–468.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Uppugundla, N., Engelberth, A., Vandhana Ravindranath, S., Clausen, E. C., Lay, J. O., Gidden, J., & Carrier, D. J. (2009). Switchgrass Water Extracts: Extraction, Separation and Biological Activity of Rutin and Quercitrin. Journal of Agricultural and Food Chemistry, 57(17), 7763–7770.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

VanWallendael, A., Bonnette, J., Juenger, T. E., Fritschi, F. B., Fay, P. A., Mitchell, R. B., Lloyd-Reilley, J., Rouquette, F. M., Bergstrom, G. C., & Lowry, D. B. (2020). Geographic variation in the genetic basis of resistance to leaf rust between locally adapted ecotypes of the biofuel crop switchgrass (Panicum virgatum). New Phytologist.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wei, Z., Gu, Y., Friman, V.-P., Kowalchuk, G. A., Xu, Y., Shen, Q., & Jousset, A. (2019). Initial soil microbiome composition and functioning predetermine future plant health. Science Advances, 5(9), eaaw0759–eaaw0759.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhou, S., Kremling, K. A., Bandillo, N., Richter, A., Zhang, Y. K., Ahern, K. R., Artyukhin, A. B., Hui, J. X., Younkin, G. C., Schroeder, F. C., Buckler, E. S., & Jander, G. (2019). Metabolome-Scale Genome-Wide Association Studies Reveal Chemical Diversity and Genetic Control of Maize Specialized Metabolites. The Plant Cell, 31(5), 937.
Pubmed: Author and Title
Google Scholar: Author Only Title Only Author and Title