Supplementary Data 1. MS Identification of flax phenolics

Elucidation of the MS^2 spectra and the sequencing terminology of the first product ions was based on the lignin oligomer / (neo)lignan sequencing approach mentioned in Morreel *et al*. (2010b) and on the fragmentation rules of the different linkage types described in Morreel *et al*. (2010a). Briefly, the three types of linkages, i.e. 8–O–4 (β-aryl ether, AE), 8–5 (phenylcoumaran, PC) and 8–8 (resinol), either loose small neutral molecules that are indicative of the type of linkage (referred to as pathway I, I), or are cleaved, hence, yielding information on the units that are connected by the linkage (referred to as pathway II, II). In the case of a β-aryl ether, pathway II cleavage leads to first product ions corresponding with the phenolic 8-end (A^- ion) and aliphatic 4-end (B^- ion) moieties. Whenever two 8–O–4-linkages are present, the particular linkage from which the ions were cleaved is indicated by a superscript ($^8A^-$ and $^8B^-$ for the 8-end-located linkage and $^4A^-$ and $^4B^-$ for the 4-end-located linkage). If an 8–O–4 linkage is still present in the first product ion, a further pathway I (AE^2 I) and pathway II (AE^2 II) fragmentation can occur. The first product ion from which these new product ions result are indicated partially as a superscript to the AE annotation (e.g. AE^{2A}), but more completely as "$X|^yX^-$" with X = A or B and y = 4 or 8.

p-Coumaryl alcohol hexoside (**1**)

The peak with m/z 311.11488 ($C_{15}H_{19}O_7^-$) and a retention time of 4.87 min showed a base peak at m/z 149 in its MS^2 spectrum due to the loss of hexose (-162 Da) upon collision-induced dissociation (CID). In the MS^3 spectrum of this first product ion, a further water loss leading to the ion at m/z 131 appeared. In the gas-phase, *p*-coumaryl alcohol should easily loose water by a charge-induced process starting from the phenoxide anion. Furthermore, as no methoxy groups – leading to methyl radical loss - are present, no other fragmentations are expected. Based on a search for *p*-coumaryl alcohol 4–*O*–β–D–glucopyranoside in the Scifinder database, this compound was identified for the first time in *Arum italicum* (Della Greca *et al*., 1993), but has not been detected in flax up to now.

G(8–O–4)**S** hexoside (**2**)

At 10.97 min, a peak eluted with a m/z value of 627.23060 ($C_{29}H_{39}O_{15}^-$). Its MS^2 spectrum indicated that the compound appeared as an acetate adduct (loss of 60 Da). A further loss of a hexose moiety (162 Da) yielded the peak at m/z 405. The MS^3 spectrum of the latter peak has been obtained before during phenolic profiling of poplar xylem and was identified as **G**(8–O–4)**S** (Morreel *et al*., 2004a). The hexoside is likely attached to the phenolic function of the aglycone, preventing it from being ionized in the gas phase, hence, explaining why this compound was mainly detected as the acetate adduct. **G**(8–O–4)**S** 4–*O*–β–D–glucopyranoside was isolated for the first time from orange and lemon and called citrusin B (Matsubara *et al*., 1985), but not been detected in flax up to now.

Secoisolariciresinol hexoside (**3**)

The MS^2 spectrum of the peak with retention time 11.40 min and m/z 523.21976 ($C_{26}H_{35}O_{11}^-$) showed the loss of a hexose moiety (162 Da) yielding a product ion at m/z 361. The MS^3 spectrum

of the latter ion was identical to a previously analyzed standard, namely secoisolariciresinol (Morreel *et al*., 2010a). Another isomer of secolariciresinol hexoside was observed at 13.41 min. They likely represent secoisolariciresinol 4–*O*–β–D–glucopyranoside and secoisolariciresinol 9–*O*–β–D–glucopyranoside. Based on the Scifinder database, both compounds were identified for the first time in *Urtica dioica* (Kraus and Spiteller, 1990) and *Pinus sylvestris* (Popoff and Theander, 1977), respectively. The 9–*O*–β–D–glucopyranoside has been identified following acid hydrolysis of secoisolariciresinol diglucoside oligomers from defatted flaxseed powder (Li *et al*., 2008).

Lariciresinol hexoside (**4**)

A peak with m/z 521.20407 ($C_{26}H_{33}O_{11}^-$) eluted at 11.55 min. This compound fragmented in the gas phase by the loss of a hexose moiety (m/z 359 in MS^2 spectrum). Another major product ion was observed at m/z 329. Such a predominant formaldehyde loss (-30 Da) from the aglycone is characteristic for lariciresinol (Morreel *et al*., 2010a). Another lariciresinol hexoside isomer was observed at 12.51 min. Lariciresinol 9–*O*–β–D–glucopyranoside was identified for the first time in *Pteris vittata* (Satake *et al*., 1978), and lariciresinol 4´–*O*–β–D–glucopyranoside and lariciresinol 4–*O*–β–D–glucopyranoside in *Arum italicum* (Della Greca *et al*., 1993) and in *Asclepias subulata* (Jolad *et al*., 1986), respectively. To our knowledge, none of these were previously detected in flax.

G(8–O–4)**S**(8–O–4)**G** (**5**)

The MS^2 spectrum of the m/z 601.23007 peak ($C_{31}H_{37}O_{12}^-$) eluting at 12.00 min was elucidated with the sequencing approach proposed by Morreel *et al*. (2010b). The first product ion at m/z 553 (AE I) represents the pathway I cleavage (loss of formaldehyde and water) typical of an 8–O–4-linkage. The pathway II cleavage is represented by the ion at m/z 421 (AE II $^4A^-$), yet no first product ion corresponding with the AE II $^4B^-$ ion was observed. A combined formaldehyde and water loss from the AE II $^4A^-$ ion at m/z 421, explained the first product ion at m/z 373 (AE^{2A} I). This indicated the presence of a second 8–O–4-linkage. Cleavage of this latter 8–O–4-linkage yielded the ions at m/z 405 (AE II $^8B^-$) and 195 (AE II $^8A^-$), the latter corresponding to a **G** unit. Finally, cleavage of the remaining 8–O–4-linkage of the first product ion at m/z 421 lead to the ions at m/z 225 (AE^{2A} II $B|^4A^-$) and m/z 195 (AE^{2A} II $A|^4A^-$), corresponding with an **S** and a **G** unit, respectively. Other G(8–O–4)S(8–O–4)G stereomers were found at 15.28 and 15.99 min. Based on the Scifinder database, this compound has not been identified in any plant species up to now.

Isodihydrodehydrodiconiferyl alcohol (IDDDC) hexoside (**6**)

At 12.58 min, a compound eluted with m/z 581.22511 ($C_{28}H_{37}O_{13}^-$). The MS^2 spectrum showed that this compound was detected as an acetate adduct (60 Da loss). The further loss of a hexose moiety (162 Da) provided a first product ion at m/z 359. The latter aglycone dissociated further by methyl radical loss (15 Da), yielding the base peak at m/z 341, and by formaldehyde loss (30 Da) giving rise to the product ion at m/z 329. These losses are typical for isodihydrodehydrodiconiferyl alcohol (IDDDC). As this compound is mainly detected as the acetate adduct, the hexoside likely resides on a phenolic function of the aglycone (see **G**(8–O–4)**S** hexoside). Two other isomers of this compound,

observed at 13.89 min 14.50 min, probably have the hexoside attached to a primary alcohol group as they were detected in their deprotonated form. IDDDC 4′–*O*–β–D–glucopyranoside has been previously detected in *Epimedium diphyllum* (Miyase *et al*., 1989) and in several other plant species as well – but not in flax yet. In contrast and based on the retrieved articles from the Scifinder database, IDDDC 4–*O*–β–D–glucopyranoside, IDDDC 9–*O*–β–D–glucopyranoside and IDDDC 9′–*O*–β–D–glucopyranoside have never been observed in any plant species.

H(8–5)**G** hexoside (**7**)

The compound eluting at 12.73 min was detected as an acetate adduct with m/z 549.19884 ($C_{27}H_{33}O_{12}^-$) based on its MS^2 spectrum (-60 Da loss). Following acetate loss, the first product ion at m/z 489 lost an additional hexose moiety (-162 Da) yielding the peak at m/z 327. Water (-18 Da) and formaldehyde (-30 Da) loss of the latter peak explains the product ions at m/z 309 and 297 that are typical for a phenylcoumaran (Morreel *et al*., 2010a). Because the acetate adduct was detected rather than the deprotonated molecule, the hexoside likely resides on the phenolic function of the aglycone. Although both **G**(8–5)**H** and **H**(8–5)**G** are possible structures for the aglycone, the latter is more likely given the higher oxidation potential of *p*-coumaryl alcohol as compared to that of coniferyl alcohol (Syrjänen and Brunow, 1998). No publications of **H**(8–5)**G** 4–*O*–β–D–glucopyranoside were retrieved using the Scifinder database.

G(8–O–4)**G**/**G**^{red}(8–8)**G**^{red}/**G** hexoside (**8**)

A m/z 717.27850 ($C_{36}H_{45}O_{15}^-$) peak appeared at 13.58 min. The MS^2 spectrum showed a base peak at m/z 555 due to the loss of a hexose residue (162 Da). MS^3 fragmentation of this first product ion lead to second product ions at m/z 537, 525 and 507 due to the loss of water (18 Da), formaldehyde (30 Da) and the combined loss of water and formaldehyde (48 Da), i.e. characteristic pathway I cleavages of a β-aryl ether linkage (AE I ions). The pathway II-typical cleavages yielded ions at m/z 195 (AE II $^8A^-$) and m/z 359 (AE II $^8B^-$) indicating that a **G** unit was connected via an 8–O–4-linkage to a dimeric moiety. The second product ion at m/z 329 results from the further dissociation of the second product ion at m/z 359 and unravels the dimeric structure as lariciresinol. Both **G**(8–O–4)**G**(8–8)**G**^{red} and **G**(8–O–4)**G**^{red}(8–8)**G** have been purified from plants: the former compound for the first time from *Brassica fruticulosa* (Cutillo *et al*., 2003) and the latter from *Ehretia ovalifolia* (Yoshikawa *et al*., 1995). Their hexosides have never been described in any plant species.

Feruloyl coniferin (**9**)

The m/z 517.17154 ($C_{26}H_{29}O_{11}^-$) peak eluting at 13.62 min showed a base peak at m/z 337 in its MS^2 spectrum. The loss of 180 Da was also observed in the full MS spectrum due to ion source fragmentation. Based on the full MS spectrum, an accurate mass for this product ion could be determined (m/z 337.09313, $C_{16}H_{17}O_8^-$), indicating that the neutral loss of 180 Da was associated with a coniferyl alcohol moiety. The MS^2 base peak fragmented further to m/z 193 by losing 144 Da corresponding to a hexose moiety. This was confirmed by MS^3 fragmentation. Second product ions at m/z 193, 149 and 134 indicated that the hexose was attached to a ferulic acid moiety. Losses of -60,

-90 and -120 Da confirmed that the hexose was present as an ester. Thus, this molecule was elucidated as feruloyl coniferin. Based on the Scifinder database, this compound has not yet been described in a plant species.

Pinoresinol hexoside (**10**)

At 14.08 min, a peak eluted with m/z 519.18875 ($C_{26}H_{31}O_{11}^-$). Its MS^2 spectrum was dominated by a hexose loss (162 Da) yielding a first product ion at m/z 357. MS^3 fragmentation of the latter ion revealed a spectrum identical to that of pinoresinol (Morreel *et al*., 2010a). Pinoresinol glucoside has been detected in various plants but not yet in flax (Scifinder database). It was described for the first time in *Forsythia suspensa* (Chiba *et al*., 1978).

G(8–O–4)**G**(8–O–4)**G´** (**11**)

The peak with m/z 569.20468 ($C_{30}H_{33}O_{11}^-$) and a retention time of 15.06 min had a MS^2 spectrum that was dominated by first product ions at m/z 177 and 162 indicating the presence of a coniferaldehyde moiety. First product ions at m/z 521 (AE I) and 391 (AE II $^4A^-$) were indicative for a β-aryl ether trilignol. No publication concerning this compound was retrieved in the Scifinder database.

G(8–O–4)**G/G^{red}**(8–5/8)**G^{red}/G** hexoside (**12**)

At 16.06 min, a peak with m/z 747.28685 ($C_{37}H_{47}O_{16}^-$) was observed that lost hexose (162 Da) upon CID yielding m/z 585 as the major first product ion. MS^3 fragmentation of this ion leads to second product ions at m/z 537 (AE I ion) and m/z 359 (AE II $^8B^-$ ion). The latter ion corresponds with a moiety derived from **G^{red}**(8–5)**G** (dihydrodehydrodiconiferyl alcohol, DDDC), **G**(8–5)**G^{red}** (isodihydrodehydrodiconiferyl alcohol, IDDDC), or **G**(8–8)**G^{red}** (lariciresinol). **G**(8–O–4)**G**(8–5)**G^{red}**, **G**(8–O–4)**G^{red}**(8–8)**G** and **G**(8–O–4)**G**(8–8)**G^{red}** have been isolated from *Vitex rotundifolia* (Ono *et al*., 1998), *Ehretia ovalifolia* (Yoshikawa *et al*., 1995) and *Brassica fruticulosa* (Cutillo *et al*., 2003). However, glycosides of these compounds have not been detected in any plant species up to now.

Sinapoyl coniferin (**13**)

An m/z 547.18218 ($C_{27}H_{31}O_{12}^-$) peak eluted at 16.64 min. CID leads to a base peak at m/z 367 in its MS^2 spectrum. This ion was also produced due to ion source fragmentation and, thus, observable in the FT-based full MS spectrum where its m/z value could be more accurately determined as 367.10416 ($C_{17}H_{19}O_9^-$). Therefore, this first product ion is associated with the loss of a coniferyl alcohol residue. MS^3 fragmentation of m/z 367 leads to second product ions characteristic of sinapoyl hexose (Dauwe *et al*., 2007). Thus, this compound is sinapoyl coniferin and has not been known to exist in plant species up to now.

N-containing derivate of feruloyl coniferin (**14**)

At 16.75 min, a compound eluted with a m/z value of 714.25241 ($C_{34}H_{40}O_{14}N_3^-$). In the MS^2 spectrum, a base peak was observed at m/z 517. Due to ion source fragmentation, the latter ion was also present in the full MS spectrum at m/z 517.17184 ($C_{26}H_{29}O_{11}^-$). Therefore, the MS^2 base peak was associated with a neutral loss corresponding to the chemical formula $C_8H_{11}O_3N_3$. Due to the ease of fragmentation, this moiety is likely attached via an ester bond, yielding $C_8H_{13}O_4N_3$ (taking water into account) as the chemical formula of the biochemical precursor of this moiety. Searching the PubChem database for this chemical formula provided 198 matches. However, none of them was a known biochemical, thus, we could not further resolve the identity of the N-containing precursor. MS^3 fragmentation of the m/z 517 first product ion yielded a spectrum identical to that of feruloyl coniferin.

G(8–O–4)**G**^{red}(8–8)**G**^{red} (**15**)

The m/z 557.24022 ($C_{30}H_{37}O_{10}^-$) ion at retention time 17.13 min showed a water loss and a combined water and formaldehyde loss in its MS^2 spectrum (AE I ions at m/z 539 and 509). The β-aryl ether pathway II-associated cleavage leads to the dimeric AE II $^8B^-$ ion at m/z 361, which corresponds with the m/z value of secoisolariciresinol (Morreel *et al*., 2010a). This compound has been previously identified in *Abies marocana* (Barrero *et al*., 1996), but not yet in flax.

G(8–O–4)**S**(8–8)**G** hexoside (**16**)

A m/z 745.27291 ($C_{37}H_{45}O_{16}^-$) peak eluted at 17.78 min. Its MS^2 spectrum showed a prominent hexose loss (162 Da) and further MS^3 fragmentation of the corresponding first product ion yielded a spectrum identical to that published for **G**(8–O–4)**S**(8–8)**G** (Morreel *et al*., 2004a). Two other isomers eluted at 17.90 and 18.59 min. **G**(8–O–4)**S**(8–8)**G** 4′′–*O*–β–D–glucopyranoside has only been purified from *Scutellaria baicalensis* (Yukinori and Tsuyoshi, 1998).

H(8–O–4)**G**(8–5)**G**′ (**17**)

At 19.12 min, a peak eluted at m/z 521.18310 ($C_{29}H_{29}O_9^-$). Its MS^2 spectrum showed losses typical for the presence of a β-aryl ether, i.e. a water loss and a combined formaldehyde and water loss yielding the first product ions at m/z 503 (AE I) and 473 (AE I). The β-aryl ether pathway II-associated product ions were not observed. In the case of the AE II $^8B^-$ ion, this is explained by the subsequent loss of a water or a formaldehyde molecule leading to the phenylcoumaran pathway I-associated (PC I) product ions at m/z 337 and 325. These latter ions indicate a **G**(8–5)**G**′ moiety, whereas the mass differences with the parent ion reveals a 8–O–4-linkage to a **H** unit. The presence of first product ions at m/z 175 and 190 are in agreement with a hydroxycinnamaldehyde-derived **G**′ endgroup (Morreel *et al*., 2004a). No articles were retrieved from the Scifinder database.

G(8–O–4)**S**/**S**^{red}(8–5/8)**G**^{red}/**G** (**18**)

A trilignol eluted at 20.11 min with an m/z value of 585.23525 ($C_{31}H_{37}O_{11}^-$). In the MS^2 spectrum, an 8–O–4-linkage could be deduced from the AE I ions at m/z 567 and 537. The AE II $^8B^-$ ion at m/z 389 suggests a reduced phenylcoumaran or resinol dimeric moiety to which a **G** unit is connected (AE II $^8A^-$ ion at m/z 195). **G**(8–O–4)**S**(8–5)**G**^{red} and **G**(8–O–4)**S**(8–8)**G**^{red} have been purified for the first time from the bark of *Acer nikoense* (Morikawa *et al*., 2003) and from *Campylotropis hirtella* (Han *et al*., 2008), but have not been observed in flax yet.

G(8–O–4)**G**(8–O–4)**G**(8–5)**G´** (**19**)

At 20.20 min, a tetramer eluted (m/z 747.26721, $C_{40}H_{43}O_{14}^-$). The MS^2 ions at m/z 729 and 699 are the β-aryl ether pathway I-associated AE I ions. The corresponding pathway II-associated AE II $^8B^-$ ion at m/z 551 was subjected to another β-aryl ether pathway I-associated cleavage as indicated by the product ion at m/z 503 (AE^{2B} I). Pathway II cleavage of the latter 8–O–4-linkage did not yield a dimeric product ion as it fragmented further by water and formaldehyde loss to the PC I product ions at m/z 337 and 325, indicating a terminal **G**(8–5)**G´** moiety. No articles describing this compound have been retrieved via Scifinder.

H(8–O–4)**S**(8–5)**G´** (**20**)

The compound at retention time 20.92 min showed an m/z value (551.19354, $C_{30}H_{31}O_{10}^-$) that is approximately 30 Da higher than that of **H**(8–O–4)**G**(8–5)**G´**, suggesting the presence of an additional methoxy substituent. Indeed, the first product ions at m/z 533 (AE I) and 503 (AE I) result from a β-aryl ether-associated pathway I dissociation, yet the absence of the corresponding pathway II product ions and the detection of the PC I product ions at m/z 367 and 355, reveals a phenylcoumaran dimeric moiety with a hydroxycinnamaldehyde-derived **G´** endgroup. The latter is further substantiated by the presence of a product ion at m/z 218 (Morreel *et al*., 2004a). No articles describing this compound have been retrieved via Scifinder.

References cited in Table 2

Ref 1. Dauwe R, Morreel K, Goeminne G, Gielen G, Rohde A, Van Beeumen J, Ralph J, Boudet AM, Kopka J, Rochange SF, Halpin C, Messens E and Boerjan W. (2007). Molecular phenotyping of lignin-modified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration. *Plant J*. 52: 263-285.

Ref 2. Vanholme R, Storme V, Vanholme B, Christensen JH, Goeminne G, Rohde A, Morreel K, Boerjan W. A systems biology view of the plant's response to lignin perturbations. The Plant Cell, submitted.

Ref 3. Matsuda F, Yonekura-Sakakibara K, Niida R, Kuromori T, Shinozaki K and Saito K. (2009). MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. *Plant J.* 57(3) : 555-577.

Ref. 4. Vanholme R, Ralph J, Akiyama T, Lu FC, Pazo JR, Kim H, Christensen JH, Van Reusel B, Storme V, De Rycke R, Rohde A, Morreel K, Boerjan W (2010) Engineering traditional monolignols out of lignin by concomitant up-regulation of F5H1 and down-regulation of COMT in Arabidopsis. *Plant Journal* **64**, 885-897

Ref 5. Morreel, K., Ralph, J., Kim, H., Lu, F., Goeminne, G., Ralph, S., Messens, E. and Boerjan, W. (2004a) Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. *Plant Physiol.* **136**, 4032-4036

Ref 6. Morreel, K., Kim, H., Lu, F., Dima, O., Akiyama, T., Vanholme, R., Niculaes, C., Goeminne, G., Inzé, D., Messens, E., Ralph, J. and Boerjan, W. (2010a) Mass spectrometry-based fragmentation as an identification tool in lignomics. *Anal. Chem.* **82**, 8095-8105

Ref 7. Morreel, K., Dima, O., Kim, H., Lu, F., Niculaes, C., Vanholme, R., Dauwe, R., Goeminne, G., Inzé, D., Messens, E., Ralph, J. and Boerjan, W. (2010b) Mass spectrometry-based sequencing of lignin oligomers. *Plant Physiol.* **153**,1464-1478

Ref 8. Morreel K, Ralph J, Lu F, Goeminne G, Busson R, Herdewijn P, Goeman JL, Van der Eycken J, Boerjan W and Messens E (2004b). Phenolic profiling of caffeic acid *O*-methyltransferase-deficient poplar reveals novel benzodioxane oligolignols. *Plant Physiol* **136**: 4023–4036

Réferences for compounds previously identified in flax (F. Ref. in Table 2).

F. Ref. 1. Ford, J.D., Huang, K.-S., Wang, H.-B., Davin, L.B., Lewis, N.G. (2001) Biosynthetic pathway to the cancer chemopreventive secoisolariciresinol diglucoside-hydroxymethyl glutaryl ester-linked lignan oligomers in flax (Linum usitatissimum) seed. Journal of Natural Products 64, 1388-1397

F. Ref 2. Beejmohun, V., Fliniaux, O., Hano, C., Pilard, S., Grand, E., Lesur, D., Cailleu, D., Lamblin, F., Laine, E., Kovensky, J., et al. (2007) Coniferin dimerization in lignan biosynthesis in flax cells. Phytochemistry 68, 2744-2752

F. Ref. 3. Ibrahim, R.K., Shaw, M. (1970) Phenolic constituents of the oil flax (Linum usitatissimum) Phytochemistry 9, 1855-1858

F. Ref. 4. Keen, N.T., Littlefield, L.J. (1979) The possible association of phytoalexins with resistance gene expression in flax to Melampsora lini Physiological Plant Pathology 14, 265-280

F. Ref. 5. Attoumbre, J., Charlet, S., Baltora-Rosset, S., Hano, C., Raynaud-Le Grandic, S., Gillet, F., Bensaddek, L., Mesnard, F., Fliniaux, M.-A. (2006) High accumulation of dehydrodiconiferyl alcohol-4-β-D-glucoside in free and immobilized Linum usitatissimum cell cultures. Plant Cell Reports 25, 859-864

F. Ref 6. Hano, C., Addi, M., Bensaddek, L., Cronier, D., Baltora-Rosset, s., Doussot, J., Maury, S., Mesnard, F., Chabbert, B., Hawkins, S., et al. (2006) Differential accumulation of monolignol-derived compounds in elicited flax (Linum usitatissimum) cell suspension cultures. Planta 223, 975-989

F. Ref. 7. Schmidt, T.J., Hemmati, S., Fuss, E., Alfermann, A.W. (2006) A combined HPLC-UV and HPLC-MS method for the identification of lignans and its application to the lignans of Linum usitatissimum L. and L. bienne Mill. Phytochemical Analysis, 17, 299-311

F. Ref 8. Sicilia, T., Niemeyer, H.B., Honig, D.M., Metzler, M. (2003) Identification and stereochemical characterization of lignans in flaxseed and pumpkin seeds. Journal of Agricultural and Food Chemistry 51, 1181-1188

References for compounds 1-20 in Fig 3.

Barrero, A.F., Haidour, A., Dorado, M.M., Cuerva, J.M. (1996) Two sesquilignans from the wood of *Abies marocana*. *Phytochemistry* **41,** 605-609

Chiba M., Hisada, S. and Nishibe, S. (1978) Studies on the Chinese drug "Forsythiae fructus" III. On the constituents of fruits of *Forsythia viridissima* and *F. suspense*. *Shoyakugaku Zasshi* **32,** 194-197

Cutillo, F., D'Abrosca, B., DellaGreca, M., Fiorentino, A. and Zarrelli, A. (2003) Lignans and neolignans from Brassica fruticulosa: effects on seed germination and plant growth. *Journal of Agricultural and Food Chemistry* **51,** 6165-6172

Dauwe, R., Morreel, K., Goeminne, G., Gielen, B., Rohde, A., Van Beeumen, J., Ralph, J., Boudet, A.M., Kopka, J., Rochange, S.F., Halpin, C., Messens, E. and Boerjan, W. (2007) Molecular phenotyping of lignin-modified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration. *Plant J*. **52**, 263-285

Della Greca, M., Molinaro, A., Monaco, P. and Previtera, L. (1993) Two new lignan glucosides from *Arum italicum*. *Heterocycles* **9,** 2081-2086

Han, H.-Y., Liu, H.-W., Wang, N.-L. and Yao, X.-S. (2008) Sesquilignans and dilignans from *Campylotropis hirtella* (Franch.) Schindl. *Natural Product Research* **22,** 984-989

Jolad, S.D., Bates, R.B., Cole, J.R., Hoffmann, J.J., Siahaan, T.J. and Timmermann, B.N. (1986) Cardenolides and a lignan from *Asclepias subulata*. *Phytochemistry* **25,** 2581-2590

Kraus, R. and Spiteller, G. (1990) Lignan glucosides from roots of *Urtica dioica*. *Liebigs annalen der chemie* **12,** 1205-1213

Li, X., Yuan, J.-P., Xu, S.-P., Wang, J.-H. and Liu, X. (2008) Separation and determination of secoisolariciresinol diglucoside oligomers and their hydrolysates in the flaxseed extract by high-performance liquid chromatography. *Journal of Chromatography A* **1185,** 223-232

Matsubara, Y., Kumamoto, H., Sawabe, A., Iizuka, Y. and Okamoto, K. (1985) Structure and physiological acitivity of phenyl propanoid glycosides in lemon, unshiu and orange peelings. *Kinki Daigaku Igaku Zasshi* **10,** 51-58

Miyase, T., Ueno, A., Takizawa, N., Kobayashi, H. and Oguchi, H. (1989) Ionone and lignan glycosides from *Epimedium diphyllum*. *Phytochemistry* **28,** 3483-3485

Morikawa, T., Tao, J., Ueda, K., Matsuda, H. and Yoshikawa, M. (2003) Medicinal foodstuffs. XXXI. Structures of new aromatic constituents and inhibitors of degranulation in RBL-2H3 cells from a Japanese folk medicine, the stem bark of *Acer nikoense*. *Chemical & Pharmaceutical Bulletin* **51,** 62-67

Morreel, K., Ralph, J., Kim, H., Lu, F., Goeminne, G., Ralph, S., Messens, E. and Boerjan, W. (2004a) Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. *Plant Physiol.* **136**, 4032-4036

Morreel, K., Kim, H., Lu, F., Dima, O., Akiyama, T., Vanholme, R., Niculaes, C., Goeminne, G., Inzé, D., Messens, E., Ralph, J. and Boerjan, W. (2010a) Mass spectrometry-based fragmentation as an identification tool in lignomics. *Anal. Chem.* **82**, 8095-8105

Morreel, K., Dima, O., Kim, H., Lu, F., Niculaes, C., Vanholme, R., Dauwe, R., Goeminne, G., Inzé, D., Messens, E., Ralph, J. and Boerjan, W. (2010b) Mass spectrometry-based sequencing of lignin oligomers. *Plant Physiol.* **153**,1464-1478

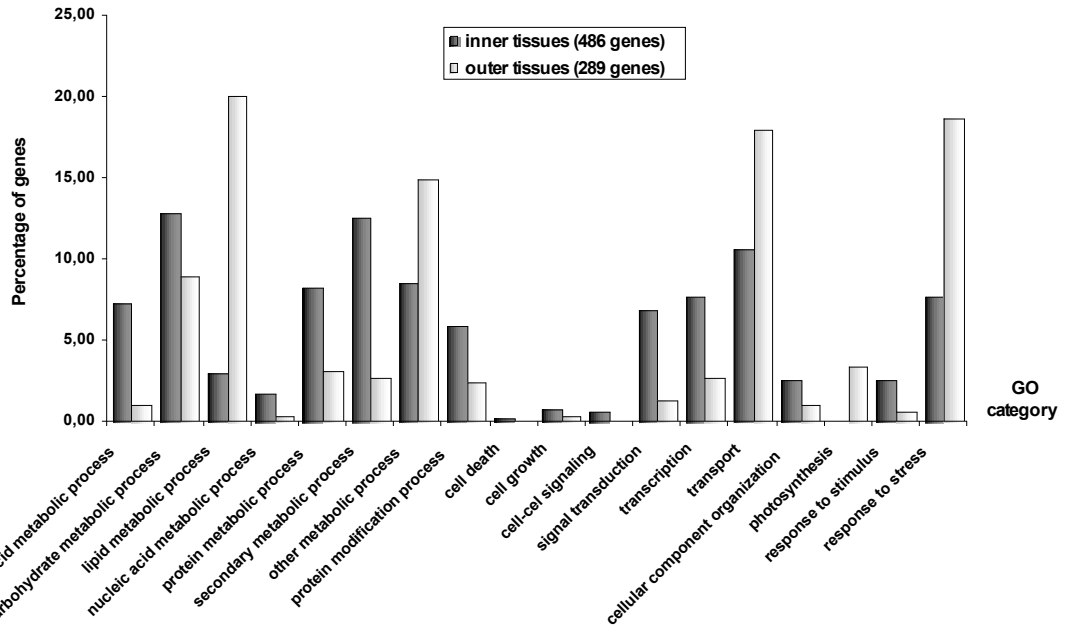
Ono, M., Masuoka, C., Ito, Y. and Nohara, T. (1998) Antioxidative constituents from *Viticis trifoliate Fructus* (fruit of *Vitex rotundifolia* L.) *Food Science and Technology International* **4,** 9-13

Popoff, T. and Theander, O. (1977) The constituents of conifer needles. VI. Phenolic glycosides from Pinus Sylvestris. *Acta Chemica Scandinavica, Series G: Organic Chemistry and Biochemistry* **B31,** 329-337

Satake, T., Murakami, T., Saiki, Y. and Chen, C.-M. (1978) Chemical and chemotaxonomic studies of the genus Pteris and related genera (Pteridaceae). Part XIX. Chemical studies of the constituents of Pteris vittata L. *Chemical & Pharmaceutical Bulletin* **26,** 1619-1622

Yoshikawa, K., Kageyama, H. and Arihara, S. (1995) Phenolic glucosides and lignans from *Ehretia ovalifolia*. *Phytochemistry* **39,** 659-664

Yukinori, M. and Tsuyoshi, T. (1998) Studies on constituents of *Scutellaria* species. XIX. Lignan glycosides of roots of *Scutellaria baicalensis* GEORGI. *Natural medicines* **52**, 82-86



Supplementary data 3. Gene ontology (GO) biological process categories. Plot displays percentage of annotated genes that were identified as differentially expressed in outer *vs*. inner tissues of flax shoots. The total number of genes selected as up-regulated in each type of tissue for log2ratio >1.5 and <-1.5 is listed in the legend. The complete list of genes is provided in Supplementary data file 1.