Available online at www.sciencedirect.com
SCIENCE DIRECT
Plant Physiology and Biochemistry 44 (2006) 106–114
Plant Physiology and Biochemistry
www.elsevier.com/locate/plaphy

Research article

Isoflavonoids are present in *Arabidopsis thaliana* despite the absence of any homologue to known isoflavonoid synthases

O. Lapcik [a,*], D. Honys [b,c], R. Koblovska [a], Z. Mackova [a], M. Vitkova [a], B. Klejdus [d]

[a] *Department of Chemistry of Natural Compounds, Faculty of Food and Biochemical Technology, Institute of Chemical Technology, Technicka 5, CZ-166 28 Prague, Czech Republic*
[b] *Institute of Experimental Botany, Academy of Sciences of The Czech Republic, Rozvojova 135, CZ-165 02 Prague, Czech Republic*
[c] *Department of Plant Physiology, Faculty of Sciences, Charles University, Vinicna 5, CZ-128 44 Prague 2, Czech Republic*
[d] *Department of Chemistry and Biochemistry, Mendel University of Agriculture and Forestry, Zemedelska 1, 613 00 Brno, Czech Republic*

Received 30 May 2005; received in revised form 13 June 2005; accepted 25 November 2005
Available online 06 March 2006

Abstract

Extracts from *Arabidopsis thaliana* leaves and inflorescence stalks and from *Lepidium sativa* (Brassicaceae) seedlings were analysed by HPLC-MS-SIM and by five isoflavonoid-specific ELISA methods after the HPLC fractionation of samples, in order to determine presence of isoflavonoids. Individual ELISAs were specific for daidzein, genistein, biochanin A and for their derivatives substituted either at the 4′- or at the 7- positions. Both analytical approaches revealed homologous spectra of isoflavonoids in both plant species. As the ononin specific immunoassay was not available this compound was only detected by HPLC-MS. Formononetin and prunetin represented the main aglycones followed by biochanin A, daidzein and genistein; sissotrin was the most abundant isoflavonoid glycoside followed by ononin, daidzin and genistin. The content of individual compounds ranged from a few micrograms up to 2.2 mg kg^{-1} (dry weight). Expression profiles of *A. thaliana* genes homologous to enzymes involved in isoflavonoid synthesis and metabolism were extracted from publicly available transcriptomic datasets for various tissues. Genes likely to be involved in important steps of the isoflavonoid metabolism in *A. thaliana* were identified. However, in accord with the previously published data, no homologue to isoflavone synthases known from the Fabaceae plants was found. These aryl migrating enzymes belong to the CYP93C family that is absent in *A. thaliana*. We conclude that another gene must be responsible for biosynthesis of the isoflavone skeleton in Brassicaceae.

Keywords: Arabidopsis thaliana; Brassicaceae; HPLC-MS; Immunoassay; Isoflavone; Isoflavone synthase; *Lepidium sativa*

1. Introduction

Due to their numerous biological activities, isoflavones (3-phenylchromones) have been drawing attention of scientists for a long time. They are involved in the interactions between a plant and its environmental partners from bacterial and fungal pathogens up to animal herbivores [6,29]. During last two decades, there has been an increasing interest in health-protecting and health-promoting effects of these compounds and in their possible use in human medicine [4].

The isoflavone synthase (IFS) gene encoding the first enzyme of the isoflavone metabolic pathway has been cloned from several leguminous species and sequenced [1,11,23,26]. Soybean IFS has been expressed in yeast and in several non-leguminous plants, which were considered to be isoflavone non-synthesizing, namely *Arabidopsis thaliana*, *Nicotiana tabacum* and *Zea mays* [11,20,30]. The IFS transgenes not only produced the expected isoflavone aglycone (i.e. genistein) but also its glycosides (namely genistin in *A. thaliana*, genistin and malonyl genistin in *N. tabacum*). This observation led the authors to the conclusion that the endogenous enzymes of non-leguminous plants were able to metabolise genistein as a novel substrate. It was, however, not taken into account, that

Abbreviations: ELISA, enzyme-linked immunosorbent assay; HPLC-MS-SIM, high performance liquid chromatography–mass spectroscopy-SIM mode; IFS, isoflavone synthase; RT-PCR, reverse-transcription polymerase chain reaction.

* Corresponding author. Tel.: +420 22 044 3240; fax: +420 22 044 1082.
E-mail address: oldrich.lapcik@vscht.cz (O. Lapcik).

0981-9428/$ - see front matter © 2006 Elsevier SAS. All rights reserved.
doi:10.1016/j.plaphy.2005.11.006

these supposed isoflavone non-producers may rather be low-level producers [20].

Reports on isoflavonoids in the Brassicaceae family have been limited to only a few notes so far. Mazur [21,22] found low levels of daidzein and genistein in two cultivars of *Brassica oleracea* during his extensive HPLC-MS screening for dietary sources of phytoestrogens; an immunochemical indication of formononetin in *Lepidium sativa* was given by Davidova [7]. Here we bring the evidence of the whole spectrum of isoflavonoids in wild type *A. thaliana* and *L. sativa*. In addition, the overview of *A. thaliana* homologues of genes encoding isoflavonoid biosynthetic enzymes in legumes is given together with their expression profiles based on microarray analyses.

2. Results

2.1. HPLC-ELISA and HPLC-MS analysis of the extracts

Water/ethanolic extracts from both, *Arabidopsis thaliana* and *Lepidium sativa*, were positively tested for several types of isoflavonoid-like immunoreactivity. After HPLC fractionation, the retention times of the immunoreactive entities corresponded to those of authentic standards (Fig. 1). Two independent immunoassays were available for daidzein (lines 1 and 2) and another two for biochanin A (lines 3 and 5). In both cases, the results of parallel assays corresponded satisfactorily. The HPLC-MS-SIM analysis confirmed the presence of all immunoreactive isoflavones and moreover, ononin, for which an immunoassay was not available, was detected (Table 1). Representative chromatograms of *A. thaliana* extract for selected ions are given in Fig. 2. The spectrum of isoflavones was similar in both plants, with the estimated content of individual substances ranging from a few micrograms to about 2.2 mg kg^{-1} (dry weight). The aglycones as well as the 7-*O*-glycosides were present in all samples, and the 4′-methoxy isoflavones were more abundant than 4′-hydroxy isoflavones (Table 2).

2.2. Searching in A. thaliana genome

Protein sequences of genes involved in the isoflavonoid synthesis and metabolism (chalcone synthase, chalcone-flavanone isomerase, chalcone reductase, isoflavone synthase, isoflavone reductase, isoflavone 2′-hydroxylase, isoflavone 7-*O*-methyltransferase) downloaded from GenBank were used for the homology-based search using WU-BLAST2 on TAIR. 51 candidate genes forming eight gene families were identified (Table 2). Publicly available Affymetrix ATH1 microarray-based transcriptomic datasets [5] were used to determine spatial and temporal expression patterns of candidate genes. Expression profiling analyses led to the identification of several genes likely to be involved in the isoflavone synthesis and metabolism in *Arabidopsis*. Out of the 51 candidates, 46 (90%) genes were present on the ATH1 genome array. For these genes, expression profiles in individual organs and tissues were obtained (Table 3). In each gene family, there was at least one

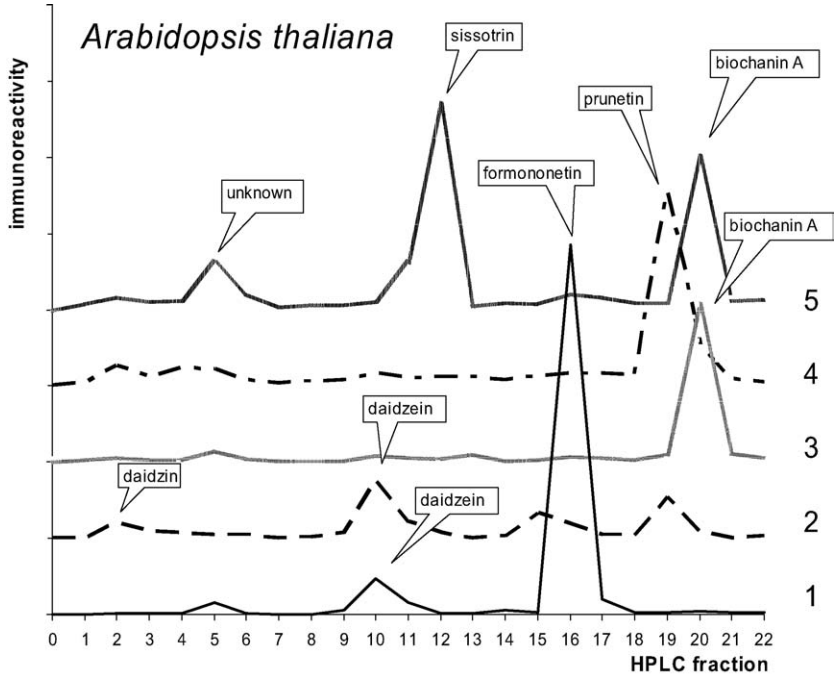

Fig. 1. Isoflavonoid immunoreactivity in HPLC fractions of *A. thaliana* leaves.
Extract from 12.5 mg of *A. thaliana* leaves was chromatographed, the fractions collected, dried and reconstituted in the assay buffer for immunoanalysis. Five ELISAs highly specific to individual isoflavonoid subgroups were used. One segment on the Y axis corresponds to 0.1 mg kg^{-1} of a respective substance in dry matter.
Column: Purospher Star RP18e 125/4 column with a guard column Purospher Star RP18e 4/4 (Merck, Germany). Gradient elution: A (0.05% acetic acid in water) and B (methanol) as followed (all steps linearly): 0 min, A = 60%, B = 40%; 5 min, B = 52%; 20 min, B = 70%; 25 min, B = 100%; 25–45 min, B = 100; then step to A = 60% and reconditioning of the column for 10 min. Flow rate was 0.8 ml min^{-1} and the temperature was set at 25 °C. 1.0 ml fractions were collected.

Table 1
Isoflavonoids identified in the extract from *Arabidopsis thaliana* and their MS-SIM characteristics
The Atlantis dC_{18} (150 × 2.1 mm^2; 3 μm, Waters, Milford, USA) analytical column was used. The gradient elution with mobile phase A (acetonitrile) and B (0.3% formic acid in water) was as follows (all steps linearly): 0 min, A = 12%, B = 88%; 15 min, B = 80%; 32 min, B = 48%; 45 min, B = 12%. Flow rate was 0.2 ml min^{-1}; the column temperature was 36 °C.

Peak number	t_R (min)	$[M + H]^+$ (m/z)	Fragment ion (m/z)	Compound identified
1	14.76	417	255	Daidzin
2	21.87	433	271	Genistin
3	26.39	431	269	Ononin
4	27.50	255	137	Daidzein
5	30.13	447	285	Sissotrin
6	31.97	271	153	Genistein
7	33.96	269	152	Isoformononetin
8	34.81	269	137	Formononetin
9	38.70	285	167	Prunetin
10	39.09	285	153	Biochanin A

gene exhibiting strong and constitutive expression pattern (Fig. 3).

Several putative isoflavone-metabolising genes were more abundantly expressed in roots and stem base than in other sporophytic tissues. This was especially the case for genes *At5g13930* (naringenin-chalcone isomerase), *At3g55120* and *At5g05270* (chalcone-flavanone isomerases), *At1g32100* and *At4g13660* (isoflavone-reductase-related) and *At5g54160* (isoflavone 7-*O*-methyltransferase). Surprisingly, two of these genes, *At5g13930* and *At5g54160* also showed abundant expression in generative organs, namely in flower buds and inflorescences.

3. Discussion

Jung et al. [11,12] observed several isoflavonoid metabolites in *A. thaliana*, tobacco and maize transgenes with IFS gene from soybean (*Glycine max*). However, they did not detect any isoflavones in wild type plants of these species and considered them to be isoflavone non-producers [30,31]. Particularly in the case of *A. thaliana*, for which the genomic data were available, they accented the absence of a sequence with sufficient homology to IFS in this plant [12]. Since other two species of the same family, namely *B. oleracea* and *L. sativa*, were previously shown to contain low levels of isoflavones [7,22] we aimed to analyse *A. thaliana*. Our recent detection of the spectrum of isoflavones in wild type *A. thaliana* is concordance with the above-mentioned data as well as with the observation that *A. thaliana* IFS transgenes were able to further metabolize genistein [11,20]. Individual isoflavones in *A. thaliana* corresponded to those in *L. sativa* qualitatively as well as quantitatively, not exceeding units of mg per kg (dry weight) level. Especially the genistein and genistin levels were low when compared to those of biochanin A, formononetin, prunetin and sissotrin, rarely exceeding 0.1 mg kg^{-1} dry weight. This finding could explain why genistein has not been detected in wild type *A. thaliana* so far, as these levels are close to or below the detection limit of commonly used analytical methods. The spectrum of isoflavones detected in *A. thaliana* implies the presence of necessary biosynthetic apparatus (Fig. 2), particularly of an enzyme with isoflavone synthase activity. Therefore we searched for homologues of the following soybean enzymes using WU-BLAST2 on TAIR: chalcone synthase, chalcone-flavanone isomerase, chalcone reductase, isoflavone synthase, isoflavone reductase, isoflavone 2′-hydroxylase, isoflavone 7-*O*-methyltransferase. We failed to identify a close homologue of IFS in *A. thaliana*. This was in accord with the previously published results [12]. However, we identified 51 genes (Table 3) showing remarkable degree of homology to their proposed leguminous counterparts. Moreover, expression profiling revealed a subset of identified genes that are likely to act as enzymes in isoflavonoid synthesis and metabolism. The abundant expression of several genes encoding putative isoflavone-metabolising enzymes demonstrated in roots is in harmony with the presumed function of their substrates in plant defence. Further experiments will be necessary for final verification of their biochemical activity.

Despite the absence of a homologue to leguminous IFS in *A. thaliana* genome, we undoubtedly detected several isoflavones in this plant. Our finding is supported by the demonstration of isoflavones in other brassicaceous species, namely in *L. sativa* [7] *B. oleracea* [22], *B. napus* and *Raphanus sativus* (preliminary data, not shown). The necessity to explain this discrepancy makes us hypothesize that other gene(s) must be responsible for the aryl migrating activity in Brassicaceae. All IFS genes described up to now belong to the CYP93C family of P450 and they come almost exclusively from legumes [1,23,26,30] with the only exception of sugar beet (*Beta vulgaris*, the Chenopodiaceae) [11]. In phylogenetically distant taxa, the knowledge is limited to the level of metabolites; the genetic data are lacking. Recently, Sawada et al. [25] demonstrated that a single amino acid mutation is able to switch the catalytic activity of IFS from the aryl migration to the 3β-hydroxylation of flavanone skeleton. Moreover, certain mutants produced both types of metabolites, i.e. isoflavones and 3-hydroxyflavanones. The latter type of reaction has been assigned to non-P450 enzymes, namely to 2-oxoglutarate-dependent dioxygenases, in several recent plants [3,24]. Sawada, however, hypothesized that flavanone 3β-hydroxylase of the P450 type was present in past plant species, and could be the evolutionary predecessor of IFS. He expressed a further hypothesis that the aryl migration activity of P450 IFS emerged through a rare gain-of-function event by mutations from ancestral *CYP*93 genes, and have survived and become dominant in present-day legumes. Extrapolating this hypothesis, we may further speculate that:

- the aryl migration activity may have developed several times independently in the phylogeny of higher plants. Would it be true, than the enzyme(s) responsible for the aryl migration in phylogenetically distant taxa need not display a noticeable degree of homology to the known CYP93C IFS enzymes from legumes;

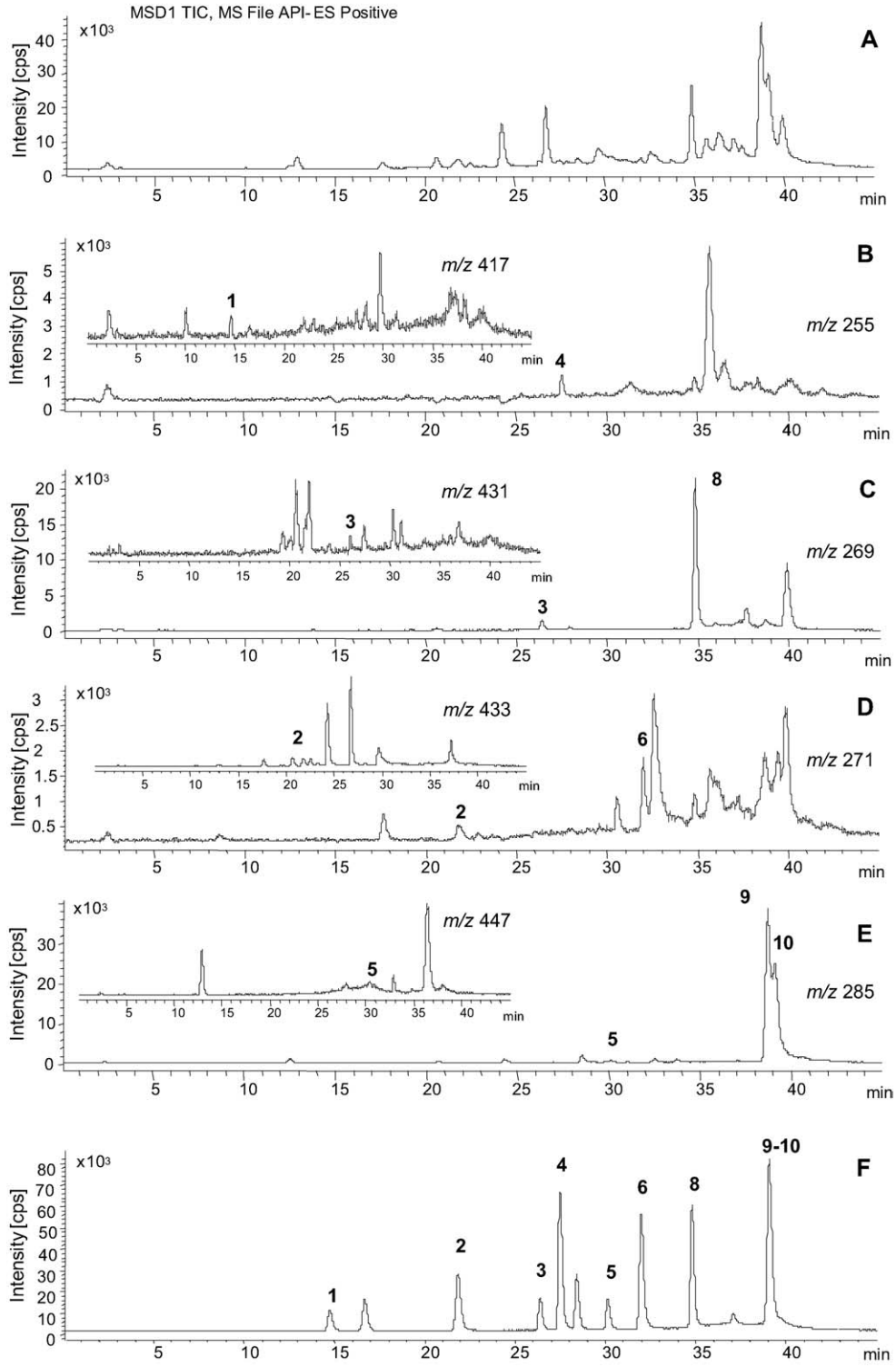

Fig. 2. HPLC/MS chromatograms of an extract from *A. thaliana* stalks.
The Atlantis dC_{18} (150 × 2.1 mm^2; 3 μm, Waters, Milford, USA) analytical column was used. The gradient elution with mobile phase A (acetonitrile) and B (0.3% formic acid in water) was as follows (all steps linearly): 0 min, A = 12%, B = 88%; 15 min, B = 80%; 32 min, B = 48%; 45 min, B = 12%. Flow rate was 0.2 ml min^{-1}; the column temperature was 36 °C.
A: TIC chromatogram of *A. thaliana* extract, B–E: SIM chromatograms for selected *m/z*, F: TIC chromatogram of standards. Peak identification: 1 – daidzin, 2 – genistin, 3 – ononin, 4 – daidzein, 5 – sissotrin, 6 – genistein, 8 – formononetin, 9 – prunetin, 10 – biochanin A.

- the possibility cannot be excluded in advance, that aryl migration may be a collateral activity of some enzyme(s) known to catalyse another reaction(s) in certain taxa.

Generally low levels of isoflavones in *Arabidopsis* together with significantly lower levels of genistein and daidzein, the first isoflavones synthesized, compared to their metabolites

Table 2
Isoflavonoids in *Arabidopsis thaliana* and *Lepidium sativa*
L: leaves, S: stalks, D: seedlings. With the exception of ononin, only the isoflavones detected consistently by both, HPLC-ELISA and HPLC-MS-SIM, are listed here. Numeric values are given in μg g^{-1} (dry weight). MS: the quantitative data obtained by HPLC-MS-SIM; ELISA: the quantitative data obtained by HPLC-ELISA; tr.: traces; n.d.: not detected.

	Arabidopsis thaliana							*Lepidium sativa*	
Sample number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Tissue	L	L	S	S	L	L	S	D	D
Method	MS	MS	MS	MS	ELISA	ELISA	ELISA	MS	ELISA
Daidzin	0.19	0.11	–	0.08	0.13	0.03	0.78	0.20	n.d.
Genistin	0.51	1.01	0.14	0.23	0.09	0.06	0.32	n.d.	0.09
Ononin	0.53	0.28	0.13	0.13	–	–	–	0.36	–
Daidzein	0.01	0.04	–	0.02	0.20	0.07	0.08	–	0.39
Sissostrin	1.60	0.57	0.49	0.58	0.63	0.34	1.54	0.20	2.16
Genistein	0.04	0.04	0.01	0.05	0.01	0.02	0.04	0.06	0.05
Isoformononetin	tr.	tr.	tr.	tr.	0.06	0.05	0.12	tr.	0.08
Formononetin	0.75	0.38	0.24	0.64	0.65	0.50	0.90	0.54	1.01
Prunetin	0.87	0.42	0.07	0.42	1.00	0.32	1.76	–	0.92
Biochanin A	0.33	0.11	0.11	0.60	0.22	0.22	0.91	0.21	0.50

strongly suggest that the isoflavone synthesis represents the rate-limiting step in this metabolic pathway. The proposed weak activity of isoflavone-synthesising enzyme(s) in *Arabidopsis* is in accord with both presented speculations.

Additional research will be necessary for elucidating the genetic background of aryl migration in non-legumes. To our best knowledge this is the first report of native isoflavonoids in wild type *A. thaliana*.

4. Materials and methods

4.1. Chemicals

Daidzein, genistein and genistin were purchased from Sigma (St. Louis, Missouri, USA), isoformononetin was prepared by partial methylation of daidzein as described elsewhere [14], and all other isoflavones were obtained from Indofine (Somerville, New Jersey, USA). Organic solvents were from Merck (Darmstadt, Germany). All other substances were of analytical grade, from Pliva-Lachema (Brno, Czech Republic). The immunochemicals were prepared by us [14,28].

4.2. Plant material

A. thaliana ecotype Col-0 plants were grown in controlled-environment cabinets at 21 °C under illumination of 150 μmol m^{-2} s^{-1} with a 16-hour photoperiod. Leaves and inflorescence stalks were collected at the growth stage 5.1 [2].

The seeds of *L. sativa* var. *capitata* were obtained from SEMO s.r.o., (Smrzice, Czech Republic). Seeds were germinated 7 days on a sheet of filtration paper using tap water for their moisturising once per day. Samples were taken every day and kept deep frozen until lyophilised.

The dry material was disintegrated in Fex IKA A11 Basic grinder (IKA-Werke, GmbH and Co, KG, Staufen, Germany) and extracted with 80% ethanol (20 ml g^{-1}). The extracts were centrifuged and filtered to remove the debris and the supernatant was analysed. Substantial attention was paid to the effort to avoid the risk of cross-contamination of the samples resulting eventually in false positive findings. Except for the lyophilisation, all preanalytical procedures were done independently in two different labs for HPLC-MS and for HPLC-ELISA, respectively. A sheet of filter paper was processed in the same way as the samples in order to get a blank.

4.3. Analytical methods

All analytical methods were described in separate papers. The recently described ELISA methods [28] are non-radioactive variants of our previously published radioimmunoassays [14,16]. The methods for genistein and daidzein are highly selective for the abovementioned isoflavones and their derivatives either at the carbon 4′ (i.e. biochanin A and formononetin, respectively) or at the carbon 7 (i.e. prunetin and genistin, isoformononetin and daidzin, respectively). The method for biochanin A is also able to detect its derivatives at the position 7- (e.g. sissotrin). No cross-reactivity was recorded for flavonoids, lignans, chalcones and steroids. Two HPLC systems were used for fractionation of the extracts before the ELISA analysis: System 1—Purospher Star RP18e 125/4 column (Merck, Germany), linear gradient water/methanol [15]; System 2—MetaChem Polaris C18A 250 × 9.6 mm^2 semi-preparative column (MetaChem Technologies, Torrance, CA, USA), linear gradient water-acetonitrile [17]. The HPLC-MS system is specified in references [15,17], the molecular ions $[M + H]^+$ and characteristic fragments of individual isoflavones are listed in Table 1.

4.4. Searching *A. thaliana* genes

Sequences of enzymes involved in isoflavonoids synthesis and metabolism (chalcone synthase, chalcone-flavanone isomerase, chalcone reductase, isoflavone synthase, isoflavone reductase, isoflavone 2′-hydroxylase, isoflavone 7-*O*-methyl-

Table 3
Expression profiles of *Arabidopsis thaliana* homologues of genes involved in isoflavones synthesis and metabolism in particular organs and tissues based on microarray transcriptomic experiments. Transcriptome datasets were downloaded from the NASCArrays microarray database (http://www.arabidopsis.info). All genes were annotated according to The Institute for Genomic Research (TIGR; www.tigr.org). The data were normalised using dChip 1.3 freeware. Resulting expression signal value represents the mean value of all normalized experiments. Particular gene was scored as "expressed" when it gave a reliable expression signal in all replicates. Genes were given the signal value "0" when their detection call value was not "present" in all replicates. Dataset labels are as follows: SL, seedlings; WP, whole plants at rosette developmental stage; SH, shoots; LF, rosette leaves; GC, guard cell-enriched cell extract; PT, petioles; ST, stems; IN, inflorescences; BY, young flower buds; BO, old flower buds; HP, hypocotyls; XL, xylem; CO, cork; RH, root hair elongation zone; RT, roots; SU, suspension cell cultures. Dataset codes for all 90 transcriptomic datasets downloaded from the NASC website are available as supplementary information

Gene	TIGR annotation	SL	WP	SH	LF	GC	PT	ST	IN	BY	BO	HP	XL	CR	RH	RT	SU
Chalcone + stilbene synthase																	
At1g02050	chalcone and stilbene synthase family	7,96	0	0	0	0	0	0	553,64	1598,61	204,06	156,35	166,17	188,01	0	0	0
At4g00040	chalcone and stilbene synthase family	29,96	121,46	0	71,93	0	0	342,24	174,26	148,6	125,65	509,63	693,97	650,64	127,9	71,55	106,81
At4g34850	chalcone and stilbene synthase family	0	0	0	0	0	0	0	575,21	1633,71	0	0	0	0	0	0	0
Naringenin/chalcone synthase																	
At5g13930	naringenin-chalcone synthase	2727,07	621,04	620,43	1364,4	1168,31	317,26	0	2974,86	3282,56	4089,75	39,17	0	0	6881,39	1168,54	0
Chalcone-flavanone isomerase																	
At1g53520	chalcone-flavanone isomerase-related	563,42	429,54	272,34	407,04	234,11	423,9	0	382,43	405,6	400,53	55,72	0	60,11	173,1	185,32	106,02
At3g55120	chalcone-flavanone isomerase	634,77	285,21	330,25	478,67	431,25	213,35	0	682,86	783,85	1009,76	129,34	0	162,4	1850,46	326,04	0
At5g05270	putative protein	522,98	171,28	0	372,05	0	0	0	567,54	551,73	609,98	0	0	0	1131,38	258,29	0
At5g66220	chalcone-flavanone isomerase	0	0	0	28,62	0	0	0	70,17	119,49	0	0	0	0	0	0	0
Chalcone reductase																	
At2g37760	aldo/keto reductase family	597,3	399,18	714,91	335,99	193,75	48,23	403,52	272,79	539,57	707,99	553,2	670,86	512,91	80,24	377,64	1180,47
At2g37770	aldo/keto reductase family	92,05	138,75	359,38	109,53	0	0	0	0	0	264,08	141,2	0	131,49	0	35,38	413,92
At2g37790	aldo/keto reductase family	204,64	222,53	102,84	233,02	206,55	341,65	183,13	434,45	407,91	361,29	179,33	0	116,13	170,6	251,12	205,38
At3g53880	aldo/keto reductase family	203,42	192,1	160,48	165,89	147,94	243,42	142,36	277,38	215,33	187,87	155,88	142,44	116,11	254,67	245,36	326,46
At5g62420	aldo/keto reductase family	0	0	0	0	0	0	0	0	0	0	0	30,32	0	0	21,13	0
Isoflavone reductase																	
At1g19540	isoflavone reductase homolog P3	126	94,2	230,06	55,08	111,61	78,29	600,58	102,73	168,55	108,13	1221,19	1692,73	1280,21	107,5	107,53	197,78
At1g75260	isoflavone reductase family	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
At1g75280	isoflavone reductase homolog P3	1527,37	1089,6	1323,31	1474,11	872,46	959,31	5715,38	946,4	754,52	1039,3	919,89	1032,64	238,02	751,38	783,46	406,82
At1g75290	isoflavone reductase homolog P3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
At1g75300	isoflavone reductase homolog P3	0	0	0	0	0	0	0	161,75	0	295,8	0	0	0	0	0	0
At4g34540	isoflavone reductase family	233,39	267,09	265,43	266,64	359,61	241,69	280,52	307,18	262,19	218,81	259,76	272,34	218,95	193,81	272,08	412,98
At4g39230	isoflavone reductase, putative	146,42	161,02	181,59	136,74	177,69	181,66	167,92	206,94	203,94	314,28	426,7	378,95	348,24	215,24	219,13	359,95
At5g18660	isoflavone reductase, putative	1154,61	1029,56	526,03	1198,76	585,09	733,08	0	947,84	633,16	773,95	0	0	0	0	0	52,43

(continued)

Table 3 *(continued)*

Gene	TIGR annotation	SL	WP	SH	LF	GC	PT	ST	IN	BY	BO	HP	XL	CR	RH	RT	SU
Isoflavone reductase-related, contain IFR domain																	
At1g32100	pinoresinol-lariciresinol reductase, putative	345,91	303,45	143,98	103,85	0	248,1	2633,63	237,21	98,28	143,68	516,41	867,16	36,74	1011,62	443,18	45,73
At4g13660	pinoresinol-lariciresinol reductase, putative	571,45	414,27	102,54	7,56	0	0	0	0	0	0	199,38	53,7	182,36	1875,75	1181,73	256,6
Isoflavone 2'hydroxylase																	
At4g37370	cytochrome p450 family	31,77	164,17	418,72	109,37	0	169,08	443,52	0	0	0	230,44	210,69	156,26	0	106,17	407
At4g37340	cytochrome p450 family	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20,65	0
At4g37330	cytochrome p450 family	391,33	189,35	199,08	259,15	279,62	223,86	173,7	302,13	488,68	468,25	86,46	74,26	71,74	204,56	305,81	0
At3g28740	cytochrome p450 family	800,14	376,87	679,69	65,96	0	0	135,98	82,49	152,4	177,41	58,03	0	0	0	290,66	0
At4g37360	cytochrome p450 family	0	79,89	46,39	23,06	0	0	57,61	0	42,16	0	0	50,66	36,09	0	0	58,23
At4g37320	cytochrome p450 family	137,7	0	210,99	170,97	0	0	0	0	0	0	0	0	0	471,75	258,86	0
At5g36220	cytochrome p450 family	15,68	0	148	172,13	428,59	67,29	0	0	0	0	599,32	633,26	864,98	0	42,56	0
At2g23190	cytochrome p450 family	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
At2g23220	cytochrome p450 family	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
At1g66540	cytochrome p450 family	65,9	48,56	19,24	91,37	100,81	144,62	70,59	103,22	0	63,42	17,84	0	42,49	56,17	20,33	42,64
Isoflavone 7-O-methyltransferase																	
At4g35160	*O*-methyltransferase family 2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
At5g54160	*O*-methyltransferase 1	3142,68	1951,78	1515,53	1361,9	1024,16	2094,1	7058,2	1364,62	2145,68	2375,08	3623,96	4280,19	2868,7	6820,87	3229,44	1821,06
At4g35150	*O*-methyltransferase family 2	243,86	241,8	0	0	0	0	0	34,06	0	78,34	31,79	0	64,73	337,92	496,29	0
At1g51990	*O*-methyltransferase family 2	0	0	0	0	0	0	0	0	0	0	36,65	0	0	0	62,57	0
At3g53140	*O*-diphenol-O-methyl transferase, putative	43,78	0	0	0	0	0	0	429,67	504,07	448,93	0	0	0	0	0	118,67
At1g33030	*O*-methyltransferase family 2	0	94,73	190,96	48,21	137,56	0	0	0	0	0	213,75	156,28	313,07	0	45,93	566,99
At1g77520	*O*-methyltransferase family 2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
At1g77530	*O*-methyltransferase family 2	20,27	150,07	0	0	0	0	0	53,13	0	56,87	157,74	0	259,03	0	301,46	0
At5g53810	*O*-methyltransferase 1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
At1g63140	putative *O*-methyltransferase	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
At1g21100	*O*-methyltransferase 1	89,54	388,26	238,64	55,86	0	0	0	0	0	0	230,84	93,35	237,73	111,56	521,74	150,62
At1g21130	*O*-methyltransferase 1	408,97	1239,73	387,5	408,94	0	1030,07	240,8	0	0	0	63,63	44,23	237,36	0	72,56	0
At1g76790	*O*-methyltransferase family 2	890,45	1162,35	429,37	140,68	0	148,03	0	80,84	62,24	141,62	332,78	0	257,35	1920,95	1117,33	25,14
At1g21110	putative *O*-methyltransferase	197,66	240,89	93,36	17,49	0	0	108,74	0	0	0	160,76	144,66	137,81	250,27	231,09	0
At1g62900	putative *O*-methyltransferase	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
At1g21120	*O*-methyltransferase 1	7,04	98,36	389,17	103	387,07	0	234,94	0	43,34	0	53,9	16,39	38,55	39,44	141,82	0
At5g37170	*O*-methyltransferase family 2	38,82	0	79,67	16,42	0	0	0	158,84	0	136,38	0	0	0	0	39,26	0

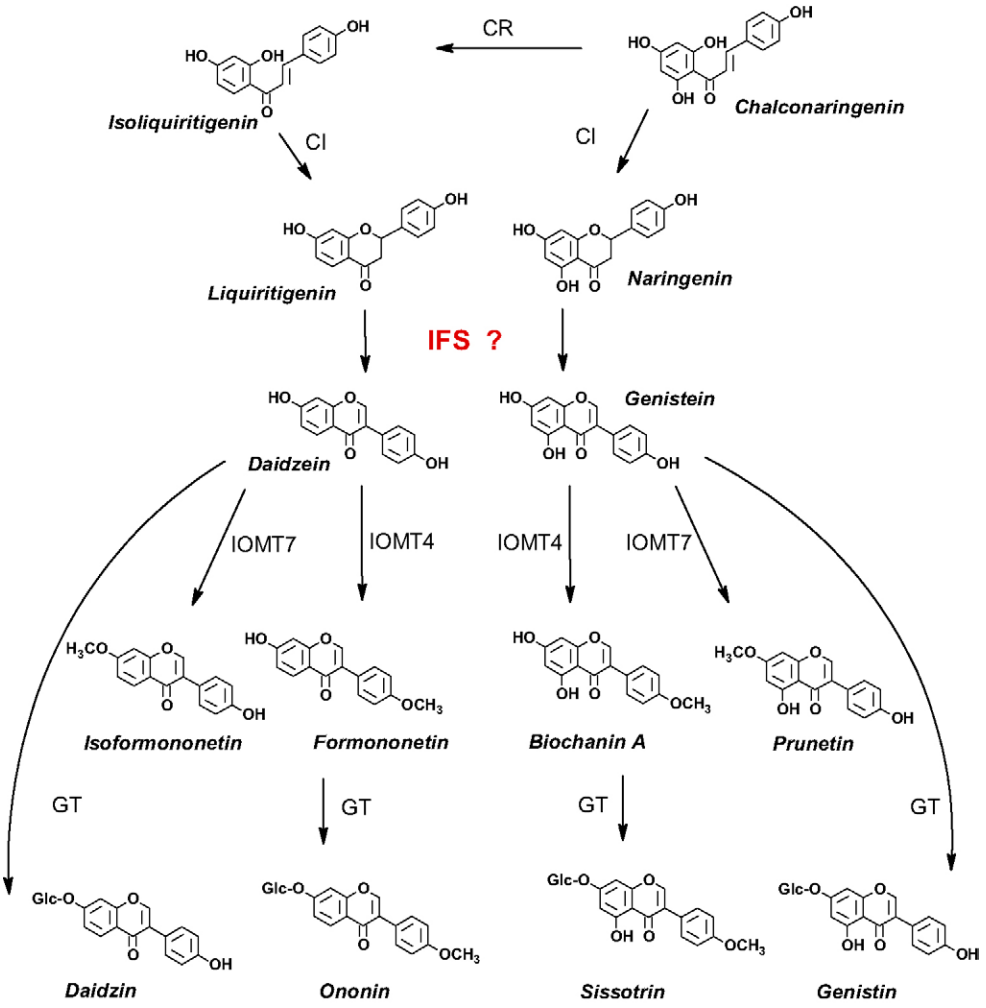

Fig. 3. Proposed scheme of isoflavonoid metabolism in *Arabidopsis thaliana*.
CR: Chalcone reductase, CI: Chalcone-flavanone isomerase, IFS: Isoflavone synthase, IOMT4: Isoflavone 4′-*O*-methyltransferase , IOMT7: Isoflavone 7-*O*-methyltransferase, GT: Isoflavone glucosyl-transferase.

transferase) downloaded from GenBank (http://www.ncbi.nlm.nih.gov) were used to conduct WU-BLAST2 [8] against Arabidopsis Annotated Protein Database at the TAIR (http://www.arabidopsis.org/wublast/index2.jsp).

4.5. GeneChip analysis

Transcriptome datasets from public baseline GeneChip experiments used for the expression analysis of the identified genes were downloaded from the NASCArray microarray database [5] (http://www.arabidopsis.info). In order to make the data from all samples comparable, individual slides were normalised using freely available dChip 1.3 software (www.dchip.org). Reliability and reproducibility of the analyses were ensured by :

- the use of duplicates or triplicates in each experiment ;
- normalization of all 90 arrays to the median probe intensity level ;
- and the use of normalized CEL intensities of all arrays for the calculation of model-based gene expression values based on the Perfect Match-only model [18,19].

A gene was scored as “expressed” when it gave a reliable expression signal in all replicates. Expression signal value “0” means that the detection call value was not “present” in all replicates provided. Although a RT-PCR validation of microarray data was not performed specifically for the purpose of this publication, our confidence in the quality of the data presented is based on our previously published RT-PCR validation of expression of 70 genes [9,10,13,27]. Resulting expression signal represents the mean value of all normalized experiments. Dataset codes for all 90 transcriptomic datasets downloaded from the NASC website are available as supplementary information.

Acknowledgments

This study was supported by GACR 525/03/0352 and GAAVCR KJB6038409 (DH) Grants.

References

[1] T. Akashi, T. Aoki, S. Ayabe, Cloning and functional expression of a cytochrome P450 cDNA encoding 2-hydroxyisoflavanone synthase involved in biosynthesis of the isoflavonoid skeleton in Licorice, Plant Physiol. 121 (1999) 821–828.

[2] D.C. Boyes, A.M. Zayed, R. Ascenzi, A.J. McCaskill, N.E. Hoffman, K.R. Davis, J. Gorlach, Growth stage-based phenotypic analysis of *Arabidopsis*: a model for high throughput functional genomics in plants, Plant Cell 13 (2001) 1499–1510.

[3] L. Britsch, L. Ruhnau-Brich, G. Forkman, Molecular cloning, sequence analysis, and in vitro expression of flavanone 3β hydroxylase from *Petunia hybrida*, J. Biol. Chem. 267 (1992) 5380–5387.

[4] P. Cos, T. De Bruyne, S. Apers, D. Vanden Berghe, L. Pieters, A.J. Vlietinck, Phytoestrogens, recent developments, Planta Medica 69 (2003) 589–599.

[5] D.J. Craigon, N. James, J. Okyere, J. Higgins, J. Jotham, S. May, NASCArrays: a repository for microarray data generated by NASC's transcriptomic service, Nucleic Acids Res. 32 (2004) D575–D577 (Database issue).

[6] F.D. Dakora, D.A. Phillips, Diverse functions of isoflavonoids in legumes transcend anti-microbial definitions of phytoalexins, Physiol. Mol. Plant Pathol. 49 (1996) 1–20.

[7] M. Davidova, Imunochemicky skrining isoflavonoidu u vybranych druhu hospodarskych plodin tropu a subtropu, Diploma Thesis, Institiute of Tropical and Subtropical Agriculture, Czech University of Agriculture in Prague, (2002) pp. 44–45.

[8] W. Gish, (1996–2004) http://blast.wustl.edu.

[9] D. Honys, D. Twell, Comparative analysis of the *Arabidopsis* pollen transcriptome, Plant Physiol. 132 (2003) 640–652.

[10] P. Hruba, D. Honys, D. Twell, V. Capkova, J. Tupy, Expression of β-galactosidase and β-xylosidase genes during microspore and pollen development, Planta 220 (2005) 931–940.

[11] W. Jung, O. Yu, S.C. Lau, D. O'Keefe, J. Odell, G. Fader, B. McGonigle, Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes, Nat. Biotechnol. 18 (2000) 208–212.

[12] W. Jung, I.M. Chung, H.Y. Heo, Manipulating isoflavone levels in plants, J. Plant Biotechnol. 5 (2003) 149–155.

[13] E. Lalanne, D. Honys, A. Johnson, G.H.H. Borner, K.S. Lilley, P. Dupree, U. Grossniklaus, D. Twell, SETH1 and SETH2, two components of the glycosylphosphatidylinositol anchor biosynthetic pathway, are required for pollen germination and tube growth in *Arabidopsis*, Plant Cell 16 (2004) 229–240.

[14] O. Lapcik, M. Hill, I. Cerny, J. Lachman, N. Al-Maharik, H. Adlercreutz, R. Hampl, Immunoanalysis of isoflavonoids in *Pisum sativum* and *Vigna radiata*, Plant Sci. 148 (1999) 111–119.

[15] O. Lapcik, J. Stursa, T. Kleinova, M. Vitkova, H. Dvorakova, B. Klejdus, J. Moravcova, Synthesis of hapten and conjugates of coumestrol and development of immunoassay, Steroids 68 (2003) 1147–1155.

[16] O. Lapcik, M. Vitkova, B. Klejdus, N. Al-Maharik, H. Adlercreutz, Immunoassay for biochanin A, J. Immunol. Methods 94 (2004) 155–163.

[17] O. Lapcik, B. Klejdus, M. Davidova, L. Kokoska, V. Kuban, J. Moravcova, Isoflavonoids in the Rutaceae family: 1. *Fortunella obovata, Murraya paniculata* and four *Citrus* species, Phytochem. Anal. 15 (2004) 293–299.

[18] C. Li, W.H. Wong, Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection, Proc. Natl. Acad. Sci. USA 98 (2001) 31–36.

[19] C. Li, W.H. Wong, Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application, Geonome Biology 2 research (2001) (0032.1–0032.11).

[20] C.-J. Liu, J.W. Blount, C.L. Steele, R.A. Dixon, Bottlenecks for metabolic engineering of isoflavone glycoconjugates in *Arabidopsis*, Proc. Natl. Acad. Sci. USA 99 (2002) 14578–14583.

[21] W. Mazur, Phytoestrogens in food, in: H. Adlercreutz (Ed.), Baillière's clinical endocrinology and metabolism: phytooestrogens. Vol. 12, Baillière Tindall, London, 1998, pp. 729–742.

[22] W. Mazur, Phytoestrogens: occurrence in foods and metabolism of lignans in man and pigs, Ph.D.Thesis, University of Helsinki, Division of Clinical Chemistry (2000) pp. 80.

[23] S. Overkamp, F. Hein, W. Barz, Cloning and characterization of eight cytochrome P450 cDNAs from chickpea (*Cicer arietinum* L.) cell suspension cultures, Plant Sci. 155 (2000) 101–108.

[24] J.L. Pelt, W.A. Downes, R.W. Schoborg, C.A. McIntosh, Flavanone 3-hydroxylase expression in *Citrus paradisi* and *Petunia hybrida* seedlings, Phytochemistry 64 (2003) 435–444.

[25] Y. Sawada, K. Kinoshita, T. Akashi, T. Aoki, S. Ayabe, Key amino acid residues required for aryl migration catalysed by the cytochrome P450 2-hydroxyisoflavanone synthase, Plant J. 31 (2002) 555–564.

[26] C.L. Steele, M. Gijzen, D. Qutob, R.A. Dixon, Molecular characterization of the enzyme catalyzing the aryl migration reaction of isoflavonoid biosynthesis in soybean, Arch. Biochem. Biophys. 367 (1999) 146–150.

[27] H. Sze, S. Padmanaban, F. Cellier, D. Honys, N.-H. Cheng, K.W. Bock, G. Conéjéro, X. Li, D. Twell, J.M. Ward, K.D. Hirschi, Expression patterns of a novel AtCHX gene family highlight potential roles in osmotic adjustment and K^+ homeostasis in pollen development, Plant Physiol. 136 (2004) 2532–2547.

[28] M. Vitkova, Z. Mackova, L. Fukal, O. Lapcik, Enzyme immunoassay for the determination of isoflavones, Chem. Listy 98 (2004) 1135–1139.

[29] S.F. Wang, T.J. Ridsdill-Smith, E.L. Ghisalberti, Role of isoflavonoids in resistance of subterranean clover trifoliates to the redlegged earth mite *Halotydeus destructor*, J. Chem. Ecol. 24 (1998) 2089–2100.

[30] O. Yu, W. Jung, J. Shi, R.A. Croes, G.M. Fader, B. McGonigle, J.T. Odell, Production of the isoflavones genistein and daidzein in non-legume dicot and monocot tissues, Plant Physiol. 124 (2000) 781–793.

[31] O. Yu, J. Shi, A.O. Hession, C.A. Maxwell, B. McGonigle, J.T. Odell, Metabolic engineering to increase isoflavone biosynthesis in soybean seed, Phytochemistry 63 (2003) 753–763.