



## Flavonol glycosides acylated with 3-hydroxy-3-methylglutaric acid as systematic characters in *Rosa*

Elaine A. Porter, Alexander A. van den Bos, Geoffrey C. Kite, Nigel C. Veitch, Monique S.J. Simmonds\*

Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK

### ARTICLE INFO

#### Article history:

Received 9 October 2011

Received in revised form 10 April 2012

Available online 20 June 2012

#### Keywords:

*Rosa spinosissima*

*Rosa*

*Rubus*

Rosaceae

Acylated flavonol glycosides

3-Hydroxy-3-methylglutaric acid

MS/MS analysis

### ABSTRACT

LC–UV–MS/MS analysis of leaf extracts from 146 accessions of 71 species of *Rosa* revealed that some taxa accumulated flavonol *O*-glycosides acylated with 3-hydroxy-3-methylglutaric acid, which are relatively uncommon in plants. The structures of two previously unrecorded examples isolated from *Rosa spinosissima* L. (syn. *Rosa pimpinellifolia* L.) were elucidated using spectroscopic and chemical methods as the 3-*O*- $\alpha$ -L-rhamnopyranosyl-(1  $\rightarrow$  2)-[6-*O*-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranosides] of kaempferol (3,5,7,4'-tetrahydroxyflavone) and quercetin (3,5,7,3',4'-pentahydroxyflavone). The corresponding 3-*O*-[6-*O*-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranoside] of quercetin was also present in *R. spinosissima*, but at lower levels, together with 17 other flavonol *O*-glycosides for which structures were assigned using LC–UV–MS/MS. The distribution of flavonol 3-hydroxy-3-methylglutaryl-galactosides in *Rosa* was limited to some species of subgenus *Rosa* section *Pimpinellifoliae* and *Rosa roxburghii* Sw. of the monotypic subgenus *Platyrhodon*, indicating that this character could be of value in phylogenetic analyses of the genus.

© 2012 Published by Elsevier Ltd.

### 1. Introduction

The taxonomy of *Rosa*, comprising nearly 200 species (Wisseman and Ritz, 2005), is problematic due to intra-specific variability, polyploidy and inter-specific hybridization (Crépin, 1893; Erlanson, 1929; Melville, 1967; Wissemann and Ritz, 2005). The classification adopted by Rehder (1940) has been widely used and was recently updated by Wissemann (2003). This divides *Rosa* into four subgenera, two of which are monotypic (*Hulthemia* and *Platyrhodon*) while subgenus *Hesperhodos* contains just two species. Thus the majority of species are placed in subgenus *Rosa*, which is divided into ten sections. Attempts at phylogenetic analyses based on DNA sequence data, however, neither fully support this traditional classification nor do they provide a consensus among themselves, and they also suggest that several sections are polyphyletic (Bruneau et al., 2007; Koopman et al., 2008; Wissemann and Ritz, 2005). This emphasises the need for additional systematically-informative characters to advance taxonomic studies of this difficult genus.

In the present work, LC–UV–MS/MS analysis of leaf extracts from 146 accessions of 71 species of *Rosa* was undertaken to assess the value of flavonoids as chemical characters. Some species were found to contain flavonol *O*-glycosides acylated with 3-hydroxy-3-methylglutaric acid, examples of which are of limited occurrence

in plants. A useful source of these compounds was identified as *Rosa spinosissima* L. (syn. *Rosa pimpinellifolia* L.; The Plant List, 2010), which is commonly known as the Burnet Rose, Pimpernel Rose or Scots Rose. The hips of this species, which is native to the British Isles, are collected from wild and cultivated specimens for use in cosmetics, and are therefore of economic value.

The major flavonoids present in leaf extracts of *R. spinosissima* were found to be the 3-*O*- $\alpha$ -L-rhamnopyranosyl-(1  $\rightarrow$  2)-[6-*O*-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranosides] of quercetin (**4**) and kaempferol (**10**), which are new compounds (Fig. 1). Among the minor flavonoids were 18 flavonol *O*-glycosides, 6 of which were acylated with 3-hydroxy-3-methylglutaric acid, including the 3-*O*-[6-*O*-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranoside] of quercetin (**11**) and **19**, which was tentatively assigned as the kaempferol analogue of **11**. Here we describe the detection and characterisation of flavonol *O*-glycosides in LC–UV–MS/MS analyses of *R. spinosissima*, the structural elucidation of **4** and **10** using spectroscopic and chemical methods, and the distribution of **4**, **10**, **11** and **19** in *Rosa*. The potential of the latter as chemical characters in phylogenetic analyses of the genus is also discussed.

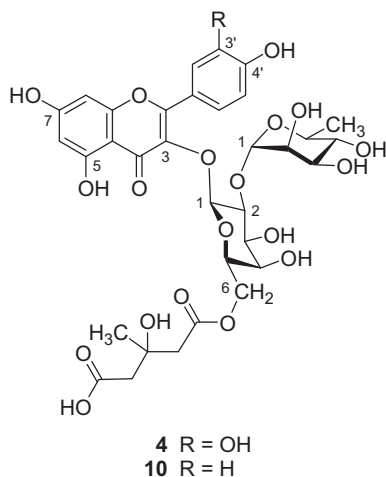
### 2. Results and discussion

#### 2.1. Flavonol glycosides detected in leaves of *R. spinosissima*

LC–UV–MS analysis of an 80% MeOH extract of leaves of *R. spinosissima* revealed the presence of flavonol *O*-glycosides, 20 of

\* Corresponding author. Tel.: +44 208 332 5328; fax: +44 208 332 5310.

E-mail addresses: [m.simmonds@kew.org](mailto:m.simmonds@kew.org), [monique.simmonds@ntlworld.com](mailto:monique.simmonds@ntlworld.com) (M.S.J. Simmonds).

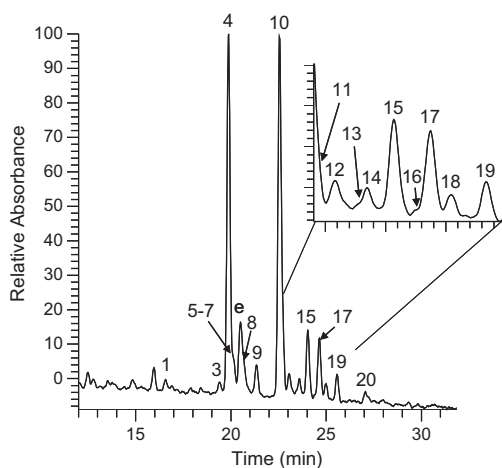


**Fig. 1.** Flavonol 3-*O*- $\alpha$ -L-rhamnopyranosyl-(1  $\rightarrow$  2)-[6-*O*-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranosides] from *Rosa spinosissima*.

which were identified (Fig. 2) either by isolation and NMR spectroscopy (**4**, **9–11**) or by LC–UV–MS techniques, making comparisons with authentic standards when available, as indicated in Table 1. An extract of the petals of *R. spinosissima* was used as a source of the 4'-*O*-glucosides of kaempferol and quercetin (Mikanagi et al., 1995) to provide comparative LC–UV–MS data for **17** and **15**, respectively. Eight of the flavonol *O*-glycosides were acylated with 3-hydroxy-3-methylglutaric acid, and are discussed further in Section 2.2. The non-acylated examples comprised the 3-*O*-galactosides, 3-*O*-glucosides, 3-*O*-glucuronides and 4'-*O*-glucosides of kaempferol and quercetin, and the 3-*O*-rhamnosyl-(1  $\rightarrow$  2)-galactosides and 3-*O*-rhamnosyl-(1  $\rightarrow$  2)-glucosides (neohesperidosides) of the same aglycones.

## 2.2. Characterisation of flavonol *O*-glycosides acylated with 3-hydroxy-3-methylglutaric acid

The UV spectra of **4** (256, 296sh, 356 nm) and **10** (264, 292sh, 350 nm), the major flavonoid constituents of the 80% aq. MeOH extract of leaves of *R. spinosissima*, were characteristic of 3-*O*-glycosides of quercetin and kaempferol, respectively (Markham, 1982).



**Fig. 2.** UV absorbance (350 nm) chromatogram from an LC–UV–MS analysis of a 80:20 MeOH/H<sub>2</sub>O extract of leaves of *Rosa spinosissima* (BI-19505). Peak numbers refer to flavonoids listed in Table 2, e = ellagic acid.

The later eluting kaempferol 3-*O*-glycoside (**10**) gave a deprotonated molecule at *m/z* 737 which fragmented following MS<sup>2</sup> to give a major product ion at *m/z* 593; accurate mass measurements of [M–H]<sup>–</sup> and *m/z* 593 gave a formula of C<sub>6</sub>H<sub>8</sub>O<sub>4</sub> for this neutral loss of 144. Negative ion MS<sup>3</sup> (*m/z* 737  $\rightarrow$  593) of **10** gave a spectrum typical of deprotonated kaempferol 3-*O*-rhamnosyl-(1  $\rightarrow$  2)-hexosides in showing an abundant rhamnose-loss ion at *m/z* 429, indicative of a (1  $\rightarrow$  2) interglycosidic linkage, as well as radical and rearrangement kaempferol aglycone ions at *m/z* 284 and 285, respectively; fragmentation of *m/z* 429 gave a spectrum with high abundance of an ion at *m/z* 339 relative to other ions which indicated that the primary hexose was galactose (Kite and Veitch, 2009). On the basis of these data, **10** was a derivative of kaempferol 3-*O*-rhamnosyl-(1  $\rightarrow$  2)-galactoside. For the corresponding quercetin 3-*O*-glycoside (**4**), negative ion MS gave a deprotonated molecule at *m/z* 753 which, like **10**, lost C<sub>6</sub>H<sub>8</sub>O<sub>4</sub> following accurate mass MS<sup>2</sup>. Serial MS analysis of [(M–H)–144]<sup>–</sup> and the resulting rhamnose-loss ion produced spectra that suggested **4** was a derivative of quercetin 3-*O*-rhamnosyl-(1  $\rightarrow$  2)-galactoside, using the reasoning applied to **10**.

Amounts of **4** and **10** suitable for further analysis were obtained from a larger scale extraction of *R. spinosissima* leaves in 80% aq. MeOH, followed by flash chromatography and HPLC (Section 3.4). NMR spectra were acquired in DMSO-*d*<sub>6</sub>+D<sub>2</sub>O, as the compounds were only sparingly soluble in DMSO-*d*<sub>6</sub> alone. The <sup>1</sup>H NMR spectrum of the quercetin 3-*O*-glycoside (**4**) included two resonances corresponding to the anomeric protons of *O*-linked sugar residues at  $\delta$ <sub>H</sub> 5.38 (1H, *d*, *J* = 7.8 Hz,  $\delta$ <sub>C</sub> 99.3) and 4.97 (1H, *d*, *J* = 1.6 Hz,  $\delta$ <sub>C</sub> 101.2), respectively. Full assignment of the remaining sugar resonances was achieved using two-dimensional NMR (Section 3.6). The multiplicities and coupling constants for the <sup>1</sup>H resonances of the sugars were as expected for a  $\beta$ -galactopyranosyl ( $\beta$ -Galp) and an  $\alpha$ -rhamnopyranosyl ( $\alpha$ -Rhap) residue (Duus et al., 2000). Acid hydrolysis of **4** followed by determination of absolute configuration for the constituent monosaccharides confirmed that these were D-Gal and L-Rha (Section 3.5). A long range correlation detected in an HMBC experiment between H-1 of the  $\alpha$ -L-Rhap residue and the downfield shifted C-2 resonance of the  $\beta$ -D-Galp residue ( $\delta$ <sub>C</sub> 76.1) indicated that these sugars were (1  $\rightarrow$  2)-linked, as predicted by MS. The <sup>1</sup>H NMR spectrum of **4** also contained resonances corresponding to an aliphatic moiety, which comprised two isolated methylene groups at  $\delta$ <sub>H</sub> 2.16 and 2.09 (both 1H, *d*, *J* = 13.3 Hz), and 2.06 and 1.89 (both 1H, *d*, *J* = 15.4 Hz), and a quaternary methyl group at  $\delta$ <sub>H</sub> 0.87 (3H, *s*). Use of indirect detection methods (HSQC, HMBC) indicated that a 3-hydroxy-3-methylglutaryl (HMG) residue was present, which was represented by a characteristic set of <sup>13</sup>C resonances at  $\delta$ <sub>C</sub> 27.4, 46.7, 47.0, 69.7, 171.5 and 177.6 (Jung et al., 1993; Sugiyama et al., 1993). The site of acylation was established using HMBC data, in which a long range correlation from 6-CH<sub>2</sub> of  $\beta$ -D-Galp to the carbonyl carbon of the HMG residue at  $\delta$ <sub>C</sub> 171.5 was detected. Compound **4** was therefore quercetin 3-*O*- $\alpha$ -L-rhamnopyranosyl-(1  $\rightarrow$  2)-[6-*O*-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranoside].

The <sup>1</sup>H NMR spectrum of **10** was similar to that of **4**, except that the aromatic resonances were those of kaempferol, rather than quercetin (Section 3.7). The two compounds contained the same acylated disaccharide moiety; thus acid hydrolysis of **10** followed by determination of absolute configuration for the constituent monosaccharides revealed the presence of D-Gal and L-Rha, the <sup>1</sup>H and <sup>13</sup>C NMR assignments of the glycosyl residues of **10** were similar to those of **4** (Section 3.7), and the same long range correlations were observed in HMBC data (i.e. H-1 of  $\alpha$ -L-Rhap to C-2 of  $\beta$ -D-Galp, and 6-CH<sub>2</sub> of  $\beta$ -D-Galp to the carbonyl carbon of the HMG group). Compound **10** was therefore kaempferol 3-*O*- $\alpha$ -L-rhamnopyranosyl-(1  $\rightarrow$  2)-[6-*O*-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranoside].

**Table 1**  
Flavonoids identified in a 80:20 MeOH/H<sub>2</sub>O leaf extract of *Rosa spinosissima* (BI-19505). HPLC retention times ( $t_R$ ), experimental high resolution  $m/z$  values of  $[M-H]^-$ , calculated molecular formulae (of M), relative abundance (Rel. Ab.) of  $[M-H]^-$  (as % of total abundance of all deprotonated flavonoids assigned) and means by which the compound was identified (Det.<sup>a</sup>), are listed.

No	$t_R$	$[M-H]^-$ ( $m/z$ )	Molecular formula	Rel. Ab. (%)	Identity	Det. <sup>a</sup>
1	16.62	609.1460	C <sub>27</sub> H <sub>30</sub> O <sub>16</sub>	0.9	Quercetin-3-O- $\alpha$ -rhamnosyl-(1 $\rightarrow$ 2)- $\beta$ -galactoside	UV, MS
2	17.11	609.1467	C <sub>27</sub> H <sub>30</sub> O <sub>16</sub>	<0.1	Quercetin-3-O- $\alpha$ -rhamnosyl-(1 $\rightarrow$ 2)- $\beta$ -glucoside	MS
3	19.46	593.1513	C <sub>27</sub> H <sub>30</sub> O <sub>15</sub>	1.1	Kaempferol-3-O- $\alpha$ -rhamnosyl-(1 $\rightarrow$ 2)- $\beta$ -galactoside	UV, MS
4	19.97	753.1876	C <sub>33</sub> H <sub>38</sub> O <sub>20</sub>	31.9	Quercetin 3-O- $\alpha$ -rhamnosyl-(1 $\rightarrow$ 2)-[(6-O-3-hydroxy-3-methylglutaryl)- $\beta$ -galactoside]	NMR
5	20.07	593.1520	C <sub>27</sub> H <sub>30</sub> O <sub>15</sub>	0.1	Kaempferol-3-O- $\alpha$ -rhamnosyl-(1 $\rightarrow$ 2)- $\beta$ -glucoside	MS
6	20.19	463.0884	C <sub>21</sub> H <sub>20</sub> O <sub>12</sub>	2.0	Quercetin 3-O- $\beta$ -galactoside (hyperoside)	Std
7	20.25	753.1882	C <sub>33</sub> H <sub>38</sub> O <sub>20</sub>	1.0	Quercetin 3-O- $\alpha$ -rhamnosyl-(1 $\rightarrow$ 2)-[(X-O-3-hydroxy-3-methylglutaryl)- $\beta$ -glucoside]	MS
8	20.74	463.0887	C <sub>21</sub> H <sub>20</sub> O <sub>12</sub>	0.7	Quercetin 3-O- $\beta$ -glucoside (isoquercitrin)	Std
9	21.39	477.0674	C <sub>21</sub> H <sub>18</sub> O <sub>13</sub>	2.5	Quercetin-3-O- $\beta$ -glucuronide (miquelianin)	NMR
10	22.58	737.1931	C <sub>33</sub> H <sub>38</sub> O <sub>19</sub>	32.3	Kaempferol 3-O- $\alpha$ -rhamnosyl-(1 $\rightarrow$ 2)-[(6-O-3-hydroxy-3-methylglutaryl)- $\beta$ -galactoside]	NMR
11	22.74	607.1305	C <sub>27</sub> H <sub>28</sub> O <sub>16</sub>	6.1	Quercetin 3-O-[(6-O-3-hydroxy-3-methylglutaryl)- $\beta$ -galactoside]	NMR
12	23.11	607.1306	C <sub>27</sub> H <sub>28</sub> O <sub>16</sub>	1.9	Quercetin 3-O-[(X-O-3-hydroxy-3-methylglutaryl)- $\beta$ -glucoside]	MS
13	23.28	447.0936	C <sub>21</sub> H <sub>20</sub> O <sub>11</sub>	0.6	Kaempferol-3-O- $\beta$ -galactoside (trifolin)	UV, MS
14	23.64	737.1938	C <sub>33</sub> H <sub>38</sub> O <sub>19</sub>	2.1	Kaempferol 3-O- $\alpha$ -rhamnosyl-(1 $\rightarrow$ 2)-[(X-O-3-hydroxy-3-methylglutaryl)- $\beta$ -glucoside]	MS
15	24.07	463.0882	C <sub>21</sub> H <sub>20</sub> O <sub>12</sub>	5.3	Quercetin 4'-O- $\beta$ -glucoside (spiraeoside)	UV, MS
16	24.49	447.0938	C <sub>21</sub> H <sub>20</sub> O <sub>11</sub>	0.6	Kaempferol 3-O- $\beta$ -glucoside (astragalin)	Std
17	24.69	447.0935	C <sub>21</sub> H <sub>20</sub> O <sub>11</sub>	5.3	Kaempferol 4'-O- $\beta$ -glucoside	UV, MS
18	25.04	461.0732	C <sub>21</sub> H <sub>18</sub> O <sub>12</sub>	1.6	Kaempferol-3-O- $\beta$ -glucuronide	Std
19	25.62	591.1364	C <sub>27</sub> H <sub>28</sub> O <sub>15</sub>	2.8	Kaempferol 3-O-[(X-O-3-hydroxy-3-methylglutaryl)- $\beta$ -galactoside]	UV, MS
20	27.10	591.1358	C <sub>27</sub> H <sub>28</sub> O <sub>15</sub>	1.1	Kaempferol 3-O-[(X-O-3-hydroxy-3-methylglutaryl)- $\beta$ -glucoside]	MS

<sup>a</sup> Identification was by NMR spectroscopy of isolated compounds (NMR), interpretation of UV and MS<sup>n</sup> spectra (UV, MS) according to Kite and Veitch (2009), or by comparison of analytical data with an authentic standard (Std). For **15** and **17** comparison was made with a petal extract of *R. spinosissima* (BI-19506) in which these compounds have been reported (Mikanagi et al., 1995). For **15** the 4'-O-linkage was supported by additional comparison with quercetin 7-O-glucoside which has a very similar UV spectrum but which was found to elute before quercetin 3-O-glucoside. 'X' indicates that the site of acylation was not determined.

In LC–UV–MS analyses, six other *R. spinosissima* flavonoids (**7**, **11**, **12**, **14**, **19**, **20**) gave deprotonated molecules in negative ion MS that showed a neutral loss of 144 following MS<sup>2</sup>, in common with **4** and **10**. Accurate mass determinations for the former confirmed the losses as C<sub>6</sub>H<sub>8</sub>O<sub>4</sub>, suggesting that these compounds were also acylated with 3-hydroxy-3-methylglutaric acid. As minor constituents of the *R. spinosissima* extract, only **11** was obtained in sufficient quantity for <sup>1</sup>H NMR analysis, which confirmed its identity as quercetin 3-O-[6-O-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranoside], a compound obtained previously by Wald et al. (1986) from MeOH extracts of the fruits of *Rubus fruticosus* (Rosaceae). UV and MS data indicated that the HMG derivative **19** was the kaempferol analogue of **11**; in particular, the negative ion MS<sup>3</sup> ( $m/z$  591  $\rightarrow$  447) of **19** was the same as the MS<sup>2</sup> of deprotonated kaempferol 3-O-galactoside (**13**). However the site of acylation could not be established unequivocally for **19**, nor for **7**, **12**, **14** and **20**. Use of serial MS indicated that the latter were the glucosyl analogues of **4**, **11**, **10** and **19**, respectively.

### 2.3. Distribution of flavonol O-glycosides acylated with 3-hydroxy-3-methylglutaric acid in *Rosa* and other Rosaceae

Flavonoid glycosides acylated with 3-hydroxy-3-methylglutaric acid (HMG-glycosides) are of limited occurrence in plants, with just over 50 examples reported to date, including the present results (Table S1, and references therein). They have been found not only in flowering plants, but also in ferns and liverworts. The types of flavonoid associated with acylation by 3-hydroxy-3-methylglutaric acid are the O-glycosides of flavones, flavonols and flavanones, C-glycosides of flavones, and one example of an anthocyanin. Flavone HMG-glycosides (O-linked) have mainly been reported from liverworts, although ferns and the angiosperm families Asteraceae and Rutaceae are also represented. HMG-glycosides of C-glycosylflavones have been found in the Leguminosae, Liliaceae and Poaceae. Existing sources of HMG-glycosides of flavanones are restricted to ferns and the genus *Citrus* (Rutaceae). Conversely, flavonol O-glycosides acylated with 3-hydroxy-3-methylglutaric acid have only been described from angiosperms, and account for almost half of the known flavonoid HMG-glyco-

sides. The glycosylation patterns of HMG-glycosides differ among the flavonoid classes represented (Table S1); thus for flavones, flavanones and the sole anthocyanin, the acylated sugar is always a primary  $\beta$ -glucopyranoside. In the case of flavonols, the HMG-glycosides are mainly  $\beta$ -glucopyranosides (64%), but  $\beta$ -galactopyranosides (28%) and  $\alpha$ -rhamnopyranosides (8%) are also found. Where the glycosyl moiety is a disaccharide, the acylation site is always on the primary sugar residue. Although the structures of the new HMG-glycosides (**4**, **10**) isolated from *R. spinosissima* conform to these trends, they are also the first examples of O- $\alpha$ -L-rhamnopyranosyl-(1  $\rightarrow$  2)- $\beta$ -D-galactopyranosides acylated with 3-hydroxy-3-methylglutaric acid to be reported. The relative scarcity of flavonoid HMG-glycosides and the unique attributes of **4** and **10** suggested that the distribution of these compounds, and the related examples **11** and **19**, would be of systematic interest in *Rosa* and other Rosaceae.

Among species analysed from 13 genera of Rosaceae (other than *Rosa*), leaves of species from 12 genera did not contain either **4** or **10**, nor the other two flavonol HMG-galactosides (**11** and **19**), characterised from *R. spinosissima* (Table 2). The exception was *Rubus idaeus* L., in which **4**, **10**, **11** and **19** were detected as minor flavonoid components. In previous work, **11** was isolated from fruits of *R. fruticosus* (Wald et al., 1986); however, we did not detect this compound in leaves of a wild-collected sample of this species, nor in leaves of any of the 12 other *Rubus* species analysed. Based on DNA sequence data, *Rubus* is considered to be the most closely related genus to *Rosa* (Potter et al., 2007). Reports of other HMG-glycosides in Rosaceae are limited to isorhamnetin 3-O-[6-O-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranoside], which occurs in the inflorescences of *Sorbus torminalis* (Olszewska and Roj, 2011).

In *Rosa* itself, 147 accessions were examined from 71 species among the various sub-genera and sections (Table 2). Of these, **4** and **10** were only found among 9 species, of which 8 were members of subgenus *Rosa* section *Pimpinellifoliae* and the ninth, *R. roxburghii*, was the sole species in the monotypic subgenus *Platyrrhodon* (Table 2). A consistent feature of the 10 specimens of *R. spinosissima* analysed was the presence of **4** and **10** as the major flavonoid components, and **4** and **10** were also major flavonoids in

**Table 2**

Distribution of foliar flavonol HMG-galactosides (**4**, **10**, **11** and **19**) amongst species of *Rosa* and Rosaceae (figures for each compound give the peak height (maximum ion abundance) of [M–H]<sup>–</sup> expressed as a percentage of the peak height of the flavonoid giving the biggest peak in the LC–MS analysis).

Species	Collection reference <sup>a</sup>	Date collected	<b>4</b>	<b>10</b>	<b>11</b>	<b>19</b>	Sample no. <sup>b</sup>
<b>Subgenus <i>Rosa</i> section <i>Pimpinellifoliae</i> (DC.) Ser. (14/15)<sup>c</sup></b>							
<i>Rosa altaica</i> Willd.	RF Murchison 1 (K)	14/09/1916	42	3	63	50	BI-20531
<i>Rosa ecae</i> Aitch.	RB & L Gibbons 435 (K)	29/05/1971	–	–	–	–	BI-20525
<i>Rosa ecae</i> Aitch.	KH Reching 35420 (K)	02/04/1967	–	–	–	–	BI-20526
<i>Rosa ecae</i> Aitch.	P Furse 8004 (K)	22/06/1966	–	–	–	–	BI-20527
<i>Rosa ecae</i> Aitch.	S Dixon 16/99 (K)	31/07/1999	–	–	–	–	BI-20528
<i>Rosa farreri</i> Stapf. ex Cox	EA Bowles 1, UK (Cult.) (K)	14/08/1928	–	–	–	–	BI-20523
<i>Rosa foetida</i> Herrm.	Lemann 3/80 (K)	1852	–	–	–	–	BI-20519
<i>Rosa foetida</i> Herrm.	S Ormant & F Karirum 38093 (K)	16/05/1971	–	–	–	–	BI-20520
<i>Rosa foetida</i> Herrm.	P Furse 6404 (K)	29/05/1994	–	–	–	–	BI-20521
<i>Rosa foetida</i> Herrm.	B Gilliat-Smith 1570 (K)	05/05/1926	–	–	–	–	BI-20522
<i>Rosa graciliflora</i> Rehder & E.H. Wilson	2001-1323	26/10/2010	–	–	–	–	BI-20229
<i>Rosa graciliflora</i> Rehder & E.H. Wilson	CR Lancaster L940 (K)	14/09/1983	–	–	–	–	BI-20324
<i>Rosa gracilipes</i> Chrshan.	1995-855	06/05/2010	43	100	20	31	BI-19507
<i>Rosa gracilipes</i> Chrshan	M Flanagan & A Kirkham ESUS 26 (K)	12/09/1994	86	100	38	31	BI-20536
<i>Rosa hemisphaerica</i> Herrm.	AC Frost 1031	24/04/1941	–	–	–	–	BI-20532
<i>Rosa hemisphaerica</i> Herrm.	M Grace, E Bailey, D Bull, S Carr 95 (K)	18/07/1981	–	–	–	–	BI-20533
<i>Rosa hugonis</i> Hemsl.	1973-16453	06/05/2010	6	32	86	100	BI-19503
<i>Rosa hugonis</i> Hemsl.	EH Wilson 3635 (K)	25/05/1908	–	–	–	–	BI-20534
<i>Rosa hugonis</i> Hemsl.	Lackham Sch Agr UK 2111/67 (Cult.) (K)	08/06/1967	4	1	100	69	BI-20535
<i>Rosa omeiensis</i> Rolfe	Fliegner, Erskine, Howick, McNamara TIBT 10 (K)	22/09/1995	–	–	35	2	BI-20545
<i>Rosa omeiensis</i> Rolfe	EH Wilson 3596 (K)	Sep. 1908	–	–	–	–	BI-20546
<i>Rosa omeiensis</i> Rolfe	Fliegner, Erskine, Howick, McNamara TIBT 35 (K)	23/09/1995	–	–	–	–	BI-20547
<i>Rosa omeiensis</i> Rolfe	Zhao Qing-Sheng 290 (K)	18/07/1989	–	–	–	–	BI-20548
<i>Rosa omeiensis</i> Rolfe	SINO British Expd Cangshan 537 (K)	13/05/1981	15	100	94	88	BI-20549
<i>Rosa primula</i> Boulenger	Arnold Arb Harvard 7446 (Cult.) (K)	09/05/1974	–	–	–	–	BI-20530
<i>Rosa sericea</i> Lindl.	1981-5230	29/09/2010	94	100	25	7	BI-20048
<i>Rosa sericea</i> Lindl.	2001-4387	26/10/2010	–	–	75	8	BI-20223
<i>Rosa sericea</i> Lindl.	1973-6394	26/10/2010	3	1	5	0	BI-20225
<i>Rosa sericea</i> var. <i>omeiensis</i> (Rolfe) G.D. Rowley	1981-3920	29/09/2010	79	100	46	12	BI-20050
<i>Rosa sericea</i> var. <i>omeiensis</i> (Rolfe) G.D. Rowley	1981-3971	29/09/2010	–	–	–	–	BI-20251
<i>Rosa sikangensis</i> T.T. Yu & T.C. Ku	1994-854	06/05/2010	11	4	18	4	BI-20055
<i>Rosa sikangensis</i> T.T. Yu & T.C. Ku	Fliegner, Erskine, Howick, McNamara SICH 1427 (K)	29/09/1994	–	–	–	–	BI-20540
<i>Rosa sikangensis</i> T.T. Yu & T.C. Ku	Fliegner, Howick, Staniforth, McNamara SICH 1035 (K)	01/10/1992	1	1	1	2	BI-20541
<i>Rosa spinosissima</i> L.	2009-409	05/10/2010	100	16	49	6	BI-20130
<i>Rosa spinosissima</i> L.	Kew, John Innes M3231 (Cult.) (K)	07/05/1963	100	60	64	11	BI-20550
<i>Rosa spinosissima</i> L.	Commercial sample	15/10/2008	85	100	34	11	BI-18236
<i>Rosa spinosissima</i> L.	1999-2167	06/05/2010	86	100	22	18	BI-19505
<i>Rosa spinosissima</i> L.	1927-2803	29/09/2010	100	16	48	6	BI-20030
<i>Rosa spinosissima</i> L.	E Ferrari, G Gola, O Mattiolo 1860 (K)	09/12/2010	100	100	59	38	BI-20319
<i>Rosa spinosissima</i> L.	M Pitman, A Wickham Turx 249 (K)	13/09/1993	100	78	51	9	BI-20315
<i>Rosa spinosissima</i> L.	VG Sobko 30	15/05/1993	100	86	70	2	BI-20477
<i>Rosa spinosissima</i> L.	Halliwel, Mason, Smallcombe 1207 (K)	03/07/1975	61	100	39	26	BI-20478
<i>Rosa spinosissima</i> L.	RB Hooglana 1979 (K)	25/05/1949	76	59	100	58	BI-20479
<i>Rosa spinosissima</i> var. <i>micrantha</i> (DC.) Ser.	NY Sandwith 5011 (K)	08/07/1957	100	21	35	4	BI-20562
<i>Rosa xanthina</i> Lindl.	Herb J Hers 336 (K)	24/04/1921	–	–	–	–	BI-20543
<i>Rosa xanthina</i> var. <i>ecae</i> (Aitch.) Boul.	R Ecreh, Aitch & Hemsl s.n. (K)	Dec. 1909	–	–	–	–	BI-20544
<i>Rosa xanthina</i> var. <i>kokanica</i> Lindl.	1978-806	06/05/2010	88	100	48	25	BI-19511
<b>Subgenus <i>Rosa</i> section <i>Gallicanae</i><sup>d</sup> (DC.) Ser. (1/1)</b>							
<i>Rosa gallica</i> var. <i>officinalis</i> L.	2009-396	29/09/2010	–	–	–	–	BI-20075
<b>Subgenus <i>Rosa</i> section <i>Caninae</i><sup>d</sup> (DC.) Ser. (13/~60)</b>							
20 accessions, 13 species <sup>d</sup>	–	–	–	–	–	–	–
<b>Subgenus <i>Rosa</i> section <i>Carolinae</i><sup>d</sup> Crép. (2/~5)</b>							
5 accessions, 2 species	–	–	–	–	–	–	–
<b>Subgenus <i>Rosa</i> section <i>Cinnamomeae</i><sup>d</sup> (DC.) Ser. (21/~80)</b>							
30 accessions, 21 species	–	–	–	–	–	–	–
<b>Subgenus <i>Rosa</i> section <i>Synstylae</i><sup>d</sup> DC. (12/~35)</b>							
34 accessions, 12 species	–	–	–	–	–	–	–
<b>Subgenus <i>Rosa</i> section <i>Banksianae</i> Lindl. (2/2)</b>							
3 accessions, 2 species	–	–	–	–	–	–	–
<b>Subgenus <i>Rosa</i> section <i>Indicae</i> Thory (1/2)</b>							
<i>Rosa chinensis</i> Jacq.	1929-81306	29/09/2010	–	–	–	–	BI-20069
<b>Subgenus <i>Rosa</i> section <i>Laevigatae</i> Thory (1/1)</b>							
<i>Rosa laevigata</i> Michx.	1986-527	29/09/2010	–	–	–	–	BI-20094
<b>Subgenus <i>Rosa</i> section <i>Bracteatae</i> Thory (1/2)</b>							
<i>Rosa bracteata</i> J.C.Wendl.	1966-26307	29/09/2010	–	–	–	–	BI-20067
<b>Subgenus <i>Rosa</i> section <i>Platyrhodon</i> (Hurst) Rehder (1/1)</b>							
<i>Rosa roxburghii</i> Tratt.	1988-4596	29/09/2010	80	61	93	100	BI-20090
<i>Rosa roxburghii</i> Tratt. var. <i>hirtula</i> (Regel) Rehder & E. H. Wilson (as <i>R. hirtula</i> (Regel) Nakai)	2000-2153	29/09/2010	20	62	51	53	BI-20077

(continued on next page)



Table 2 (continued)

Species	Collection reference <sup>a</sup>	Date collected	4	10	11	19	Sample no. <sup>b</sup>
<b>Subgenus Hesperodos</b> Cockerell (1/2)							
<i>Rosa stellata</i> subsp. <i>stellata</i>	Wallace 126 (K)	10/07/1897	–	–	–	–	BI-20316
<b>Subgenus Hulthemia</b> (Dumort.) Focke (1/1)							
<i>Rosa persica</i> J.F. Gmel.	Townsend 69/174 (K)	25/05/1969	–	–	–	–	BI-20325
<b>Rosaceae</b>							
<i>Rubus fruticosus</i> L.	Wild	11/10/2010	–	–	–	–	BI-20136
<i>Rubus idaeus</i> L.	Commercial sample	14/07/2009	36	19	86	33	BI-18845
<i>Rubus idaeus</i> L.	Wild	15/06/2009	1	6	21	72	BI-18688
<i>Rubus idaeus</i> subsp. <i>strigosus</i> (Michx.) Focke	1963-14808	21/04/2011	6	1	19	2	BI-20601
13 other species of <i>Rubus</i>	–	–	–	–	–	–	–
11 other genera of Rosaceae 21 accessions <sup>d</sup>	–	–	–	–	–	–	–

<sup>a</sup> Royal Botanic Gardens Kew living collections accession number or herbarium sheet number.

<sup>b</sup> RBG Kew phytochemical sample number (reference for LC–MS analysis in Kew's collections).

<sup>c</sup> Number of species examined and number estimated in group (from Wissemann and Ritz, 2007).

<sup>d</sup> Sample details – species, collection reference, date collected (sample number) – listed below: **Subgenus Rosa section Banksianae**: *Rosa banksiae* Aiton, SINO British Expd Cangshan K146 (K), 30/04/81 (BI-20323); –, *Camillio-Schneider* 1219 (K) 15/05/1914 (BI20539); *R. cymosa* Tratt., *Cavalerie* 1422 (K), 1922 (BI-20317); **Subgenus Rosa section Carolinae**: *R. palustris* Marshall, 2001-1325, 29/09/10 (BI-20081); –, 2003-901, 29/09/10 (BI-20082); –, 1981-3119, 29/09/10 (BI-20083); *R. virginiana* Mill., 2009-410, 05/10/10 (BI-20134); –, *ML Fernald & KM Wiegand* 3624 (K) 05/09/1910 (BI-20322); **Subgenus Rosa section Caninae**: *R. agrestis* Savi, 1975-5456, 29/09/10 (BI-20064); *R. canina* L., 1995-1662, 29/09/10 (BI-20068); *R. corymbifera* Borkh., 1995-1797, 29/09/10 (BI-20070); –, 1989-8263, 29/09/10 (BI-20092); *R. dumalis* subsp. *antalyensis* (Manden) O.Nilsson, 1994-841, 29/09/10 (BI-20072); *R. elliptica* Tausch ex Tratt., 1968-40401, 29/09/10 (BI-20073); *R. glauca* Pourr., 1986-3562, 26/10/10 (BI-20226); *R. iberica* Steven ex M.Bieb., 1998-1333, 29/09/10 (BI-20078); *R. jundzillii* Besser, 1968-4401, 29/09/10 (BI-20080); *R. rubiginosa* L., 1988-8664, 05/10/10 (BI-20131); *R. sherardii* Davies, 1994-856, 29/09/10 (BI-20084); *R. tomentosa* Sm., 1993-267, 29/09/10 (BI-20087); –, Baldacici 263, 14/01/99 (BI-20321); –, G. Degstat (K), 16/09/1897 (BI-20484); –, A. De Degen (K), 19/05/23 (BI-20485); –, A. Primavesi & A.P. Conolly (K), 09/10/77 (BI-20487); *R. t. var. umbelliflora* (Sw. ex Scheutz) Mattsson, B. Florstrom 750 (K), 17/09/1908 (BI-20486); *R. turcica* Rouy, 1978-4926, 29/09/10 (BI-20085); *R. villosa* L., R.M.A. Nesbitt 1369 (K), 26/09/87 (BI-20314); –, A. Prior 402 (K), Apr 1903 (BI-20480); **Subgenus Rosa section Cinnamomeae**: *R. acicularis* Lindl., 1995-838, 29/09/10 (BI-20010); –, 1999-3332, 29/09/10 (BI-20011); *R. a.* subsp. *bourgeauina* Crép., 1982-954, 29/09/10 (BI-20013); *R. a.* subsp. *sayi* (Schwein.) W.H. Lewis, 1973-16447, 29/09/10 (BI-20012); *R. amblyotis* C.A.Mey, 1998-1330, 29/09/10 (BI-20014); –, C.A.Mey, 1995-1069, 29/09/10 (BI-20015); *R. arizonica* (Gray) Rybd., 1980-1448, 29/09/10 (BI-20016); *R. arkansana* Porter, 1992-1043, 29/09/10 (BI-20065); *R. blanda* Aiton, 1956-56717, 29/09/10 (BI-20066); *R. caudata* Baker, 1913-52839, 29/09/10 (BI-20017); *R. davidii* Crép., 2009-2180, 29/09/10 (BI-20018); *R. davurica* Pall., 1998-1332, 29/09/10 (BI-20071); *R. forrestiana* Boulenger, Osbourne 84 Cult. (K), Oct-1936 (BI-20524); *R. gymnocarpa* Torr. & A. Gray, 1989-8409, 29/09/10 (BI-20023); –, J.P. Tracy 17814 (K), 05/0747 (BI-20481); –, J.B. Davy 315 (K), Sep 1893 (BI-20482); –, O.D. Allen 72 (K), 07/07/1894 (BI-20483); *R. jacutica* Juz., 1995-1347, 29/09/10 (BI-20079); *R. laxa* Retz., 1906-52603, 29/09/10 (BI-20034); *R. majalis* Herrm., 1968-15501, 29/09/10 (BI-20038); *R. moyesii* Hemsl. & E.H. Wilson, 1981-3939, 29/09/10 (BI-20040); –, 1992-3449, 29/09/10 (BI-20041); *R. rugosa* Thunb., 1996-1441, 29/09/10 (BI-20089); *R. sertata* Rolfe, 1996-1443, 29/09/10 (BI-20052); *R. setigera* Michx., 1936-94101, 29/09/10 (BI-20053); *R. setipoda* Hemsl. & E.H. Wilson, 1988-8810, 29/09/10 (BI-20054); *R. sweginowii* Koehne, 1988-8688, 29/09/10 (BI-20088); *R. webbiana* Royle, 1998-1336, 29/09/10 (BI-20060); –, 1991-1970, 29/09/10 (BI-20060); *R. willmottiae* Hemsl., 2004-595, 29/09/10 (BI-20062); *R. woodsii* Lindl., 2009-2182, 29/09/10 (BI-20063); **Subgenus Rosa section Synstylae**: *R. brunonii* Lindl., 1989-8263, 29/09/10 (BI-20093); *R. filipes* Rehder & E.H. Wilson, 1947-44801, 29/09/10 (BI-20074); *R. glomerata* Rehder & E.H. Wilson, 2000-2152, 29/09/10 (BI-20019); *R. helena* Rehder & E.H. Wilson, 1956-56706, 29/09/10 (BI-20024); –, 1981-8694, 29/09/10 (BI-20025); *R. henryi* Boulenger, 2003-905, 29/09/10 (BI-20033); –, 1994-845, 29/09/10 (BI-20035); *R. longicuspis* Bertol., 1999-2166, 29/09/10 (BI-20036); *R. luciae* var. *fujisanensis* Makino, 1994-846, 29/09/10 (BI-20037); *R. moschata* Herrm., 1994-858, 29/09/10 (BI-20039); *R. mulliganii* Boulenger, 1956-56763, 29/09/10 (BI-20042); *R. multiflora* Thunb., 1982-8236, 29/09/10 (BI-20043); –, 2000-2155, 29/09/10 (BI-20044); *R. m.* var. *cathayensis* Rehder & E.H. Wilson, BI-20026, 2009-2234, 29/09/10; *R. m.* var. *formosana* Cardot., 1994-848, 29/09/10 (BI-20028); *R. murielae* Rehder & E.H. Wilson, EH Wilson 3635 (K), 04/06/1902 (BI-20318); –, Wang Zhong-Tao 870266 (K), 29/07/1987 (BI-20551); *R. onoei* Makino, 2001-1324, 05/10/10 (BI-20132); *R. paniculigera* var. *awaensis* Makino ex Momiyama, 2001-1326, 29/09/10 (BI-20091); *R. paniculigera* var. *awaensis* Makino ex Momiyama, 2003-900, 29/09/10 (BI-20096); *R. rubus* L.H. Bailey, 1993-266, 29/09/10 (BI-20032); *R. sambucina* Koidz., 1993-907, 29/09/10 (BI-20045); *R. sempervirens* L., 2007-1415, 29/09/10 (BI-20046); *R. soulieana* Crép., 1973-20724, 29/09/10 (BI-20056); **Rosaceae**: *Alchemilla aronica* (Buser) Rothm., 1999 1133, 29/10/97 (BI-16924); *A. erythropoda* Juz., 1976-475, 29/10/97 (BI-16926); *A. mollis* (Buser) Rothm., 1973-10120, 29/10/97 (BI-16930); *A. subsericea* Reut., 1968-38106, 29/10/97 (BI-16925); *A. xanthochlora* Rothm., 1963-49201, 29/10/97 (BI-16923); *Agrimonia eupatoria* L., 2003-1182, 28/07/09 (BI18909); *Cotoneaster nanshan* A.Vilm. ex Mollet, 1988-8795, 14/10/10 (BI-20178); *Crataegus laevigata* DC., 1973-16498, 13/07/09 (BI-18836); *C. monogyna* Jacq., Wild (WP), 15/06/09 (BI18647); *Filipendula ulmaria* (L.) Maxim., 2003-1181, 14/10/10 (BI20177); *F. vulgaris* Moench., 1970-579, 14/10/10 (BI-20176); *Fragaria chiloensis* (L.) Mill., 2000-2501, 14/10/10 (BI-20174); *F. vesca* L., 1986-2325, 14/10/10 (BI20175); *Geum aleppicum* Jacq., 1988-5734, 14/10/10 (BI-20173); *G. urbanum* L., Wild (WP), 15/06/09 (BI-18656); *Potentilla erecta* Maiden, Wild (WP), 15/06/09 (BI-18681); *P. nepalensis* Hook., 1989-1446, 14/10/10 (BI-20172); *Prunus spinosa* L., 1987-2363, 28/07/09 (BI-18895); *Pseudocyclonia sinensis* C.K.Schneid., 1994-1771, 26/02/09 (BI-18507); *Rubus adenophorus* Rolfe, 1908-908, 21/04/2011 (BI20598); *R. caucasicus* Focke, 1890-15902, 21/04/2011 (BI-20587); *R. crataegifolius* Bunge, 1955-30801, 21/04/2011 (BI20595); *R. diversifolius* Lindl., 1990-16008, 21/04/2011 (BI-20600); *R. frondosus* Bigelow, 1909-67808, 21/04/2011 (BI-20594); *R. mesogaeus* Focke ex Diels., 1907-59901, 21/04/2011 (BI-20590); *R. neomexicanus* A. Gray, 1980-6391, 21/04/2011 (BI-20589); *R. niveus* Wall., 1983-772, 21/04/2011 (BI-20597); *R. odoratus* L., 1980-543, 21/04/2011 (BI-20599); *R. parviflorus* Nutt., 1965-28314, 21/04/2011 (BI-20588); *R. spectabilis* Pursh., 1961-48403, 21/04/2011 (BI-20591); *R. spectabilis* Pursh., 1961-48401, 21/04/2011 (BI-20592); *R. villicaulis* var. *calvatus* (Lees) Focke, 1973-11882, 21/04/2011 (BI-20596); *Sanguisorba minor* Scop., 2003-1174, 28/07/09 (BI-18913); *Sorbus aucuparia* L., Wild (WP), 15/06/09 (BI-18694).

the two specimens each of *R. gracilipes* and *R. roxburghii* examined, but among specimens of the other seven species in section *Pimpinellifoliae* in which **4** and **10** were detected their occurrence was more erratic. In some specimens, the low levels recorded, or failure to detect the compounds, may have been due to the age of the material (e.g. the E.H. Wilson 3635 collection of *R. hugonis* and the two herbarium specimens of *R. xanthina* were all more than 90 years old), but in more recent collections of *R. omeiensis* and living material of *R. sericea* the detection and levels of **4** and **10** also varied among specimens. This either indicates some intra-specific variability in the accumulation of **4** and **10** or problematic species identification; *R. sericea* and *R. omeiensis* have been united by some authors under the name *R. sericea* (Rowley, 1959), and likewise, *R. hugonis* and *R. xanthina* have been united under *R. xanthina* (Roberts, 1977). Flowers of *R. spinosissima* also contained **4**, **10**, **11**

and **19** as the major flavonoid components, but these four compounds were not detected in the hips. We did not study flowers or hips of other species.

Mikanagi et al. (1995) previously reported that members of section *Pimpinellifoliae* were distinct, chemically, within *Rosa* in containing the 4'-*O*-glucosides of kaempferol and quercetin in their flowers. The chemical distinction of some members of this section is supported by the accumulation of **4**, **10**, **11** and **19** among the foliar flavonoids. However, the failure to detect these compounds in other members of *Pimpinellifoliae* also supports DNA sequence analyses in which the monophyly of *Pimpinellifoliae* has been called into question, with some *Pimpinellifoliae* placed among members of sections *Cinnamomeae* and *Carolinae* (Koopman et al., 2008; Matsumoto et al., 2000, 2001; Wissemann and Ritz, 2005). It is of interest to note that nrITS-1 data suggests a close relationship of *R. ecae*,

*R. foetida* and *R. primula* in a clade distant from *R. spinosissima* (Wissemann and Ritz, 2005). These are three species in *Pimpinellifoliae* in which we could not detect flavonol HMG-galactosides.

Molecular studies place subgenus *Platyrhodon* within subgenus *Rosa*, so this monotypic group is not considered to warrant subgeneric status (e.g. Wissemann and Ritz, 2005). Its affinities within *Rosa* are, however, unclear. Analyses of chloroplast DNA and ribosomal spacer sequences have suggested that *R. roxburghii* is sister to, or near the base of, a major clade mainly containing species of subgenus *Rosa* other than those species mostly in sections *Pimpinellifoliae*, *Cinnamomeae* and *Carolinae* (Bruneau et al., 2007; Wissemann and Ritz, 2005). This position would not be congruent with the presence of flavonol HMG-galactosides in *R. roxburghii*, which suggests at least a chemical affinity with some members of *Pimpinellifoliae*. In contrast, the use of AFLP markers to reconstruct species relationships in *Rosa* placed *R. roxburghii* near the base of the genus sister to *R. hugonis* (Koopman et al., 2008), a member of section *Pimpinellifoliae* in which flavonol HMG-galactosides were detected. There has been a strong indication from molecular data that *R. persica* (sub-genus *Hulthemia*) is sister to the rest of the genus and should retain its subgeneric status (Wissemann and Ritz, 2005). However, other analyses are less conclusive and place both *R. persica* and *R. roxburghii* in early-derived positions. We did not detect flavonol HMG-galactosides in the one sample of *R. persica* analysed.

Given that the low level of DNA sequence divergence in *Rosa* creates difficulties in distinguishing *Rosa* species at the molecular level, we suggest that the occurrence of flavonol HMG-galactosides should be considered as a character in future combined 'morphological-molecular' phylogenetic analyses of the genus. Our current state of phytochemical knowledge suggests that the glycosylation profile of these acylated flavonol *O*-glycosides is unusual, and the compounds appear to have a discrete distribution in *Rosa*.

### 3. Experimental

#### 3.1. General instrumentation

LC–UV–MS/MS analyses were performed on a Thermo Scientific system consisting of an 'Accela' U-HPLC unit with a photodiode array detector and an 'LTQ Orbitrap XL' mass spectrometer fitted with an electrospray source. Chromatography was performed on 5  $\mu$ l sample injections onto a 150  $\times$  3 mm i.d., 3  $\mu$ m, Luna C18(2) column (Phenomenex) using a 400  $\mu$ l/min linear mobile phase gradient of MeOH/H<sub>2</sub>O/ACN + 1% formic acid changing from 0:90:10 to 40:50:10 over 30 min or from 0:90:10 to 90:0:10 over 20 min, followed by a column wash phase and equilibration of the column in start conditions for 3 min before the next injection. The ESI source of the mass spectrometer was operated in both positive and negative modes under the recommended manufacturers' conditions for the mobile phase parameters. The orbitrap mass analyser was set to scan in range *m/z* 200–2000 at 30,000 resolutions in one polarity while the linear ion-trap analyser performed MS<sup>n</sup> analyses on the most abundant ions in both polarities using an ion isolation window of  $\pm 2$  *m/z* and relative collision energy of 35%. For accurate mass analyses of product ions generated by MS<sup>2</sup> in the ion trap, the ions were scanned by the orbitrap at 7000 resolutions.

NMR spectra were acquired either in DMSO-*d*<sub>6</sub>+D<sub>2</sub>O at 37 °C on a Bruker Avance 400 MHz instrument, or in MeOH-*d*<sub>4</sub> at 30 °C on a Bruker Avance II +700 MHz instrument equipped with a 5 mm 1H/13C/15N triple-resonance PFG cryoprobe. Standard pulse sequences and parameters were used to acquire one-dimensional <sup>1</sup>H and two-dimensional COSY, TOCSY, HSQC and HMBC spectra. Chemical shift referencing was carried out using the internal solvent resonances at either  $\delta_{\text{H}}$  2.50 and  $\delta_{\text{C}}$  39.5 (DMSO-*d*<sub>6</sub>+D<sub>2</sub>O) or

$\delta_{\text{H}}$  3.31 and  $\delta_{\text{C}}$  49.1 (MeOH-*d*<sub>4</sub>), after calibration to TMS at 0.00 ppm.

#### 3.2. Plant material

Dried aerial material of *R. spinosissima* was obtained from a commercial supplier (RBG Kew sample numbers BI-18236 and BI-18454). Leaves from a specimen of *R. spinosissima* of verified identity growing at RBG Kew [Kew Accession No. 1999-2167] were collected on 06/05/10 (BI-19505). Flowers from this plant were collected on 06/05/10 (BI-19506) and hips were gathered on 26/10/10 (BI-20216). For the survey of species of *Rosa* and other Rosaceae, leaf samples were collected from plants growing in the living collections at RBG Kew or wild at RBG Kew (Wakehurst Place), or from specimens held in the Herbarium RBG Kew (K) (Table 2). All fresh material was freeze dried.

#### 3.3. Standards

Kaempferol 3-*O*-glucoside (astragalín), quercetin 3-*O*-galactoside (hyperoside), and quercetin 3-*O*-glucoside (isoquercitrin) were obtained from Apin Chemicals Ltd., and kaempferol 3-*O*-glucuronide was from the J.B. Harborne collection.

#### 3.4. Extraction, analysis, and isolation of flavonol glycosides

For analytical-scale work, freeze dried plant material was ground to a powder in a pestle and mortar and 100 mg extracted for 48 h with 1 ml of 80% aq. MeOH at room temperature (~20 °C). The extracts were then clarified by centrifugation and the supernatants analysed by LC–UV–MS/MS.

For the isolation of flavonol glycosides, 20 g of the commercially sourced sample of *R. spinosissima* was ground to a powder, and extracted for 48 h in 1 l of 80% aq. MeOH. The extract was filtered and dried *in vacuo* and subjected to flash C18 chromatography (Silica C18 40–70 $\mu$ m) on a 20 g, 750  $\times$  40 mm column, eluted with equal volumes (50 ml) of increasing percentages of aq. MeOH (10%, 30%, 50%, 80%, 100%). Fractions (10 ml) were collected and combined, according to the results of LC–MS analyses, into five fractions (A–E). Fraction B was subjected to semi-preparative HPLC (Waters 600 Controller, 717plus Autosampler and 2996 Photodiode Array Detector) with repeated injections (170  $\mu$ l) made onto a 250 mm  $\times$  10 mm i.d., 10  $\mu$ m particle size, Lichrospher RP-18e column (Phenomenex), operating with a 3 ml/min linear mobile phase gradient of 20:75:5 to 70:25:5 MeOH/H<sub>2</sub>O/MeOH + 5% acetic acid over 10 min. This yielded nine fractions (B1–B9), of which B2 and B3 were further purified by HPLC with repeated injections (50  $\mu$ l) made onto a 250 mm  $\times$  4.6 mm, 5  $\mu$ m particle size, Lichrospher RP18e column (Phenomenex) with a 1 ml/min mobile phase gradient of 25:75 to 75:25 MeOH/H<sub>2</sub>O over 10 min. Manual collection of the peaks of interest afforded **4** (1.0 mg), **9** (0.5 mg), **10** (1.2 mg) and **11** (0.5 mg), as pale yellow amorphous solids.

#### 3.5. Sugar analysis

Acid hydrolysis of **4** and **10** (0.5 mg of each in 100  $\mu$ l DMSO) was carried out using a standard protocol (0.5 ml 2 M HCl, 110 °C, 1.5 h). After cooling, particulates were spun down by microcentrifugation (5 min) and the supernatants removed and dried under a stream of nitrogen. The absolute configurations of the monosaccharides of **4** and **10** released by acid hydrolysis were determined from GC–MS analysis of their trimethylsilylated thiazolidine derivatives, which were prepared using the method of Ito et al. (2004). GC–MS analyses were performed on a Perkin Elmer Autosystem XL GC coupled to a Perkin Elmer Turbo Mass MS (EI, 70 eV) using a 30 m  $\times$  0.25 mm, 0.25  $\mu$ m, DB5-MS column (Agilent), an oven temp. program of 180–

300 °C at 6 °C/min, injection temp. 350 °C, and helium as the carrier gas at 1 ml/min. The acid hydrolysates of **4** and **10** each gave L-rhamnose and D-galactose at  $t_R$  = 10.2 and 12.2 min, respectively (identical to authentic standards).

### 3.6. Quercetin 3-O- $\alpha$ -L-rhamnopyranosyl-(1 $\rightarrow$ 2)-[6-O-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranoside] (**4**)

UV (LC-PDA)  $\lambda_{\max}$  nm: 256, 292 (sh), 356;  $^1\text{H}$  NMR (DMSO- $d_6$ +D $_2$ O):  $\delta$  7.51 (1H, dd,  $J$  = 8.4, 2.1 Hz, H-6'), 7.44 (1H, d,  $J$  = 2.1 Hz, H-2'), 6.78 (1H, d,  $J$  = 8.5 Hz, H-5'), 6.11 (1H, br s, H-8), 5.92 (1H, br s, H-6);  $\beta$ -Gal (primary): 5.38 (1H, d,  $J$  = 7.8 Hz, Gal H-1), 3.71 (1H, dd,  $J$  = 9.6, 7.7 Hz, Gal H-2), 3.55 (1H, dd,  $J$  = 9.7, 3.4 Hz, Gal H-3), 3.65 (1H, br d,  $J$  = 3.6 Hz, Gal H-4), 3.56 (1H, m, Gal H-5), 3.94 (1H, dd,  $J$  = 11.3, 7.8 Hz, Gal H-6a), 3.87 (1H, dd,  $J$  = 11.3, 5.0 Hz, Gal H-6b);  $\alpha$ -Rha (terminal): 4.97 (1H, d,  $J$  = 1.6 Hz, Rha H-1), 3.77 (1H, dd,  $J$  = 3.4, 1.7 Hz, Rha H-2), 3.53 (1H, dd,  $J$  = 9.6, 3.4 Hz, Rha H-3), 3.15 (1H, t,  $J$  = 9.6 Hz, Rha H-4), 3.81 (1H, dd,  $J$  = 9.6, 6.2 Hz, Rha H-5), 0.83 (3H, d,  $J$  = 6.2 Hz, Rha H-6); HMG: 2.16 (1H, d,  $J$  = 13.3 Hz, H-2a), 2.09 (1H, d,  $J$  = 13.3 Hz, H-2b), 0.87 (3H, s, 3-CH $_3$ ), 2.06 (1H, d,  $J$  = 15.4 Hz, H-4a), 1.89 (1H, d,  $J$  = 15.4 Hz, H-4b);  $^{13}\text{C}$  NMR (DMSO- $d_6$ +D $_2$ O):  $\delta$  122.4 (C-6'), 116.0 (C-2'), 115.8 (C-5'), C-6 and C-8 broadened beyond detection;  $\beta$ -Gal (primary): 99.3 (Gal C-1), 76.1 (Gal C-2), 73.9 (Gal C-3), 69.1 (Gal C-4), 73.0 (Gal C-5), 63.4 (Gal C-6);  $\alpha$ -Rha (terminal): 101.2 (Rha C-1), 70.9 (Rha C-2), 71.0 (Rha C-3), 72.3 (Rha C-4), 68.8 (Rha C-5), 17.5 (Rha C-6); HMG: 171.5 (C-1), 46.7 (C-2), 69.7 (C-3), 47.0 (C-4), 177.6 (C-5), 27.4 (3-CH $_3$ );

LC-HRESIMS (orbitrap)  $m/z$ : 753.1876 [M-H] $^-$  (calc. for C $_{33}$ H $_{37}$ O $_{20}$  $^-$ , 753.1884); LC-ESI-MS/MS (ion trap) of [M-H] $^-$ ,  $m/z$  (rel. int.): 691 (3), 651 (5), 609 [(M-H)-HMG] $^-$  (100); LC-ESI-MS $^3$  ( $m/z$  753  $\rightarrow$  609),  $m/z$  (rel. int.): 489 (13), 445 (loss of rhamnose) (12), 409 (2), 343 (4), 301 [quercetin-H] $^-$  (31), 300 [quercetin-2H] $^{\bullet-}$  (100), 271 (16), 255 (9).

### 3.7. Kaempferol 3-O- $\alpha$ -L-rhamnopyranosyl-(1 $\rightarrow$ 2)-[6-O-(3-hydroxy-3-methylglutaryl)- $\beta$ -D-galactopyranoside] (**10**)

UV (LC-PDA)  $\lambda_{\max}$  nm: 264, 292 (sh), 348;  $^1\text{H}$  NMR (DMSO- $d_6$ +D $_2$ O):  $\delta$  7.97 (2H, d,  $J$  = 8.9 Hz, H-2'/6'), 6.86 (2H, d,  $J$  = 8.9 Hz, H-3'/5'), 6.41 (1H, d,  $J$  = 2.1 Hz, H-8), 6.17 (1H, d,  $J$  = 2.1 Hz, H-6);  $\beta$ -Gal (primary): 5.39 (1H, d,  $J$  = 7.7 Hz, Gal H-1), 3.70 (1H, dd,  $J$  = 9.6, 7.7 Hz, Gal H-2), 3.56 (1H, dd,  $J$  = 9.7, 3.4 Hz, Gal H-3), 3.65 (1H, br d,  $J$  = 3.5 Hz, Gal H-4), 3.57 (1H, m, Gal H-5), 4.00 (1H, dd,  $J$  = 11.5, 8.0 Hz, Gal H-6a), 3.85 (1H, dd,  $J$  = 11.5, 4.1 Hz, Gal H-6b);  $\alpha$ -Rha (terminal): 4.98 (1H, d,  $J$  = 1.7 Hz, Rha H-1), 3.78 (1H, dd,  $J$  = 3.3, 1.7 Hz, Rha H-2), 3.52 (1H, dd,  $J$  = 9.5, 3.4 Hz, Rha H-3), 3.15 (1H, t,  $J$  = 9.6 Hz, Rha H-4), 3.78 (1H, dd,  $J$  = 9.6, 6.2 Hz, Rha H-5), 0.80 (3H, d,  $J$  = 6.2 Hz, Rha H-6); HMG: 2.11 (1H, d,  $J$  = 13.4 Hz, H-2a), 2.01 (1H, d,  $J$  = 13.4 Hz, H-2b), 0.82 (3H, s, 3-CH $_3$ ), 2.03 (1H, d,  $J$  = 15.4 Hz, H-4a), 1.86 (1H, d,  $J$  = 15.4 Hz, H-4b);  $^{13}\text{C}$  NMR (DMSO- $d_6$ +D $_2$ O):  $\delta$  131.5 (C-2'/6'), 115.8 (C-3'/5'), 99.6 (C-6), 94.7 (C-8);  $\beta$ -Gal (primary): 99.3 (Gal C-1), 76.1 (Gal C-2), 73.9 (Gal C-3), 69.1 (Gal C-4), 73.3 (Gal C-5), 63.6 (Gal C-6);  $\alpha$ -Rha (terminal): 101.3 (Rha C-1), 70.9 (Rha C-2), 71.0 (Rha C-3), 72.4 (Rha C-4), 68.9 (Rha C-5), 17.5 (Rha C-6); HMG: 171.5 (C-1), 46.5 (C-2), 69.6 (C-3), 46.9 (C-4), 177.3 (C-5), 27.4 (3-CH $_3$ ); LC-HRESIMS (orbitrap)  $m/z$ : 737.1931 [M-H] $^-$  (calc. for C $_{33}$ H $_{37}$ O $_{19}$  $^-$ , 737.1935);

LC-ESI-MS/MS (ion trap) of [M-H] $^-$ ,  $m/z$  (rel. int.): 675 (13), 635 (18), 593 [(M-H)-HMG] $^-$  (100); LC-ESI-MS $^3$  ( $m/z$  737  $\rightarrow$  593),  $m/z$  (rel. int.): 473 (7), 429 (loss of rhamnose) (43), 393 (10), 327 (9), 285 [kaempferol-H] $^-$  (97), 284 [kaempferol-2H] $^{\bullet-}$  (100), 255 (26).

## Acknowledgements

The authors thank Boots the Chemist (Nottingham) for the initial supply of leaf material and financial support, the Medical Research Council Biomedical NMR Centre, National Institute for Medical Research, Mill Hill, London, for access to high-field NMR facilities, Mr. Jonathon Reichl and Miss. Lauren Gooch for technical assistance, and staff at the Herbarium, Royal Botanic Gardens, Kew, for guidance in the sampling of herbarium material.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.phytochem.2012.05.006>.

## References

- Bruneau, A., Starr, J.R., Joly, S., 2007. Phylogenetic relationships in the genus *Rosa*: new evidence from chloroplast DNA sequences and an appraisal of current knowledge. *Syst. Bot.* 32, 366–378.
- Crépin, F., 1893. *Rosae hybridae. Études sur les roses hybrides.* Bull. Soc. Roy. Bot. Belgique 32, 52–55.
- Duus, J.Ø., Gotfredsen, C.H., Bock, K., 2000. Carbohydrate structural determination by NMR spectroscopy: Modern methods and limitations. *Chem. Rev.* 100, 4589–4614.
- Erlanson, E.W., 1929. Cytological conditions and evidences for hybridity in North America wild roses. *Bot. Gaz.* 87, 443–506.
- Ito, A., Chai, H.-B., Kardono, L.B.S., Setowati, F.M., Afriastini, J.J., Riswan, S., Farnsworth, N.R., Cordell, G.A., Pezzuto, J.M., Swanson, S.M., Kinghorn, A.D., 2004. Saponins from the bark of *Nephelium maingayi*. *J. Nat. Prod.* 67, 201–205.
- Jung, K.Y., Do, J.C., Son, K.H., 1993. Kaempferol 3-O-[6''-O-(3-hydroxy-3-methylglutaryl)glucoside] from leaves of *Polygala japonica*. *Phytochemistry* 34, 1196–1197.
- Kite, G.C., Veitch, N.C., 2009. Assigning glucose or galactose as the primary glycosidic sugar in 3-O-mono-, di- and triglycosides of kaempferol using negative ion electrospray and serial mass spectrometry. *Rapid Commun. Mass Spectrom.* 23, 3125–3132.
- Koopman, W.J.M., Wisseman, V., De Cock, K., Van Huylbroeck, A., De Riek, J., Sabatino, G.J.H., Visser, D., Vosman, B., Ritz, C.M., Maes, B., Werlemark, G., Nybom, H., Debener, T., Linde, M., Smulders, M.J.M., 2008. AFLP markers as a tool to reconstruct complex relationships: A case study in *Rosa* (Rosaceae). *Am. J. Bot.* 95, 353–366.
- Markham, K.R., 1982. *Techniques of Flavonoid Identification.* Academic Press, London.
- Matsumoto, S., Kouchi, M., Fukui, H., Ueda, Y., 2000. Phylogenetic analyses of the subgenus *Eurosa* using the ITS nrDNA sequence. *Acta Hort.* 521, 193–202.
- Matsumoto, S., Nishio, H., Ueda, Y., Fukui, H., 2001. Phylogenetic analysis of genus *Rosa*: polyphyly of section *Pimpinellifoliae* and origin of *Rosa x fortuniana* Lindl. *Acta Hort.* 547, 357–363.
- Melville, R., 1967. The problem of classification in the genus *Rosa*. *Bull. Jard. Bot. Natl. Belg.* 37, 39–44.
- Mikanagi, Y., Yokoi, M., Ueda, Y., Saito, N., 1995. Flower flavonol and anthocyanin distribution in subgenus *Rosa*. *Biochem. Syst. Ecol.* 23, 183–200.
- Olszewska, M.A., Roj, J.M., 2011. Phenolic constituents of the inflorescences of *Sorbus torminalis* (L.) Crantz. *Phytochem. Lett.* 4, 151–157.
- Potter, D., Eriksson, T., Evans, R.C., Oh, S., Smedmark, J.E.E., Morgan, D.R., Kerr, M., Robertson, K.R., Arsenault, M., Dickinson, T.A., Campbell, C.S., 2007. Phylogeny and classification of Rosaceae. *Plant Syst. Evol.* 266, 5–43.
- Rehder, A., 1940. *Manual of cultivated trees and shrubs hardy in North America*, 2nd edition. Macmillan, New York.
- Roberts, A.V., 1977. Relationship between species in the genus *Rosa*, section *Pimpinellifoliae*. *Bot. J. Linn. Soc.* 74, 309–328.
- Rowley, G.D., 1959. Some naming problems in *Rosa*. *Bull. Jard. Bot. Ét. Bruxelles* 29, 205–211.
- Sugiyama, S., Umehara, K., Kuroyanagi, M., Ueno, A., Taki, T., 1993. Studies on the differentiation inducers of myeloid leukemic cells from *Citrus* species. *Chem. Pharm. Bull.* 41, 714–719.
- The Plant List, 2010. Published on the Internet at <http://www.theplantlist.org/>.
- Wald, B., Galensa, R., Herrmann, K., Grotjahn, L., Wray, V., 1986. Quercetin 3-O-[6''-(3-hydroxy-3-methylglutaryl)- $\beta$ -galactoside] from blackberries. *Phytochemistry* 25, 2904–2905.
- Wisseman, V., 2003. Conventional taxonomy (wild roses). In: Roberts, A.V., Debener, T., Gudin, S. (Eds.), *Encyclopedia of Rose Science.* Elsevier, Amsterdam, pp. 111–117.
- Wisseman, V., Ritz, C.M., 2005. The genus *Rosa* (Rosaceae, Rosaceae) revisited: molecular analysis of nrITS-1 and *atpB-rbcL* intergenic space (IGS) versus conventional taxonomy. *Bot. J. Linn. Soc.* 147, 275–290.
- Wisseman, V., Ritz, C.M., 2007. Evolutionary patterns and processes in the genus *Rosa* (Rosaceae) and their implications for host-parasite co-evolution. *Plant Syst. Evol.* 266, 79–89.