Screening for potential co-products in a *Miscanthus sinensis* mapping family by liquid chromatography with mass spectrometry detection

Ifat Parveen,[a*] Thomas Wilson,[a*] Michael D. Threadgill,[b] Jakob Luyten,[c] Ruth E. Roberts,[a] Paul R. H. Robson,[a] Iain S. Donnison,[a] Barbara Hauck[a] and Ana L. Winters[a]

[a]Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Gogerddan, Aberystwyth, Ceredigion SY23 3EB, UK

[b]Medicinal Chemistry, Department of Pharmacy and Pharmacology, University of Bath, Claverton Down, Bath BA2 7AY, UK

[c]K. H. Kempen University College, Department of Health Care & Chemistry, Kleinhoefstraat 4, 2440 Geel, Belgium

[*]Corresponding author: Tel: +44 (0)1970 823216; Fax: +44 (0)1970 828357; E-mail: ifp@aber.ac.uk (Ifat Parveen), thw7@aber.ac.uk (Thomas Wilson)

Abstract

Society is demanding more green chemicals from sustainable sources. *Miscanthus* is a potential source of platform chemicals and bioethanol through fermentation. *Miscanthus sinensis* (*M. sinensis*) has been found to contain particularly high levels of soluble phenols (hydroxycinnamates and flavonoids) which may have application in the nutraceutical, cosmetic and pharmaceutical industries. Here, we describe the first study on the identification and quantification of phenols from the leaf tissue of a bi-parental *M. sinensis* mapping family. Parents and progeny showed complex profiles of phenols with highly related structures which complicated characterisation of individual

phenotypes. Separation of semi-purified extracts by reverse-phase liquid chromatography, coupled with detection by diode array and ESI-MS/MS, enabled distinction of different profiles of phenols. Ten hydroxycinnamates (*O*-cinnamoylquinic acids) and several flavones (one mono-*O*-glycosyl flavone, eight mono-*C*-glycosyl flavones, two di-*C*-glycosyl flavones, five *O*-glycosyl-*C*-glycosyl flavones and nine 2”-*O*-glycosyl-*C*-glycosyl flavones) were identified and quantified in leaf tissue of two hundred progeny and maternal and paternal plants during the seedling stage. Progeny exhibiting high, moderate and low amounts of hydroxycinnamates and flavonoids and both parents were selected and screened at seven months’ growth to determine the abundance of these phenols at their highest biomass and compared with seedlings. Concentrations of phenols generally decreased as leaves matured. Several novel flavone-glycosides were identified. This technique can be used for rapid screening of plants in a mapping family to identify genotypes with high phenol content to add value in the biorefinery chain. This comparative study provides information on the content of potentially valuable compounds from readily renewable resources and possible biomarkers for identification in breeding programmes.

Highlights

- *Miscanthus sinensis* mapping family: sources of biofuel & platform chemicals
- LC-PDA-ESI-MS^n identified >30 phenols in seedlings and mature leaves
- Luteolin- and apigenin-glycosides were observed in leaves and novel compounds were identified
- Concentrations of phenols decreased as leaves matured; phenols ranged between 0.53 and 12.4 mg/g FW
- Genotypes with high content of phenols were selected to add value in the biorefinery chain

Keywords

Miscanthus sinensis; mapping family; hydroxycinnamates; flavonoids; LC-DAD-ESI-MS^n; green chemicals

1. Introduction

Society is becoming increasingly concerned about the negative impact on the environment associated with the chemical industries and this has led to a demand for the supply of "green" chemicals from sustainable sources. The identification of high-value chemicals within energy crops would increase value in the biorefinery chain, which could increase the adoption of bioenergy and reduce dependence on fossil fuels. A biorefinery using plant-based feedstocks could generate multiple products, including platform chemicals, fuel and power (Cherubini, 2010). Advances in cleaner processing technologies, particularly in fermentation, molecular and genetic engineering, have

increasingly allowed chemical industries to use plant-based, as opposed to petrochemical, feedstocks for the manufacture of commercially important chemicals (Xu et al., 2008). Green chemicals are not limited to the fuel sector but are also relevant to a range of commercial sectors, including the pharmaceutical, nutraceutical, cosmetic, food and beverage industries (Dong et al., 2011; Crozier et al., 2009; An et al., 2008; Aburjai and Natsheh, 2003). The world production of biomass each year is thought to exceed 1×10^{14} Kg (Xu et al., 2008). In the US alone, 2.5×10^{11} Kg of the plant biomass produced each year is wasted and exceeds the current total consumption of 1×10^{11} Kg used for the production of organic chemicals, plastic resins and fibres (Xu et al., 2008). Moreover, only 5% of chemicals are presently sourced from renewable resources and there may be potential for bio-chemicals to share markets with their petrochemical-based counterparts (Xu et al., 2008).

There is huge interest in using perennial grasses, such as *Miscanthus* spp. and *Panicum virgatum*, as a source of renewable biomass (Lewandowski et al., 2000). Components of biomass, such as lignin, cellulose and hemicellulose (Kleinert and Barth, 2008), are used to provide bioenergy in the form of heat, electricity and liquid fuels (Cherubini, 2010). There are numerous other chemicals with potential commercial value present in grasses (Heaton, et al., 2008). *Miscanthus* is a good source of biomass for biorefining, owing to its high lignocellulose yield, producing nearly three times more harvestable biomass at cool temperatures (3×10^4 Kg hectare^{-1}) than does *P. virgatum* (1×10^4 Kg hectare^{-1}) (Heaton, et al., 2008). In addition, *Miscanthus* is able to tolerate marginal lands and flooding (Heaton, et al., 2008; Lewandowski et al., 2000). Biomass from *Miscanthus* has a high lignin content, comprising around 24-25% of the cell wall, making it closer to a woody material than to a grass (Lygin et al., 2011; Villaverde et

al., 2010). Furthermore, hydroxycinnamic acids (specifically ferulic acid and *para*-coumaric acid) are covalently linked into the cell wall of grass species, where they serve to cross-link hemicellulose and lignin by ester and ether bonds, and their concentration is thought to be correlated with ease of hydrolysis of the biomass (Akin, 2007).

Plants produce a wide array of mono- and poly-phenols, with roles in strengthening cell walls, UV protection, stress tolerance and resistance to pathogens (Parveen et al., 2011, 2010). These include hydroxycinnamic acid conjugates and flavonoids from the phenylpropanoid and shikimate pathways and have antioxidant and UV-absorbing properties. The identification of commercially important molecules from crops to maximise value within a biorefinery has resulted in a demand for high-throughput screening tools for the extraction, identification and quantification of bio-chemicals. High performance liquid chromatography (HPLC) is the most widely used analytical technique for the characterisation of a wide range of polyphenolic compounds. LC-electrospray ionisation-tandem mass spectrometry (LC-ESI-MS/MS) has been used to characterise and quantify rapidly hydroxycinnamic acid conjugates and flavonoids in a wide range of plants (Parveen et al., 2011; Valls et al., 2009; Clifford et al., 2006).

Miscanthus is currently being investigated as a biofuel plant; however, little work has been carried out to identify high-value chemicals in this perennial grass. Mapping families are produced by crossing parents with contrasting phenotypes for a selected trait. The progeny are then screened for variation (segregation) in the trait which can then be associated with differences between the progeny in genomic sequence. The overall aim of this study was to screen a mapping family to identify commercially important chemicals which increase the value in the biorefinery chain and to identify

genotypes and growth stages which show highest abundance of these compounds. Examples are luteolin and its glycosides, which are reported to have anti-melanogenic activity (An et al., 2008), radical-scavenging activity, anti-inflammatory activity (Odontuya et al., 2005) and antibacterial activity (Dillard and German, 2000).

Characterisation of individual progeny is complicated by complex profiles of phenols and a high abundance of related structures and the objective was to determine the application of LC-PDA-ESI-MS/MS for detecting variation within the mapping family. The soluble phenol profiles in the leaf tissue of a biparental *Miscanthus sinensis* mapping family (two hundred progeny and both maternal and paternal plants) were investigated in the early development growth period (at one month's growth) and at their highest biomass at seven months' growth. We discuss the variation that exists for high-value chemicals that may be used to develop a genetic screen for such chemicals.

2. Results and Discussion

A *Miscanthus sinensis* (*M. sinensis*) mapping family, consisting of a bi-parental cross, was selected for study due to its high phenotypic variability. The maternal plant exhibited a stay-green trait, while the paternal plant produced high biomass and high numbers of seeds. LC-DAD-ESI-MS/MS, in negative ion mode, was used to screen two hundred progeny and maternal and paternal plants for soluble phenols in early developmental growth (one month seedlings). More than thirty hydroxycinnamates and flavonoids were identified and quantified in leaf tissue of progeny and both parents (Fig. 1). This method proved effective for distinguishing differences in profiles of phenols between progeny. The progeny were divided into categories (low, medium and high)

according to their content of hydroxycinnamates and flavonoids (Fig. 2). Eleven progeny exhibiting high, moderate and low concentrations of hydroxycinnamates and flavonoids and both parents were selected and screened at seven months' growth to determine the abundance of these soluble phenols at their peak of production of biomass and their levels were compared with their profiles of phenols when seedlings. Information on how the content of phenols within the leaf tissue differs between the early and late periods of growth may provide insight into how value may be increased within a biorefinery chain. Ratios of the hydroxycinnamates to flavonoids on a weight-per-weight basis were higher in seedlings (73:60) and lower in mature leaves (12:18). Phenols are known to be synthesised in the cytoplasm and chloroplasts and, as the plant matures, phenols accumulate in cell vacuoles or polymerise into lignin to strengthen the secondary cell walls (Kefeli et al., 2003).

The results showed that the progeny varied significantly in their concentrations of hydroxycinnamic acid conjugates, flavonoids and total phenols in seedlings and mature leaves (Fig. 3). In general, the levels of the phenols decreased as the leaves matured . This is not surprising, as seedlings contain only a primary cell wall and, as secondary cell wall thickens in mature tissue, soluble phenols are incorporated into lignin (Kefeli et al., 2003). This is also in agreement with Martin, who investigated the effect of stage of maturity of perennial ryegrass on phenolic compounds and found that increasing maturity of the herbage was linearly correlated to the decrease in the proportion of soluble phenols in leaf tissue and increase in the lignin content of the herbage (Martin, 1970).

A significant decrease was observed in the concentration of total phenols from one month to seven months' growth, with the most abundant sample decreasing from 8.0 mg / g FW to 0.1 mg/g FW. Plants synthesise phenols as a protective mechanism against herbivores, microbial pathogens, invertebrate pests and hostile environmental stresses (Parveen et al., 2010). To date, studies on the effect of maturation of the leaf on the concentrations of polyphenols are scarce. Although there was a decrease in the total content of phenols of the paternal plant from 5.3 mg / g FW to 2.1 mg / g FW and a reduction in the maternal plant from 1.6 mg / g FW to 1.1 mg/g FW, this may suggest that differences in the total contents of phenols of parents are less marked after seven months' growth, in comparison to the seedling stage, and may indicate that soluble plant phenols may be rapidly lost in the first few weeks of development. Interestingly, *ca*. twenty-five percent of the progeny had a profile of phenols similar to that of the paternal parent, while seventy-five percent showed a profile similar to that of the maternal parent. This population represents the F1 progeny of a biparental cross and the 3:1 ratio may indicate that the chemical composition is inherited in a classical Mendelian fashion and that the paternal phenotype is linked with recessive alleles for genes involved in the expression of this trait. If the chemical phenotype were also associated with simple Mendelian segregation of markers, it would allow for the robust association of genetic markers with chemical phenotypes and for the rapid selection of chemical profiles using marker-assisted selection. With plant phenols being of such high economic interest, this could be investigated further within the *M. sinensis* mapping family to ensure that maximum value within the biorefinery process is achieved. On the other hand, plants from the lower end of the spectrum may be favoured for fermentation (Akin, 2007). It would be worth characterising the individual phenols to assess their

potential economic value, as this could add value to the biorefinery chain and furthermore may facilitate conversion into other industrially useful products.

Ten hydroxycinnamates were identified with UV absorptions typically in the region 240-340 nm, while twenty-six flavones were identified by their characteristic UV absorption profiles in the region 240-380 nm. The UV absorbance maxima could not be recorded for some compounds, either because of low abundance or because they were masked by co-eluting peaks. The identification method employed used UV absorption profiles coupled with the fragmentation pattern observed in tandem mass spectra using LC-ESI-MS^n. Where possible, the mass spectra and retention times of the phenols were compared with those of standards including caffeic acid, *para*-coumaric acid, ferulic acid, 5-caffeoylquinic acid (5-CQA, also known as chlorogenic acid (the IUPAC isomeric numbering convention (IUPAC Commissions on Nomenclature, 1976; Clifford et al., 2003) is used for acylated quinic acids throughout this paper)), luteolin-6-*C*-glucoside (isoorientin), luteolin-8-*C*-glucoside (orientin), luteolin-7-*O*-glucoside, apigenin-6-*C*-glucoside (isovitexin), apigenin-8-*C*-glucoside (isovitexin), apigenin-7-*O*-glucoside.

2.1 Hydroxycinnamates

A number of hydroxycinnamic acid conjugates, including *O*-caffeoylquinic acids, 3-caffeoylquinic acid (3-CQA; **1**), 4-caffeoylquinic acid (4-CQA; **4**) and 5-caffeoylquinic acid (5-CQA; **3**), *O*-feruloylquinic acids, 3-feruloylquinic acid (3-FQA; **2**), 4-feruloylquinic acid (4-FQA; **11**) and 5-feruloylquinic acid (5-FQA; **10**) and *O-para*-coumaroylquinic acids, 5-*para*-coumaroylquinic acid (5-*para*-CoQA; **7**) were identified in the leaf extracts of both parents and progeny (Table 1).

Diagnostic fragmentation ions of caffeoylquinic acids in negative-ion mode ESI-MS^n involved one of two pathways: loss of the acyl group – cleavage of the carbonyl-oxygen bond (which gives quinate *m/z* 191) and β-elimination of a carboxylic acid (which gives a dehydrated quinate *m/z* 173 and caffeate *m/z* 179 anions) (Parveen et al., 2011). For 5-CQA, an MS^2 experiment at *m/z* 353 yielded a base ion, *m/z* 191, while the predominant fragmentation of 3-CQA gave an MS^2 base ion *m/z* 191 and an intense caffeic acid-derived ion *m/z* 179 (Table 1).4-CQA was characteristic by β–elimination of caffeic acid to give MS^2 base ion *m/z* 173. 3-FQA was distinguished by its MS^2 base ion *m/z* 193 $[ferulate]^-$, while 4-FQA was discriminated by the abundant peak *m/z* 173 [M – ferulic acid]$^-$ from β-elimination and 5-CQA gave the ion *m/z* 191 $[quinate]^-$. 5-*para*-CoQA was characterized by the MS^2 base ion *m/z* 191. Free *para*-coumaric acid (**5**) was found to be relatively abundant in paternal leaf tissue. Of the hydroxycinnamates, 5-CQA was the most abundant in seedlings and mature leaves. Caffeoylquinic acids have been reported in a range of monocots (Parveen et al., 2010). Geometrical isomers (*cis* / *trans*) were evident based on photoirradiation experiments resulting in photoisomerisation (Parveen et al., 2011).

2.2 Mono-*O*-Glycosyl Flavones

Distinction of flavone-*C*-glycosides and flavone-*O*-glycosides is made by a simple hydrolysis experiment, where, in the former case, the carbon-carbon bond is resistant to cleavage and, therefore, the characteristic fragmentations are due to fission of bonds in the carbohydrate ring. In contrast, *O*-glycosides are easily distinguished due to cleavage of the interglycosidic bond, resulting in the loss of a neutral sugar molecule. Fragmentation of the molecular ion *m/z* 447 $[M-H]^-$ in negative-ion mode ESI-MS/MS

yielded MS^2 base ion *m/z* 285 Y_0^- $[(M - H) - 162]^-$, consistent with a loss of a dehydrated *O*-hexose moiety through heterolytic cleavage of the acetal *O*–*C* bond (Vukics and Guttman, 2010). A less abundant ion *m/z* 327 $[(M - H) - 120; AGly + 41]^-$ indicated internal fragmentation of the sugar moiety. The pattern of fragmentation observed for compound **20** by MS^2 and order of elution following separation by C_{18} reverse-phase chromatography under the above conditions were consistent with an authentic standard of luteolin-7-*O*-gucoside.

2.3 Mono-*C*-Glycosyl Flavones

In mono-*C*-hexosyl flavones, the presence of ions $-^{0,3}X_0$ $[(M - H) - 90]^-$, and $-^{0,2}X_0$ $[(M - H) - 120]^-$ were consistent with a *C*-linked hexose moiety. Moreover, ions [AGly + 41] and [AGly + 71] are characteristic of mono-*C*-glycosyl flavones. Plant compounds **16** and **19** showed MS^2 fragments $[(M - H) - 90]^-$ and $[(M - H) - 120]^-$ in negative-ion mode. Previous studies have reported that a highly abundant ion $[(M - H) - 120]^-$ was consistent with the presence of a hexose on *C*-8 of the aglycone, while the ion $[(M - H) - 90]^-$ was typical of the 6-*C* isomer (Ferreres et al., 2008, 2007, 2003). If this were the case, compound **16** showing base ion *m/z* 327 $[(M – H - 120]^-$ would indicate a 8-*C*-isomer (Pereira et al., 2005; Ferreres et al., 2003; Sánches-Rabaneda et al., 2003). However, the presence of an intense ion at *m/z* 429, corresponding to a loss of a water molecule, was consistent with an authentic standard of luteolin-6-*C*-glucoside. Thus compound **16** was confirmed as luteolin-6-*C*-glucoside. Similarly, **19** showed an intense base ion corresponding to the *C*-8 isomer; however, comparison of the fragmentation data and retention time under the conditions described above were consistent with commercial apigenin-6-*C*-glucoside. These inconsistencies may be attributed to

different experimental, instrumentation and fragmentation conditions, such as ionisation energies (Vukics and Guttman, 2010).

Mono-*C*-pentosyl flavone fragmentations: $-^{0,3}X_0$ [(M - H) - 60]$^-$ and $-^{0,2}X_0$ [(M - H) - 90]$^-$ were characteristic of the pentose residue. In addition, ions AGly + 41 and AGly + 71 were also observed. Plant components **24**, **34** and **35** showed abundant ions [(M - H)-60]$^-$ and [(M - H) - 90]$^-$, consistent with a *C*-linked pentose. A previous study reported that an intense ion [(M - H) - 60]$^-$ was typical for the 6-*C* isomer, while a strong ion [(M - H) - 90]$^-$ was associated with the 8-*C* isomer (Ferreres et al., 2003). However, as we observed discrepancies in the relative intensities of MS^2/MS^3 ions for compounds that are reported in the literature, we would not show where the linkage is located, unless this is confirmed by authentic standards or 2D NMR. Based upon this, **24** was tentatively identified as luteolin-*C*-pentoside and **34** and **35** as apigenin-*C*-pentosides. The latter two compounds are likely to be stereoisomers. Similarly, mono-*C*-deoxyhexosyl flavones were characterized by fragmentations: $-^{0,3}X_0$ [(M - H) - 74]$^-$ and $-^{0,2}X_0$ [(M - H) - 104]$^-$ (Vukics and Guttman, 2010). Compound **29** was identified as luteolin-*C*-deoxyhexoside and **36** as apigenin-*C*-deoxyhexoside.

2.4 Di-*C*-Glycosyl Flavones

In di-*C*-hexosyl flavones, the main MS^2 fragmentations are $-^{0,2}X_0$ [(M - H) - 120]$^-$, $-^{0,3}X_0$ [(M - H) - 90]$^-$, [(M - H) - 210; AGly + 113]$^-$ and [(M - H) - 240; AGly + 83]$^-$. Compound **13** showed fragmentations characteristic with this group of compounds and yielded a base ion *m/z* 473, showing a loss of 120 Da. Compound **13** was tentatively identified as apigenin-6,8-di-*C*-hexoside. The fragmentation pathway and relative intensity of MS^2 ions were consistent with apigenin-6,8-di-*C*-glucoside previously

identified in polar extracts of *Siparuna guianensis* and in *Cydonia oblonga* (Negri et al., 2012; Ferreres et al., 2003). By contrast, di-*C*-substituted flavones with two different sugars (such as a pentose and a hexose) linked to the aglycone gave characteristic ions $[(M - H) - 90]^-$, $[(M - H) - 120]^-$, $[(M - H) - 60]^-$ and $[(M - H) - H_2O]^-$. The position of glycosylation has previously been predicted for asymmetric di-*C*-glycosyl flavones (Ferreres et al., 2008, 2003). The higher abundance of the ion $[(M - H) - 90]^-$ indicated that the pentose moiety was *C*-linked at position-6 of the aglycone, while a more intense ion $[(M - H) - 120]^-$ indicated linkage at position-8. Compound **9** showed losses that were consistent with this group of compounds and was tentatively identified as a luteolin-*C*-pentoside-*C*-hexoside. Fragmentation and intensities of ions were consistent with luteolin-6-*C*-arabinoside-8-*C*-glucoside previously identified in leaf tissue of *Hordeum vulgare* L (Ferreres et al., 2008, 2003).

2.5 *O*-Glycosyl-*C*-Glycosyl Flavones

In negative-ion mode ESI-MS/MS, compound **14** yielded base ion $-Y_0^-$ $[(M - H) - 162]^-$ and relatively abundant ions $[(M - H) - 90]^-$ and $[(M - H) - 252; AGly + 41]^-$. Ions *m/z* 417 and 399, showing losses of 162 and 180, were consistent with an *O*-linked hexose moiety. Fragment ions *m/z* 357, 327, AGly + 71 and AGly + 41 were typical of mono-*C*-pentosyl flavones. Compound **14** was tentatively identified as a luteolin-*O*-hexoside-*C*-pentoside. Previous studies have reported on the presence of a *C*-hexosyl-*C*-pentosyl-luteolin and a *C*-heoxosyl-luteolin-*O*-pentoside in *Triticum durum* species; however, compound **14** has not been reported (Cavaliere et al., 2005).

LC-ESI-MS/MS of molecular ion *m/z* 593 (**18**) gave base ion *m/z* 341 characteristic of AGly+41 and an intense ion *m/z* 431 which shows loss of an *O*-hexose. Other MS^2 ions

which were relatively abundant were *m/z* 503 (-90), 533 (-60) and 473 (-30) consistent with a *C*-linked pentose. Furthermore, the presence of an intense ion *m/z* 578 (loss of 15 Da) indicates an *O*-methylated flavone. These data suggest the aglycone to be either chrysoeriol or diosmetin (luteolin-4'-methyl ether). Compound **18** was tentatively identified as chrysoeriol / diosmetin-*O*-hexoside-*C*-pentoside. *C*-Hexosyl-chrysoeriol-*O*-pentoside and *C*-hexosyl-chrysoeriol-*O*-hexoside have previously been reported in *Triticum durum* plants (Cavaliere et al., 2005). Compounds **23** and **25** showed losses from the *pseudo*molecular ion *m/z* 577 $[M - H]^-$, 104 and 162 (base peak), consistent with a *C*-deoxyhexose and an *O*-hexose. These were tentatively identified as apigenin-*O*-hexoside-*C*-deoxyhexosides. These are likely to be regioisomers.

2.6 2"-*O*-Glycosyl-*C*-Glycosyl-Flavones

A large number of plant components **15**, **17**, **21**, **22**, **26**, **27**, **32** and **33** yielded abundant ions; in many cases, these were base ion Z_1^- $[Y_1^{6-} - 18]^-$ and were consistent with *O*-glycosylation occurring at position-2" of the sugar, as opposed to *C*-glycosylation. Two sugars can be attached to the aglycone, either at two different positions to give (a) di-*O*-glycosides and (b) di-*C*-*O*-glycosides, or at the same position to furnish (a) *O*-diglycosides and (b) *O*,*C*-diglycosides. All the above compounds were identified as *O*,*C*-diglycosides based upon the presence of the highly abundant Z_i fragment and a low-intensity Y_i fragment, indicating a disaccharide residue (Vukics and Guttman, 2010). In compounds **22** and **27**, an intense ion $[(M - H) - 150]^-$ suggested a terminal *O*-pentose and $[(M - H) - 164]^-$ in compounds **15**, **17**, **21**, **26** and **32** (possibly also **31** and **33**) indicated a terminal *O*-deoxyhexosose. In addition, characteristic ions AGly + 71 - 18 and AGly + 41 - 18 were typical in this group of compounds. Furthermore, the

presence of ions corresponding to the loss of $^{0,2}X_0^-$ [(M - H) - 120]$^-$ (for hexoses), [(M – H) - 104]$^-$ (for deoxyhexoses) and [(M - H) - 90]$^-$ (for pentoses) and MS^3 fragment ions AGly + 41 / AGly + 41 - 18 were considered indicative of 2”-*O*-glycosyl-*C*-glycosyl flavones (Ferreres et al., 2008, 2007). Compounds **15** and **17** are stereoisomers and were characterised as 2”-*O*-deoxyhexoside-*C*-hexosides of luteolin. Compound **21** was identified as a 2”-*O*-deoxyhexoside-*C*-hexoside of apigenin. The fragmentations and relative intensity of ions for the latter compound were consistent with 2”-*O*-rhamnoside-*C*-hexoside apigenin reported in several plant species (Ferreres et al., 2007).

Compounds **22** and **27** yielded MS^2 base ions [(M - H) - 150]$^-$ typical of a loss of an *O*-glycosylated pentose unit, while MS^2 ions *m/z* [(M - H) - 90]$^-$ and [(M - H) - 60]$^-$ and AGly + 41 - 18 were typical of a *C*-pentose. Thus **16** was identified as 2”-*O*-pentoside-*C*-pentoside luteolin and **27** as 2”-*O*-pentoside-*C*-pentoside apigenin. By contrast, compound **26** showed an abundant peak *m/z* 399 [(M - H) - 164]$^-$ consistent with a loss of an *O*-glycosylated deoxyhexose, while MS^2 fragment ions *m/z* 473 [(M - H) - 90]$^-$ were consistent with a *C*-pentose. Compound **20** was provisionally identified as 2”-*O*-deoxyhexoside-*C*-pentoside luteolin. On the other hand, **32** showed losses of 164 Da and ions AGly + 71 and AGly + 71 – 18, consistent with 2”-*O*-deoxyhexoside-*C*-deoxyhexoside luteolin.

2.7 Unidentified derivatives

Compound **28** gave a UV spectrum consistent with the class of flavones but could not be identified. It yielded an MS^2 base peak at *m/z* 385 [(M - H) - 44]$^-$ which is consistent with the loss of CO_2. This suggests the presence of at least one carboxylic acid in the molecule. Intense ions *m/z* 411 [(M - H) - H_2O]$^-$, 357 [(M - H) - 72; AGly + 71]$^-$, and

325 [(M - H) - 104]$^-$ were also observed. These are typical ions for a *C*-linked deoxyhexose moiety. An ion of low abundance at *m/z* 285 may correspond to a deprotonated luteolin. A *pseudo*molecular ion at *m/z* 413 [M - H]$^-$ (**31**) gave MS^2 ions 339 [(M - H) - 74]$^-$consistent with a *C*-linked deoxyhexoside. We were unable to characterise fully both compounds based upon their fragmentation data.

Two further compounds (**31** and **33**) with *pseudo*molecular ion *m/z* 593 [M - H]$^-$, yielded base ion *m/z* 575 and a weak ion *m/z* 411, corresponding to loss of an *O*-linked deoxyhexose unit in a disaccharide from the dehydrated pseudomolecular ion. MS^3 fragmentation of *m/z* 575 gave base ion *m/z* 411 and ions *m/z* 531 [(M - H) - 44, - 18]$^-$, 513 [(M - H) - 62, - 18]$^-$, 429 [(M - H) -164]$^-$, 367 [(M - H) - 164, - 44, - 18]$^-$ and 337 [(M - H) - 164, - 74, - 18]$^-$. We were unable to elucidate the structures of these compounds on the basis of their fragmentation patterns. In addition, a molecular ion at *m/z* 575 [M - H]$^-$ (**30**) gave MS^2 base ion *m/z* 411 [(M - H) - 164]$^-$ and secondary ions identical to those observed in MS^3 for **31** and **33**, indicating that **30** is a related compound.

2.8 Quantification

The abundances of selected compounds in the various extracts were by measurement of peak area in the HPLC chromatogram at 340 nm and reference to the extinction coefficient of chlorogenic acid and luteolin. Data are presented in Table 2. 5-Caffeoylquinic acid was the most abundant hydroxycinnamate in all tissues examined; this was also true of the hybrid *M.* × *giganteus* (Parveen et al., 2011) and the parents *M. sinensis* and *M. sacchariflorus.* However, there were marked differences in the abundance of 5-CQA (**3**) between the maternal and paternal plants. For example, **3** was

present to only 0.87 μmol g^{-1} in the maternal parent plant, whereas it was more than twice as abundant in the paternal plant (2.15 μmol g^{-1}). In addition, *para*-coumaric acid (**4**) was present in high concentrations (0.38 μmol g^{-1}) in paternal leaf tissue when compared with the maternal parent (0.16 μmol g^{-1}).

Several flavones **16, 17, 26, 29, 30** were present at relatively high concentrations in leaf tissues of parents and progeny. However, there were differences in abundance of Lut-6-*C*-Glu (**16**) and 2"-Deoxyhex-*C*-Hex-Lut (**17**) between the maternal and paternal plants; **16** was present to only 0.18 μmol g^{-1} and **17** to 0.19 μmol g^{-1} in the paternal parent, whereas these compounds were present at higher concentrations in the maternal plant (**16** - 0.69 μmol g^{-1}; **17** - 0.65 μmol g^{-1}). Unknown compound **30** was present in relatively high levels in both maternal (0.38 μmol g^{-1}) and paternal (0.53 μmol g^{-1}) parent plants.

Plants produce a wide range of secondary compounds that have potential valuable/high value activities and applications in a range of industries (such as pharmaceutical, nutraceutical, food and cosmetic). Current concerns with the negative environmental impact associated with production of chemicals is driving policy towards increasing the supply and diversity of phytochemicals available for a range of applications. *Miscanthus* is an alternative and reliable source of phenols, which can be used as an alternative in the pharmaceutical, nutraceutical, cosmetic, food and beverage industries (Ghasemzadeh, 2011; Parveen et al., 2010). Flavonoids are the largest class of polyphenols, with over 6000 compounds showing a wide range of biological activities (Ghasemzadeh, 2011; Garcia-Lafuente et al., 2009). Consequently, they have important economic benefit; in particular, luteolin and apigenin compounds have been widely used

in the food and cosmetic industry (An et al., 2008; Mak et al., 2006). The most abundant flavones in the *M. sinensis* mapping family were indeed derivatives of luteolin and apigenin. Luteolin has been shown to inhibit cellular melanogenesis in murine melanoma B16F10 cells activated by α-melanocyte stimulating hormone (An et al., 2008). Luteolin also inhibits tyrosinase-catalysed oxidation of L-dihydroxyphenylalanine in cell-free extracts and living cells and softens/reduces cell pigmentation induced by expression of exogenous human tyrosinase (An et al., 2008). The anti-melanogenic effects of luteolin were attributed to inhibition of the activity of the tyrosinase enzyme (An et al., 2008). The three glycosides, luteolin-6-*C*-glucoside, luteolin-7-*O*-glucoside and luteolin-6-*C*-arabinoside, above 200 μg mL^{-1} have been shown to protect the HSC-2 carcinoma cell line from UV-induced cytotoxicity (Matsuta et al., 2011). In addition, all three glycosides showed potent antioxidant and radical scavenging activities. A previous study assessed the effect of luteolin and its glycosides, luteolin-7-*O*-glucoside, luteolin-6-*C*-glucoside and luteolin-6-*O*-glucoside in the hydrogen peroxide scavenging activity and their inhibitory effect on the enzymes involved in the arachidonic acid pathway which generate thromboxane B_2 (TXB_2) and leukotriene B_4 (LTB_4) (Odontuya et al., 2005). Both of these have attracted much attention, as TXB_2 is considered to play an important role in diseases such as cirrhosis of the liver, systemic lupus erythematosis and thrombosis, while LTB_4 is involved in asthma, psoriasis, gout and inflammatory bowel diseases (Odontuya et al., 2005). Luteolin was reported to possess good inhibition against synthesis of LTB_4, luteolin-7-*O*-glucoside showed relatively good inhibition production of both TXB_2 and LTB_4, while luteolin-6-*C*-glucoside was reported to have potent selective inhibition against

biosynthesis of TXB_2 (Odontuya et al., 2005). Luteolin has also been reported to exhibit anti-inflammatory and anti-bacterial properties (Dillard and German, 2000).

Apigenin exerts its antiproliferative effects on prostate (DU-145) and breast cancer (MDA-MB-231) cells through oestrogen receptor-β (Mak et al., 2006). This compound has been shown to have anti-proliferative effects and pro-apoptotic activities through caspase-3 activation in the two cell lines (Mak et al., 2006). It inhibits the production of TXB_2, while both luteolin- and apigenin-7-O-glycosides showed potent inhibition of biosynthesis of 12-hydroxy-5,8,10,14-eicosatetraenoic acid (12-HETE) (Zheng et al., 2005). The authors report that such compounds may be useful in treating ailments such as thrombosis and atherosclerosis (Zheng et al., 2005). Furthermore, apigenin has been shown to suppress 12-O-tetradecanoylphorbol-1,3-acetate (TPA)-mediated tumour promotion in mouse skin and has antibacterial, anti-inflammatory and diuretic properties (Dillard and German, 2000). Recently, nanosuspensions of apigenin have been shown to enhance antioxidant activity significantly and may have potential in anti-ageing products or pharmaceutical creams for protection against skin cancer (Arct and Pytkowska, 2008; Shaal et al., 2011). Flavone glycosides with identical molecular mass but different positional isomers have different chemical and biological properties. Compounds shown in Table 1 without references are novel, to the best of our knowledge. Much novelty remains in *Miscanthus* species and opens up the possibility of new developments in Intellectual Property and processes. Figure 4 shows structures of some novel flavonoids from mature leaf tissue.

3. Materials and methods

3.1 Plant materials

Miscanthus seeds were generated from a cross between two M. sinensis varieties, one a high-yielding and good seed producer (paternal plant), the other a plant of similar overall morphology but that exhibits a stay-green phenotype (maternal plant). The identity of the parental plants as *M. sinensis* species was authenticated by Dr. Steve Renvoize, Kew Botanic Gardens. Seeds were sown in December 2010 in a controlled environment in a 50:50 mixture of damp soil and vermiculite at the Institute of Biological, Environmental and Rural Sciences (IBERS), Aberystwyth. The resulting F1 progeny were similar in overall development and morphology but displayed considerable variation specifically in leaf size and leaf chlorophyll content. This population was judged to be suitable for determining variation in chemical composition since potential confounding factors such as variation in developmental timing were absent whereas large differences in leaf structure and chlorophyll composition it was hypothesised may be linked to large differences in composition. Analysis part 1: On 1 March 2011 (at one-month growth), the youngest fully grown leaf of each plant (200 progeny and maternal and paternal plants) was removed and placed in an Eppendorf tube, weighed and frozen in liquid nitrogen. Samples were stored at -80°C prior to extraction of phenols, purification and analysis. Analysis part 2: Leaves of selected progeny with high, low and average contents of hydroxycinnamates and flavonoids (at one-month growth) were harvested at seven months' growth in October 2011. As previously, plant material was placed in an Eppendorf tube, weighed and frozen in liquid nitrogen. Samples were stored at -80 ºC prior to extraction and assay.

3.2 Solid-phase extraction

Free phenols were extracted by adding *M. sinensis* leaf tissue to aqueous methanol (75%) at 70ºC for 5 min. The mixture was then transferred to a mortar and ground with a pestle to extract the phenols. The suspension was poured into a vial and centrifuged for 10 min at approximately 3000 rpm. The solvent was evaporated from the supernatant solution by vacuum centrifugation for 180 min with gentle heating applied. A Waters Sep-Pak (500 mg) C_{18} reverse-phase cartridge was used to semi-purify the crude mixture. The C_{18} reverse-phase cartridge was initially cleaned with methanol (4.0 mL) and washed several times with purified water (containing 5% acetic acid) (4.0 mL). After the sample was loaded, the column was washed with water (95%)-acetic acid (5%) to remove any unbound compounds before the target compounds were eluted with methanol (4.0 mL). The solvent was evaporated in a vacuum centrifuge for approximately 180 min and the residue was subsequently re-suspended in methanol (100 μL) and centrifuged at 14500 rpm for 2 min. The supernatant solution was transferred into to a vial for analysis by HPLC.

3.3 HPLC-PDA

Analysis by HPLC was carried out as reported previously (Parveen et al., 2011, 2008). A Waters HPLC system (Waters Corporation, USA) with an auto-sampler, a photodiode array detector and an analytical workstation was used. The column configuration consisted of a Waters C_{18} reverse-phase Nova-Pak column (4 μm, 8 mm × 100 mm). The sample vial was loaded into the auto-sampler and maintained at a constant temperature of 4ºC. Injection volumes were either 20 μL or 10 μL. The mobile phase consisted of purified water : acetic acid (A; 95:5, v/v) and HPLC-grade methanol (B;

100%). Initial conditions were a) A:B (95:5, v/v) with a linear increase in B to 75% in 70 min. The data were analysed using Empower software (version 2002, Waters Corporation, USA).

3.4 LC-PDA-ESI-MS/MS

HPLC/MS^n analysis was performed as described previously (Parveen et al., 2011) using a Thermo Finnigan LCMS System (Thermo Electron Corporation, USA). The system comprised a Finnigan Surveyor PDA Plus detector and a Finnigan LTQ (linear trap quadrupole) with an ESI source. Chromatography was performed on a Waters C_{18} reverse-phase Nova-Pak column. The column temperature was constant at 30ºC and the temperature of the auto-sampler tray was maintained at 5ºC. The sample injection volume was 10 µL, the detection wavelength was set at 240-400 nm with a flow rate of 1 mL min^{-1}. Purified water-formic acid (A; 100 : 0.1 v/v) and HPLC-grade methanol-formic acid (B; 100 : 0.1) together formed the mobile phases. The initial condition was A:B (95 : 5 v/v) and the percentage of B increased linearly over time from 5 to 75% in 70 min. Interface and MD parameters were as follows; sheath gas 30 arbitrary units, auxiliary gas 15 units, spray voltage -4 KV, capillary temperature 320ºC, capillary voltage -1.0 V and tube lens offset -68 V.

3.5 Standards

Authentic standards (purity >98.0 %) of lutoeolin-6-*C*-glucoside, luteolin-8-*C*-glucoside, apigenin-8-*C*-glucoside, apigenin-6-*C*-glucoside, apigenin-7-*O*-glucoside, luteolin-7-*O*-glucoside, 5-caffeoylquinic acid and *para*-coumaric acid were purchased from Sigma Aldrich, UK.

3.6 Quantification

Quantification was performed using response factors. The hydroxycinnamates were quantified using 5-caffeoylquinic acid (8.74×10^{-7} μg / area unit) and the flavonoids with the response factor for luteolin (5.32×10^{-7} μg / area unit). These values were used to estimate the amount of compound in fresh plant material (mg / g FW). These compounds were estimated on a fresh-weight basis to make values comparable between growth stage and different genotypes. The fibre content in leaves varied with genotype and increased with growth stage and therefor makes a highly variable contribution to the dry matter. Studies have demonstrated that water content in perennial grass species is best explained by the proportion of mesophyll protoplast (Garnier and Laurent, 1994). Therefore, expressing phenol content on a fresh weight basis gives a closer estimation of the concentration of phenols within cells which we consider is a more accurate comparison of cellular metabolism between plants and between the different developmental stages.

4. Conclusions

This study describes the application of HPLC-PDA and LC-PDA-ESI-MS/MS as a rapid screening tool for the qualitative and quantitative determination of soluble phenols in a *M. sinensis* mapping family. This technique was sufficiently sensitive to detect differences in phenol profiles between closely related progeny. This is the first study to report on the content of flavonoids in a *M. sinensis* family. More than thirty hydroxycinnamates and flavonoids were identified and quantified in leaf tissue of two hundred progeny and maternal and paternal plants in seedlings and in mature leaves (at seven months growth). The concentration of total phenols in this mapping family varied

between 0.53 and 12.4 mg / g FW. The paternal plant contained a higher total content of phenols when compared to the maternal plant. The concentrations of phenols generally decreased as leaves matured. The phenolic profile patterns of seedlings were not correlated with the patterns observed at seven months' growth. Nevertheless, several genotypes were identified as having potential to add value to the biorefinery by providing functional soluble phenol co-products. These may have application as antioxidants, UV screens or inflammatory inhibitors in cosmetic and pharmaceutical products. Alternatively, plants with low levels of anti-microbial phenols may be good feedstocks for fermentation. Flavones can also be used as markers of the parental lines in the *sinensis* mapping family. Furthermore, several novel flavone-glycosides have been observed. LC-ESI-MS^n has proven to be a powerful technique to screen plants in a mapping family for profiles of soluble phenols and to select plants that are potential sources of high-value chemicals. This study provides information on the content of potentially valuable compounds from readily renewable resources and possible biomarkers for identification in breeding programmes and development of varieties with potential for biorefining.

5. Acknowledgements

Work reported in this manuscript was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme Grant on Energy Grasses & Biorefining (BB/J0042/1). We gratefully acknowledge support for BEACON from the European Regional Development Fund through the Welsh European Funding Office, part of the Welsh Assembly. We are also grateful to Sue Youell, Dr. Charlotte Hayes and Jo Spikes for useful discussions.

References

Aburjai, T., Natsheh, F.M., 2003. Plants used in cosmetics. Phytother. Res. 17, 987-1000.

Akin D.E., 2007. Grass lignocellulose: Strategies to overcome recalcitrance. Appl. Biochem. Biotechnol. 137, 3-15.

An, S.M., Kim, H.J., Kim, J-E., Boo, Y.C., 2008. Flavonoids, taxifolin and luteolin attenuate cellular melanogenesis despite increasing tyrosinase protein levels. Phytother. Res. 22, 1200-1207.

Arct, J., Pytkowska, K., 2008. Flavonoids as components of biologically active cosmeceuticals. Clin. Dermatol. 26, 247-357.

Cherubini, F., 2010. The biorefinery concept: Using biomass instead of oil for producing energy and chemicals. Energy Convers. Manage. 51, 1412-1421.

Cavaliere, C., Foglia, P., Pastorini, E., Samperi, R., Lagana, A., 2005. Identification and mass spectrometric characterization of glycosylated flavonoids in *Triticum durum* plants by high-performance liquid chromatography with tandem mass spectrometry. Rapid Commun. Mass Spectrom 19, 3143-3158.

Clifford, M.N., Johnston, K. L., Knight, S., Kuhnert, N., 2003. Hierarchical scheme for LC-MS^n identification of chlorogenic acids. J. Agric. Food Chem. 51, 2900-2911.

Clifford, M.N., Knight, S., Surucu, B., Kuhnert, N., 2006. Characterisation by LC-MS^n of four new classes of chlorogenic acids in green coffee beans: dimethoxycinnamoylquinic acids, diferuloylquinic acids, caffeoyl-dimethoxycinnamoylquinic acids, and feruloyl-dimethoxycinnamoylquinic acids. J. Agric. Food Chem. 54, 1957-1969.

Crozier, A., Jaganath, I. B., Clifford, M.N., 2009. Dietary phenolics: chemistry, bioavailability and effects on health. Nat. Prod. Rep. 26, 1001-1043.

Dillard, C.J., German, J.B., 2000. Phytochemicals: nutraceuticals and human health. J. Sci. Food Agric. 80, 1744-1756.

Dong, A., Fang, Y. , Zhang, L., Xie, J., Wu, X., Zhang, L., Lian, X., Chen, Y., Luo, J., Liu, M., 2011. Caffeic acid 3,4-dihydroxy-phenethyl ester induces cancer cell senescence by suppressing twist expression. J. Pharmacol. Exp. Ther. 339, 238-247.

Ferreres, F., Silva, B.M., Andrade, P.B., Seabra, R.M., Ferreira, A., 2003. Approach to the study of C-glycosyl flavones by ion trap HPLC-PAD-ESI/MS/MS: application to seeds of quince (*Cydonia oblonga*). Phytochem. Anal. 14, 352-359.

Ferreres, F., Gil-Izquierdo, A., Andrade, P.B., Valentão, P., Tomás-Barberán, F-A., 2007. Characterization of *C*-glycosyl flavones *O*-glycosylated by LC-MS/MS. J. Chromatogr. A. 1161, 214-223.

Ferreres, F., Andrade, P.B., Valentão, P., Gil-Izquierdo, A., 2008. Further knowledge on barley (*Hordeum vulgare L.*) leaves *O*-glycosyl-*C*-glycosyl flavones by liquid chromatography-UV diode-array detection-electrospray ionisation mass spectrometry. J. Chromatogr. A. 1182, 56-64.

García-Lafuente, A., Guillamón, E., Villares, A., Rostagno, A., Martínez, J.A., 2009. Flavonoids as anti-inflammatory agents: implications in cancer and cardiovascular disease. Inflamm. Res. 58, 537-552.

Garnier, E., Laurent, G., 1994. Leaf anatomy, specific mass and water content in congeneric annual and perennial grass species. *New Phytol.* 128, 725-736.

Ghasemzadeh, A., Ghasemzadeh, N., 2011. Flavonoids and phenolic acids: Role and biochemical activity in plants and human. J. Med. Plants Res. 5, 6697-6703.

Heaton, E.A., Dohleman, F.G., Long, S.P., 2008. Meeting US biofuel goals with less land: the potential of *Miscanthus*. Glob. Change Biol. 14, 2000-2014.

IUPAC Commission on the Nomenclature of Organic Chemistry (CNOC) and the IUPAC-IUB Commission on Biochemical Nomenclature, 1976. Nomenclature of cyclitols. Biochem. J., 153, 23-31.

Kefeli, V.I., Kalevitch, M.V., Borsari, B., 2003. Phenolic cycle in plants and environment. J. Cell Mol. Biol. 2, 13-18.

Kleinert, M., Barth, T., 2008. Phenols from lignin. Chem. Eng. Technol. 31, 736-745.

Lewandowski, I., Clifton-Brown, J.C., Scurlock, J.M.O., Huisman, W., 2000. Miscanthus: European experience with a novel energy crop. Biomass Bioenergy 19, 209-227.

Lygin, A.V., Upton, J., Dohleman, F.G., Juvik, J., Zabotina, O.A., Widholm, J.M., Lozovaya, V.V., 2011. Composition of cell wall phenolics and polysaccharides of the potential bioenergy crop – *Miscanthus*. Global Change Biol. Bioenergy 3, 333-345.

Mak, P., Leung, Y-K., Tang, W-Y., Harwood, C., Hoy, S-M., 2006. Apigenin suppresses cancer cell growth through ERβ. Neoplasia 8, 896-904.

Martin, A., 1970. Effect of stage of maturity of perennial ryegrass on its content of some organic acids and phenolic compounds. J. Sci. Food Agr. 21, 496-501.

Matsuta, T., Sakagami, H., Satoh, K., Kanamoto, T., Terakubo, S., Nakashima, H., Kitajima, M., Oizumi, H., Oizumi, T., 2011. Biological activity of luteolin glycosides and tricin from *Sasa senanensis* Rehder. In Vivo 25, 757-762.

Negri, G., de Santi, D., Tabach, R., 2012. Chemical composition of hydroethanolic extracts from *Siparuna guianensis*, medicinal plant used as anxiolytics in Amazon region. Rev. Bras. Farmacogn. 1024-1034. http://dx.doi.org/10.1590/S0102-695X2012005000034

Odontuya, G., Hoult, J.R.S., Houghton, P.J., 2005. Structure-activity relationship for anti-inflammatory effect of luteolin and its derived glycosides. Phytother. Res. 19, 782-786.

Parveen, I., Winters, A., Threadgill, M.D., Hauck, B., Morris, P., 2008. Extraction, structural characterization and evaluation of hydroxycinnamate esters of orchard grass (*Dactylis glomerata*) as substrates for polyphenol oxidase. Phytochem. 69, 2799-2806.

Parveen, I., Threadgill, M.D., Moorby, J.M., Winters, A., 2010. Oxidative phenols in forage crops containing polyphenol oxidase enzymes. J. Agric. Food Chem. 58, 1371-1382.

Parveen, I., Threadgill, M.D., Hauck, B., Donnison, I., Winters, A., 2011. Isolation, identification and quantitation of hydroxycinnamic acid conjugates, potential platform chemicals, in the leaves and stems of *Miscanthus* × *giganteus* using LC-ESI-MS^n Phytochem. 72, 2376-2384.

Pereira, C.A.M., Yariwake, J.H., McCullagh, M., 2005. Distinction of the *C*-glycosylflavone isomer pairs orientin/isoorientin and vitexin/isovitexin using HPLC-MS exact mass measurement and in-source CID. Phytochem. Anal. 16, 295-301.

Sánches-Rabaneda, F., Jáuregui, O., Casals, I., Andrés-Lacueva, C., Izquierdo-Pulido, M., Lamuela-Raventós, R.M., 2003. Liquid chromatographic/electrospray ionization tandem mass spectrometric study of the phenolic composition of cocoa (*Theobroma cacao*). J. Mass Spectrom. 38, 35-42.

Al Shaal, L., Shegokar, R., Müller, R.H., 2011. Production and characterization of antioxidant apigenin nanocrystals as a novel UK skin protective formulation. Int. J. Pharm. 420, 133-140.

Valls, J., Millán, S., Martí, M.P., Borràs, E., Arola, L., 2009. Advanced separation methods of food anthocyanins, isoflavones and flavanols. J. Chromatogr. A 1216, 7143-7172.

Villaverde, J.J., Ligero, P., de Vega, A., 2010. *Miscanthus* × *giganteus* as a source of biobased products through organosolv fractionation: A mini review. Open Agric. J. 4, 102-110.

Vukics, V., Guttman, A., 2010. Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. Mass Spectrom. Rev. 29, 1-16.

Xu, Y., Hanna, M.A., Isom, L., 2008. “Green” chemicals from renewable agricultural biomass – A mini review. Open Agr. J. 2, 54-61.

Zheng, Y-N., Zhang, J., Han, L-K., Sekiya, K., Kimura, Y., Okuda, H., 2005. Effects of compounds in leaves of *Salix matsudana* on Arachidonic acid metabolism. Yakugaku Zasshi 125, 1005-1008.

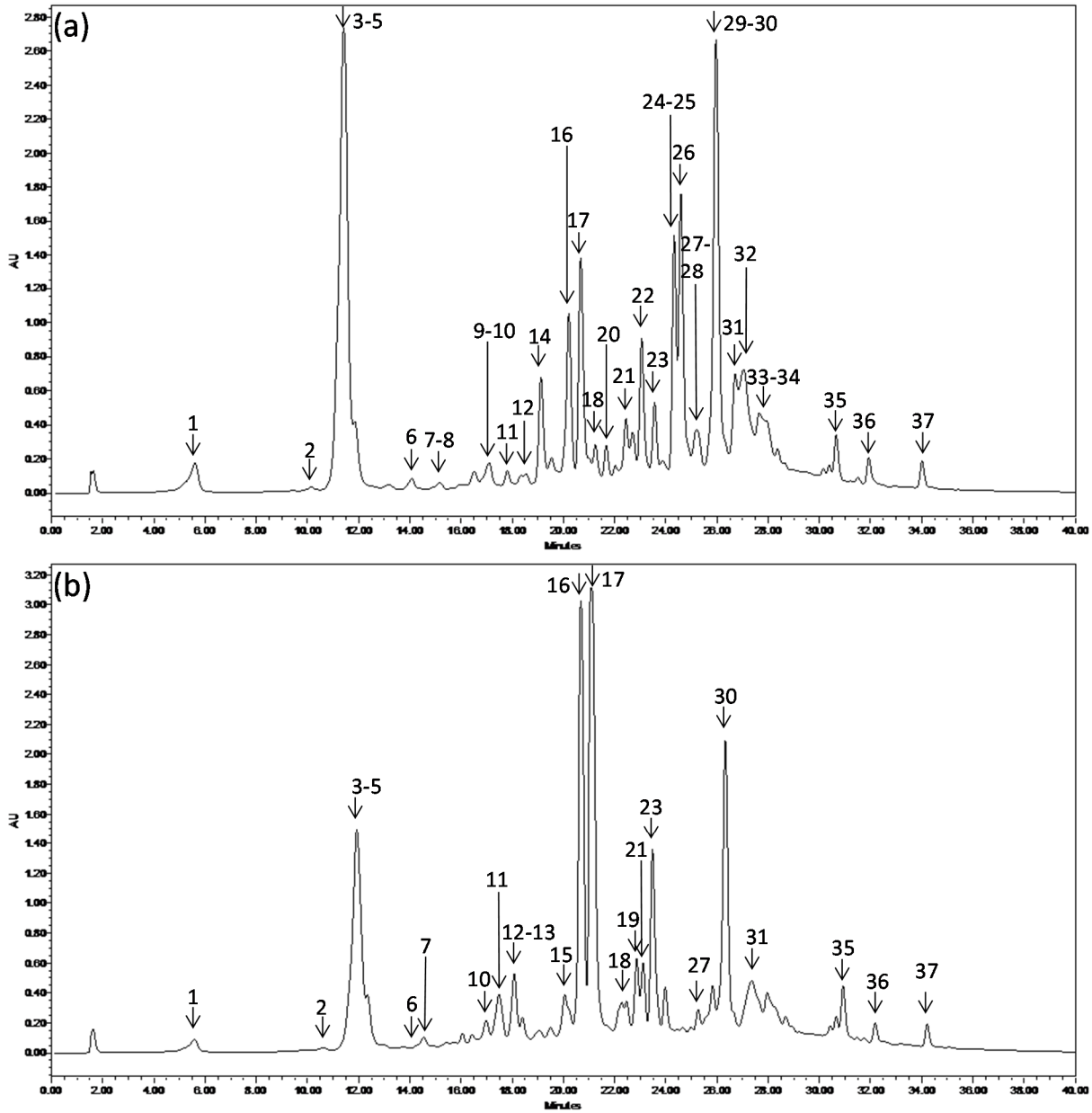

Figure 1: HPLC chromatograms (240-400 nm) in leaf tissue of Miscanthus sinensis mature paternal (a) and maternal (b) leaf tissue.

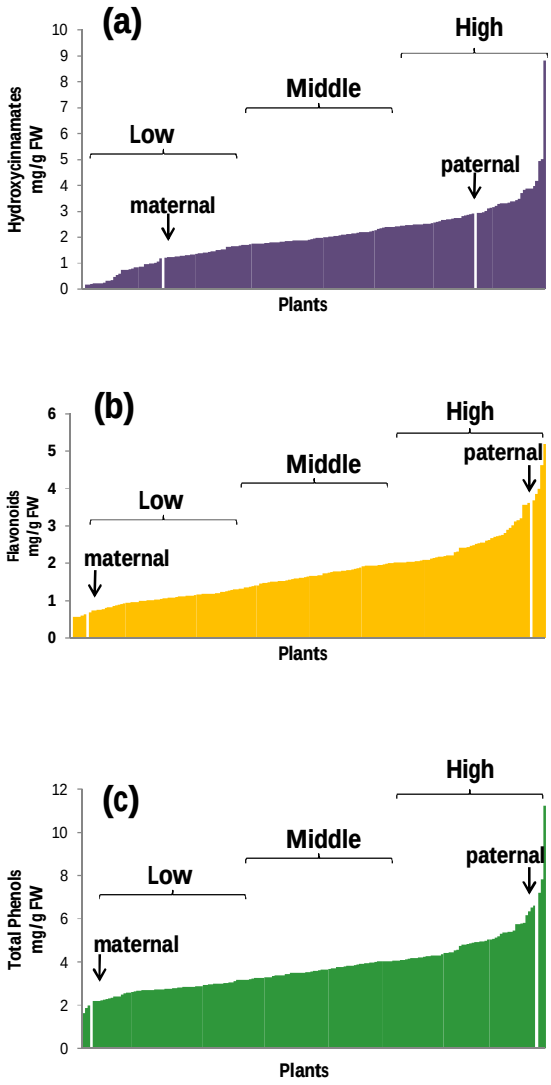

Figure 2: (a) Concentration of hydroxycinnamates in mapping family at one month growth (seedlings); (b) concentration of flavonoids in mapping family at one month growth; (c) total phenols in mapping family at one month growth.

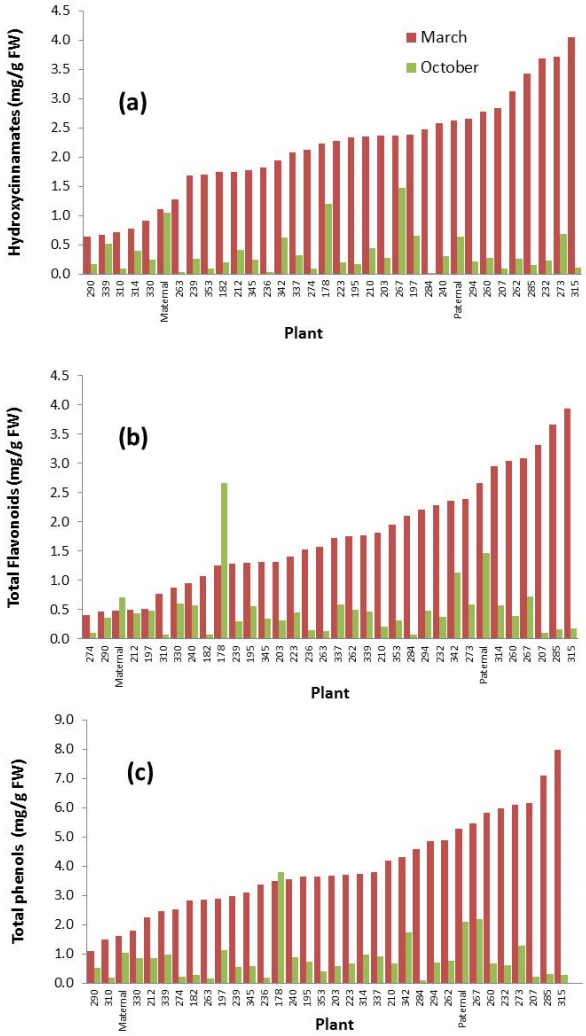

Figure 3: (a) Concentration (mg / g FW) of hydroxycinnamates in March and October; (b) concentration of flavonoids in March and October; concentration of total phenols in March and October.

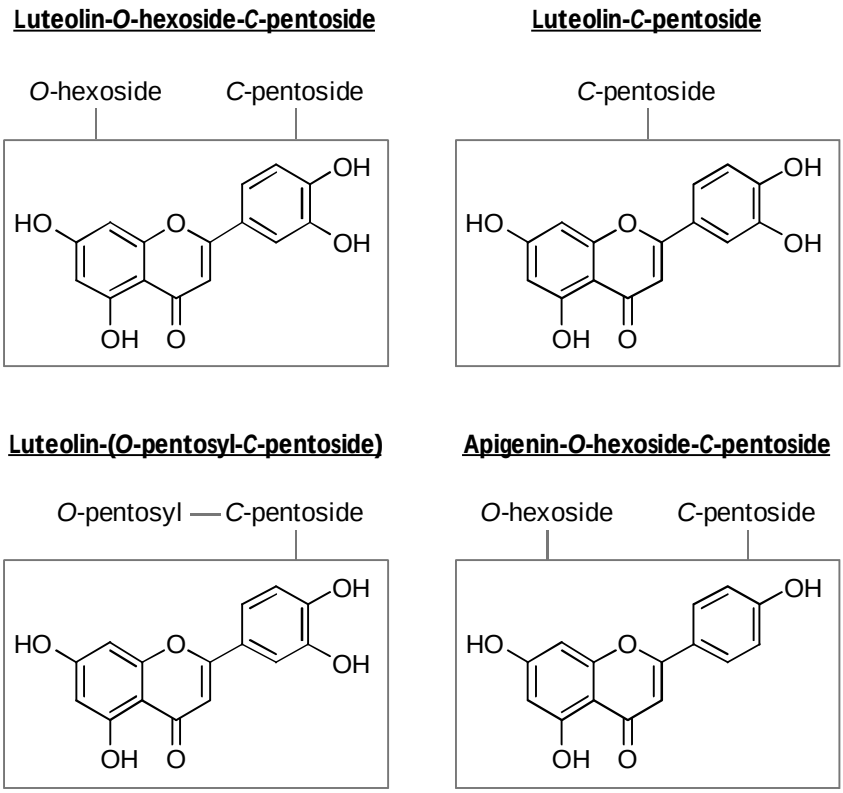

Figure 4: Structures of novel flavonoids in *Miscanthus sinensis* mature leaf tissue.

1 **Table 1**: Compounds identified in leaf of maternal (Mb 255) and paternal (Mb 111) plants using HPLC-ESI-MS^n (negative ion mode)

Peak / Cpd. No.	***m/z* (-ve mode)**	***MS^2 fragmentation (relative intensity %)***	***Compound***	***HPLC t_R (min)***	***λ_{max} (nm)***	***HPLC t_R (min)***	***λ_{max} (nm)***	***References***
				Paternal (Mb 111)		***Maternal (Mb 255)***		
1	353	**191,** 179 (43), 135 (7)	3-*O*-caffeoyl-quinic acid (*trans*)	5.6	290, 322	5.6	323	Ferreres et al. (2008)
2	367	**193,** 134 (5), 173 (3)	3-*O*-feruloyl-quinic acid (*trans*)	10.2	324	10.3	323	Ferreres et al. (2008)
3	353	**191**, 179 (6)	5-*O*-caffeoyl-quinic acid (*trans*)	11.4	325	11.9	326	Ferreres et al. (2008)
4	353	191 (25), 179 (63), **173**, 135 (9)	4-*O*-caffeoyl-quinic acid (*trans*)	11.8	326	11.7	325	Ferreres et al. (2008)
5	163	**119**	*para*-coumaric acid (*trans*)	11.8	309	12.3	n.d.	Ferreres et al. (2008)
6	353	**191**, 179 (4), 135 (1)	5-*O*-caffeoyl-quinic acid (*cis*)	14.5	319	14.2	320	Ferreres et al. (2008)
7	337	**191**, 163 (6), 119 (<0.5)	5-*O*-*para*-comaroyl-quinic acid (*trans*)	15.0	311	14.6	n.d.	Ferreres et al. (2008)
8	367	**173,** 191 (10)	4-*O*-feruloyl-quinic acid (*cis*)	15.2	319	n.d.	n.d.	Ferreres et al. (2008)
9	579	561 (16) $[(M-H)-H_2O]^-$, 519 (17) $[(M-H)-60]^-$, **489** $[(M-H)-90]^-$, 459 (21) $[(M-H)-120]^-$, 429 (4) $[(M-H)-150]^-$, 399 (14) [(M-H)-180;	Lut-6-*C*-Pent-8-*C*-Hex	17.1	272, 350	n.d.	n.d.	Lygin et al. (2011)

		AGly+113]$^-$, 369 (10) [(M-H)-210; AGly+83]$^-$, 327 (4) [(M-H)-252; AGly+41]$^-$						
10	367	**191**, 173 (2), 134 (1)	5-*O*-feruloyl-quinic acid (*trans*)	17.2	n.d.	17.0	n.d.	Ferreres et al. (2008)
11	367	193 (19), **173**, 135 (9)	4-*O*-feruloyl-quinic acid (*trans*)	17.7	325	17.5	325	Ferreres et al. (2008)
12	337	**191**, 163 (5), 119 (<0.5)	5-*O*-*para*-comaroyl-quinic acid (*cis*)	18.5	n.d.	18.0	n.d.	Ferreres et al. (2008)
13	593	575 (10) [(M-H)-H_2O]$^-$, 503 (30) [(M-H)-90]$^-$, **473** [(M-H)-120]$^-$, 455 (3) [(M-H)-138]$^-$, 413 (2) [(M-H)-180]$^-$, 383 (15) [(M-H)-210; AGly+113]$^-$, 353 (28) [(M-H)-240; AGly+83]$^-$	Apig-6,8-di-*C*-Hex	n.d.	n.d.	18.1	271, 336	Matsuta et al. (2011) and Kleinert and Barth (2008)
14	579	561 (<0.5) [(M-H)-H_2O]$^-$, 489 (2) [(M-H)-90]$^-$, 459 (2) [(M-H)-120]$^-$, **417** [(M-H)-162]$^-$, 399 (4) [(M-H)-180; AGly+113]$^-$, 357 (10) [(M-H)-222; AGly+71]$^-$, 327 (28.8) [(M-H)-252; AGly+41]$^-$	Lut-*O*-Hex-*C*-Pent	19.1	270, 347	n.d.	n.d.	
15	593	575 (5) [(M-H)-H_2O]$^-$, 503 (1) [(M-H)-90]$^-$, **473** [(M-H)-120]$^-$, 447 (4) [(M-H)-146]$^-$, 429 (66) [(M-H)-164]$^-$, 357 (24) [(M-H)-236; AGly+71]$^-$, 339 (12) [(M-H)-254; AGly+71-18]$^-$, 327 (6) [(M-H)-266; AGly+41]$^-$, 309	2”-*O*-Deoxyhex-*C*-Hex-Lut	n.d.	n.d.	20.1	270, 348	Lewandowski et al. (2000)

		(12) [(M-H)-284; AGly+41-18]$^-$						
16	447	429 (26) [(M-H)-H_2O]$^-$, 387 (3) [(M-H)-60]$^-$, 357 (79)[(M-H)-90; AGly+71]$^-$, **327** [(M-H)-120; AGly+41]$^-$**,** 285 (<0.5) [luteolin-H]$^-$	Lut-6-*C*-Glu	20.2	269, 348	20.7	270, 348	Kleinert and Barth (2008) and Martin (1970)
17	593	575 (6) [(M-H)-H_2O]$^-$, 503 (<0.5) [(M-H)-90]$^-$, **473** [(M-H)-120]$^-$, 447 (5) [(M-H)-146]$^-$, 429 (68) [(M-H)-164]$^-$, 357 (25) [(M-H)-236; AGly+71]$^-$, 339 (13) [(M-H)-254; AGly+71-18]$^-$, 327 (6) [(M-H)-266; AGly+41]$^-$, 309 (12)[(M-H)-284; AGly+41-18]$^-$, 285 (<0.5) [luteolin-H]$^-$	2”-*O*-Deoxyhex-*C*-Hex-Lut	20.7	269, 349	21.1	270, 348	Lewandowski et al. (2000)
18	593	578 (31) [(M-H)-CH_3]$^-$, 533 (12) [(M-H)-60]$^-$, 503 (65) [(M-H)-90]$^-$, 473 (17) [(M-H)-120]$^-$, 431 (86) [(M-H)-162]$^-$, 371 (28) [(M-H)-222]$^-$, **341** [(M-H)-252; AGly+41]$^-$, 327 (12) [(M-H)-266]$^-$, 298 (5) [(M-H)-295]$^-$	Chrys/Dios-*O*-Hex-*C*-Pent	21.2	271, 341	22.5	271, 335	
19	431	413 (8) [(M-H)-H_2O]$^-$, 341 (31) [(M-H)-90; AGly+71]$^-$, **311** [(M-H)-120; AGly+41]$^-$, 269 (<0.5) [apigenin-H]$^-$	Apig-6-*C*-Glu	n.d.	n.d.	22.9	271, 335	Mak et al. (2006)
20	447	327 (2) [(M-H)-120; AGly+41]$^-$, **285** [luteolin-H]$^-$	Lut-7-*O*-Glu	21.2	271, 345	n.d.	n.d.	Mak et al. (2006)
21	577	457 (7) [(M-H)-120]$^-$, **413** [(M-H)-164]$^-$, 341 (5) [(M-H)-236; AGly+71]$^-$, 323 (6) [(M-H)-254;	2”-*O*-Deoxyhex-*C*-Hex-Apig	22.4	271, 338	23.1	271, 340	Lewandowski et al. (2000)

		AGly+71-18]$^-$, 311 (2) [(M-H)-266; AGly+41]$^-$, 293 (36) [(M-H)-284; AGly+41-18]$^-$, 269 (<0.5) [apigenin-H]$^-$					
22	549	531 (6) [(M-H)-H_2O]$^-$, 459 (76) [(M-H)-90]$^-$, 417 (21) [(M-H)-132]$^-$, **399** [(M-H)-150]$^-$, 357 (27) [(M-H)-192; AGly+71]$^-$, 339 (4) [(M-H)-210; AGly+71-18]$^-$	2”-*O*-Pent-*C*-Pent-Lut	23.1	271, 348	n.d.	n.d.
23	577	559 (<0.5) [(M-H)-H_2O]$^-$, 517 (<0.5) [(M-H)-60]$^-$, 473 (4) [(M-H)-104]$^-$, **415** [(M-H)-162]$^-$, 353 (5) [(M-H)-224; AGly+83]$^-$, 311 (4) [(M-H)-266; AGly+41]$^-$	Apig-*O*-Hex-*C*-DeoxyHex	23.6	271, 349	23.5	270, 349
24	417	399 (25) [(M-H)-H_2O]$^-$, **357** [(M-H)-60; AGly+71]$^-$, 327 (49) [(M-H)-90; AGly+41]$^-$, 285 (<0.5) [luteolin-H]$^-$	Lut-*C*-Pent	24.3	269, 349	n.d.	n.d.
25	577	559 (<0.5) [(M-H)-H_2O]$^-$, 473 (3) [(M-H)-104]$^-$, **415** [(M-H)-162]$^-$, 353(2) [(M-H)-224; AGly+83]$^-$, 311 (6) [(M-H)-266; AGly+41]$^-$	Apig-*O*-Hex-*C*-Deoxyhex	24.4	271, 349	n.d.	n.d.
26	563	545 (7) [(M-H)-H_2O]$^-$, **473** [(M-H)-90]$^-$, 417 (17) [(M-H)-146]$^-$, 399 (93) [(M-H)-164]$^-$, 357 (24.6) [(M-H)-206; AGly+71]$^-$, 339 (6) [(M-H)-224; AGly+71-18]$^-$, 327 (7) [(M-H)-236; AGly+41]$^-$, 285 (<0.5) [luteolin-H]$^-$	2”-*O*-Deoxyhex-*C*-Pent-Lut	24.6	269, 351	n.d.	n.d.
27	533	515 (5) [(M-H)-H_2O]$^-$, 473 (15) [(M-H)-60]$^-$, 443 (31) [(M-H)-90]$^-$, 401	2”-*O*-Pent-*C*-Pent-Apig	25.2	271, 341	24.0	271, 348

		(3) [(M-H)-132]$^-$, **383** [(M-H)-150]$^-$, 353 (3) [(M-H)-180; AGly+83]$^-$, 341 (5) [(M-H)-192; AGly+71]$^-$, 323 (2) [(M-H)-210; AGly+71-18]$^-$, 311 (1) [(M-H)-222; AGly+41]$^-$, 293 (8) [(M-H)-240; AGly+41-18]$^-$					
28	429	411 (66) [(M-H)-H_2O]$^-$, **385** [(M-H)-44]$^-$, 357 (44) [(M-H)-72; AGly+71]$^-$, 325 (45) [(M-H)-104]$^-$, 285 (19) [luteolin-H]$^-$	Unknown Flavonoid	25.5	271, 341	n.d.	n.d.
29	431	413 (23) [(M-H)-H_2O]$^-$, 387 (3) [(M-H)-44]$^-$, **357** [(M-H)-74; AGly+71]$^-$, 327 (39) [(M-H)-104; AGly+41]$^-$, 285 (0.5) [luteolin-H]$^-$	Lut-*C*-Deoxyhex	26.0	271, 347	n.d.	n.d.
30	575	531 (16) [(M-H)-44]$^-$, 513 (3) [(M-H)-44, H_2O]$^-$, 429 (17) [(M-H)-146]$^-$, **411** [(M-H)-164]$^-$**,** 367 (29) [(M-H)-208]$^-$, 337 (8) [(M-H)-238]$^-$, 313 (2) [(M-H)-262]$^-$,	Unknown flavonoid	26.3	269, 350	26.3	269, 350
31	593	575 [(M-H)- H_2O]$^-$, 411 (2) [(M-H)-164]$^-$	Unknown flavonoid	26.7	270, 350	27.4	271, 349
32	577	559 (13) [(M-H)-H_2O]$^-$, 473 (95) [(M-H)-104]$^-$, **413** [(M-H)-164]$^-$, 357 (22) [(M-H)-220; AGly+71]$^-$, 339 (6) [(M-H)-238; AGly+71-18]$^-$, 327 (6) [(M-H)-250; AGly+41]$^-$, 309 (4) [(M-H)-268; AGly+41-18]$^-$, 285 (<0.5) [luteolin-H]$^-$	2”-*O*-Deoxyhex-*C*-Deoxyhex-Lut	27.0	270, 349	n.d.	n.d.
33	593	575 (48) [(M-H)-H_2O]$^-$, **411** [(M-H)-H_2O, 164]$^-$, 313 (10) [(M-H)-280]$^-$,	Unknown flavonoid	27.7	270, 349	n.d.	n..d

		285 (0.6) [luteolin-H]$^-$						
34	401	383 (1) [(M-H)-H_2O]$^-$, 341 (17) [(M-H)-60; AGly+71]$^-$, **311** [(M-H)-90; AGly+41]$^-$, 269 (<0.5) [apigenin-H]$^-$	Apig-8-*C*-Pent	27.9	272, 339	n.d.	n.d.	Kleinert and Barth (2008)
35	401	383 (2) [(M-H)-H_2O]$^-$, 341 (27) [(M-H)-60; AGly+71]$^-$, **311** [(M-H)-90; AGly+41]$^-$, 269 (<0.5) [apigenin-H]$^-$	Apig-8-*C*-Pent	30.7	271, 336	31.0	271, 334	Kleinert and Barth (2008)
36	415	397 (18) [(M-H)-H_2O]$^-$, **311** [(M-H)-104; AGly+41]$^-$	Apig-*C*-Deoxyhex	31.9	272, 341	32.2	n.d.	
37	413	395 (7) [(M-H)-H_2O]$^-$, **339** [(M-H)-74]$^-$, 311 (87) [(M-H)-102]$^-$, 285 (<0.5) [luteolin-H]$^-$	Unknown flavonoid	34.0	268, 341	34.2	n.d.	

Base peak (100%) is shown in bold
n.d. = not detected
Absence of references = compounds are considered novel
Hex: hexose; Glu: glucose; Pent: pentose; Deoxyhex: deoxyhexose; Apig: apigenin; Lut: luteolin; Chrys: chryseriol; Dios: diosmetin

Table 2: Quantification of hydroxycinnamates and flavone-glycosides in mature *Miscanthus sinensis* leaf tissue in paternal and maternal parental lines.

Peak / Compound No.	Compound	Paternal Leaf (µg/g fresh weight) (µmol g^{-1})	Maternal Leaf (µg/g fresh weight) (µmol g^{-1})
1	3-*O*-caffeoyl-quinic acid (*trans*)	45 (0.13)	9 (0.03)
2	3-*O*-feruloyl-quinic acid (*trans*)	1 (0.003)	1 (0.003)
3	5-*O*-caffeoyl-quinic acid (*trans*)	762 (2.15)	309 (0.87)
4	4-*O*-caffeoyl-quinic acid (*trans*)	-	-
5	*para*-coumaric acid (*trans*)	161 (0.38)	26 (0.16)
6	5-*O*-caffeoyl-quinic acid (*cis*)	-	-
7	5-*O*-*para*-coumaroyl-quinic acid (*trans*)	-	-
8	4-*O*-feruloyl-quinic acid (*cis*)	13 (0.01)	-
9	Lut-6-*C*-Pent-8-*C*-Hex	17 (0.03)	-
10	5-*O*-feruloyl-quinic acid (*trans*)	-	-
11	4-*O*-feruloyl-quinic acid (*trans*)	-	-

12	5-*O*-*para*-coumaroyl-quinic acid (*cis*)	-	-
13	Apig-6,8-di-*C*-Hex	-	2 (0.003)
14	Lut-*O*-Hex-*C*-Pent	49 (0.08)	-
15	2”-*O*-Deoxyhex-*C*-Hex-Lut	-	38 (0.06)
16	Lut-6-*C*-Glu	81 (0.18)	309 (0.69)
17	2”-*O*-Deoxyhex-*C*-Hex-Lut	114 (0.19)	384 (0.65)
18	Chrys/Dios-*O*-Hex-*C*-Pent	20 (0.03)	33 (0.06)
19	Apig-6-*C*-Glu	-	42 (0.10)
20	Lut-7-*O*-Glu	17 (0.04)	-
21	2”-*O*-Deoxyhex-*C*-Hex-Apig	35 (0.06)	40 (0.07)
22	2”-*O*-Pent-*C*-Pent-Lut	75 (0.14)	-
23	Apig-*O*-Hex-*C*-DeoxyHex	39 (0.07)	120 (0.21)
24	Lut-*C*-Pent	109 (0.26)	-
25	Apig-*O*-Hex-*C*-Deoxyhex		-

26	2”-*O*-Deoxyhex-*C*-Pent-Lut	164 (0.29)	-
27	2”-*O*-Pent-*C*-Pent-Apig	41 (0.10)	27 (0.05)
28	Unknown flavonoid *m/z* 429		-
29	Lut-6-*C*-Deoxyhex	305 (0.53)	-
30	Unknown flavonoid		218 (0.38)
31	Unknown flavonoid	61 (0.10)	83 (0.14)
32	2”-*O*-Deoxyhex-*C*-Deoxyhex-Lut	123 (0.21)	-
33	Unknown flavonoid	53 (0.09)	-
34	Apig-8-*C*-Pent	28 (0.07)	-
35	Apig-8-*C*-Pent	20 (0.05)	27 (0.07)
36	Apig-*C*-Deoxyhex	11 (0.03)	9 (0.02)
37	Unknown flavonoid	12 (0.03)	9 (0.02)

- = not detected in tissue extract / not quantifiable
Hex: hexose; Glu: glucose; Pent: pentose; Deoxyhex: deoxyhexose; Apig: apigenin; Lut: luteolin; Chrys: chryseriol; Dios: diosmetin