Metabolic Engineering Reveals *LUP5* as a Determinant of Saponin Composition and Insect Resistance in *Barbarea vulgaris*

Jincheng Shen[1], Jan Günther[1], Sebastian Kjeldgaard-Nintemann[1], Pablo D. Cárdenas[1*], Søren Bak[1*].

[1]Department of Plant and Environmental Sciences, Faculty of Science, University of Copenhagen, 1871 Frederiksberg C, Denmark.

* Corresponding author: bak@plen.ku.dk

* Co-corresponding author: pdcardenas@plen.ku.dk

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plphys/pages/General-Instructions) is Søren Bak.

Running head: *LUP5* Drives Saponin-Mediated Insect Resistance

Summary

Plant–insect coevolution has been a major driver of specialized metabolite diversification, yet the genetic basis of natural variation in defensive chemistry remains poorly understood. The wild crucifer *Barbarea vulgaris* comprises two ecotypes, an insect-resistant G-type and a susceptible P-type, characterized by distinct triterpenoid saponin profiles. To investigate the causal relationship between saponin composition and insect resistance, we established a stable transformation system for *B. vulgaris*. Expression of the G-type β-amyrin synthase gene *LUP5* in the susceptible P-type conferred up to a 95% reduction in *Plutella xylostella* feeding, accompanied by increased accumulation of three hederagenin-derived monodesmosidic saponins. Comparison of *LUP5* expression driven by its native promoter and by the constitutive 35S promoter revealed that the native promoter leads to increased hederagenin accumulation and is activated later in development, which may prevent early metabolic stress and allow coordinated expression of downstream pathway genes. In contrast, silencing of *CYP72A552* by RNAi decreased total hederagenin levels by approximately 40% without affecting resistance, indicating threshold-dependent defense. Our results provide direct *in planta* evidence that *LUP5* is a key determinant of natural variation in insect resistance in *B. vulgaris*, underscoring the pivotal role of the saponin backbone in herbivore deterrence. By linking promoter activity to metabolite structural diversity, this work provides mechanistic and conceptual insight into how plants coordinate specialized metabolism and defense.

Keywords: triterpenoid saponins, metabolic engineering, β-amyrin synthase, promoter regulation, insect resistance, *Barbarea vulgaris,* chemical ecology

38 **Introduction**

39 Plants have evolved a wide array of specialized metabolites to deter insect
40 herbivores (Fürstenberg-Hägg, Zagrobelny and Bak, 2013; Yactayo-Chang et al., 2020;
41 Beran and Petschenka, 2022). Among these, triterpenoid saponins, which consist of a
42 hydrophobic sapogenin backbone and one or more hydrophilic sugar moieties,
43 represent a structurally diverse and ecologically important class of defensive
44 compounds. Due to their structural similarity to sterols and steroid hormones,
45 triterpenoid saponins have been hypothesized to interact with eukaryotic cell
46 membranes (Cárdenas, Almeida and Bak, 2019). In addition, their inherent bitterness
47 may deter herbivores, contributing to their role as effective defense agents (Augustin
48 et al., 2011). Thousands of triterpenoid saponins have been reported to date (Xu, Fazio
49 and Matsuda, 2004; Augustin et al., 2011; Moses, Papadopoulou and Osbourn, 2014;
50 Huang et al., 2020). Importantly, specific saponin structures have been implicated in
51 insect resistance, highlighting the functional significance of their chemical diversity
52 (Stevenson, Nyirenda and Veitch, 2010; Cui et al., 2019; Hussain et al., 2019; Liu et al.,
53 2019; Dolma et al., 2021).

54 The structure–activity relationships underlying saponin bioactivity are critical to
55 their defensive function. For example, hederagenin 3-*O*-monoglucoside has higher
56 toxicity and antifeedant activity than oleanolic acid 3-*O*-monoglucoside (which lacks
57 C-23 hydroxylation), and gypsogenic acid 3-*O*-monoglucoside (which is further
58 oxidized at C-23 to the ketone) (Liu et al., 2019). Additionally, monodesmosidic
59 hederagenin and oleanolic acid exhibited higher toxicity and antifeedant activity than
60 their bidesmosidic counterparts (Tian et al., 2021; Dervishi et al., 2025, in press). These
61 observations suggest that even subtle modifications to the sapogenin backbone,
62 glycosylation pattern, or sugar linkage can significantly alter saponin bioactivity. This
63 structural diversity makes saponins a valuable system for studying molecular
64 evolutionary mechanisms of specialized metabolism, particularly the structure–
65 activity relationships underlying their ecological functions in insect defense. Despite
66 growing knowledge of these structure–activity relationships, the genetic and
67 biochemical mechanisms underlying saponin diversification remain poorly
68 understood in most species. One particularly promising ecological model to explore

69 this question is the wild crucifer *Barbarea vulgaris*, which naturally segregates in two
70 ecotypes or chemotypes that alter in both saponin profiles and insect resistance.

71 *B. vulgaris* has emerged as a powerful ecological model for investigating how
72 specialized metabolites contribute to herbivore resistance (Kuzina et al., 2011; Hauser,
73 Toneatto and Nielsen, 2012; Byrne et al., 2017). The species comprises two ecotypes:
74 the glabrous insect-resistant G-type, and the pubescent insect-susceptible P-type
75 (Agerbirk, Olsen and Nielsen, 2001). This intraspecific divergence likely arose from
76 isolation in different Ice Age refugia (Hauser, Toneatto and Nielsen, 2012) and is
77 associated with marked differences in saponin composition (Kuzina et al., 2009;
78 Khakimov et al., 2012; Khakimov et al., 2016). Gene duplication events have further
79 contributed to the chemical divergence and insect resistance (Khakimov et al., 2015;
80 Liu et al., 2019). Insect resistance in G-type *B. vulgaris* correlates with four oleanane-
81 type defensive saponins, including oleanolic acid cellobioside, hederagenin
82 cellobioside, gypsogenin cellobioside, and 4-epihederagenin cellobioside (Agerbirk et
83 al., 2003; Kuzina et al., 2011). Among these, hederagenin cellobioside is particularly
84 toxic and confers G-type insect resistance (Agerbirk et al., 2003; Nielsen et al., 2010;
85 Kuzina et al., 2011).

86 The biosynthetic pathway leading to hederagenin cellobioside in G-type *B.*
87 *vulgaris* has been elucidated up to the first glycosylation step (Fig. 1). The
88 oxidosqualene cyclase *LUP5* is highly expressed in the G-type and primarily produces
89 β-amyrin (Khakimov et al., 2015). Subsequent enzymatic steps involve CYP716A80,
90 CYP72A552, and UGT73C11, which sequentially convert β-amyrin into oleanolic acid,
91 hederagenin, and hederagenin 3-*O*-glucoside (Fig. 1) (Augustin et al., 2012; Khakimov
92 et al., 2015; Liu et al., 2019). The enzyme responsible for the final step in forming
93 hederagenin cellobioside remains unidentified (Erthmann, Agerbirk and Bak, 2018). In
94 contrast, *LUP5* is expressed at very low levels in *B. vulgaris* P-type and primarily
95 produces α-amyrin due to two amino acid substitutions as compared to the *LUP5* in
96 the G-type (Günther et al., 2021). Despite 98% sequence identity at the amino acid
97 level, the G- and P-type LUP5 enzymes differ markedly in both their gene expression
98 and product profiles. The P-type plants preferentially express *LUP2,* an oxidosqualene
99 cyclase over 80% identical at the amino acid level to *LUP5*, but it produces

predominantly lupeol (98%) and only 2% β-amyrin-derived saponins (Khakimov et al., 2015). Notably, the genes encoding for the remaining enzymes responsible for the biosynthesis of hederagenin 3-*O*-glucoside (CYP716A80, CYP72A552, and UGT73C11) are present in both P-type and G-type and exhibit similar substrate and product specificity (Augustin et al., 2012; Khakimov et al., 2015; Liu et al., 2019). However, the *cyclases* are differentially regulated suggesting that upstream regulation and oxidosqualene cyclase product specificity may be key drivers of saponin structural and insect resistance difference. Additionally, CYP72A552, which catalyzes the C23 hydroxylation of oleanolic acid to hederagenin, the precursor of the more toxic hederagenin-type saponins, has not yet been validated *in planta* for its ecological contribution to insect resistance in *B. vulgaris*.

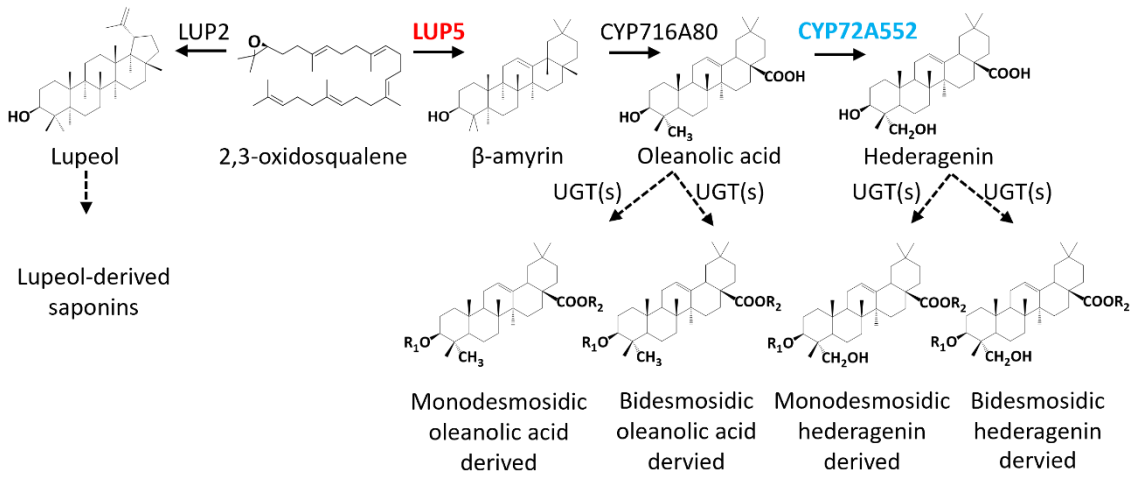


Fig. 1 Biosynthetic pathways for lupeol and β–amyrin derived triterpenoid saponins in *B. vulgaris*. Monodesmosidic saponin, a single sugar chain attached at preferentially C-3 (R1) or C-28 (R2); bidesmosidic saponin, sugar chains attached at both C-3 (R1) and C-28 (R2).

To investigate the genetic and biochemical basis of saponin-mediated insect resistance, we leveraged the natural variation in *B. vulgaris* to develop a model system linking *in planta* metabolic engineering of specialized metabolism with ecological function. We provide functional evidence that engineering saponin biosynthesis through manipulation of *LUP5* expression substantially enhances insect resistance in this wild crucifer and explains the molecular basis of the difference in insect resistance

122 between the G- and P-type *B. vulgaris*. By comparing constructs driven by the native
123 *LUP5* and constitutive 35S promoters, we further reveal how promoter activity shapes
124 metabolite structural diversity. These findings offer new mechanistic insight into
125 plant–insect coevolution and established a model system for dissecting genotype–
126 phenotype relationships in plant chemical specialization.

127 **Materials and methods**

128 **Plant and insect materials**

129 Seeds of *Barbarea vulgaris* G-type (glabrous, accession B44, Herbarium-code:
130 CP0057358) and P-type (pubescent, accession B4, herbarium-code: CP0057347) were
131 provided by Associate professor Niels Agerbirk *(Agerbirk et al., 2021)* and used for
132 stable transformation. Both transformed and wild-type plants were cultivated in soil
133 (Krukvaxtjord med lera och kisel, SW Horto AB, Hammenhog, Sweden) in a greenhouse
134 maintained at 19 °C under a 16 h light/8 h dark photoperiod. Irradiance was
135 supplemented with LED lamps whenever natural light intensity dropped below 250
136 µmol m^{-2} s^{-1}, with full-spectrum sunlight available when curtains were open. Plants
137 were fertigated on Thursdays and Fridays with a nutrient solution (electrical
138 conductivity 2.2 mS cm^{-1}; pH 5.8) prepared by dissolving 100 g L^{-1} of a compound
139 fertilizer (Pioner Basis 13-2-23+3+ME; containing 13% N, 2% P_2O_5, 23% K_2O, 3% MgO,
140 plus micronutrients) together with 0.5 g L^{-1} of iron chelate (Pioner Fe-EDDHA 6%; 6%
141 Fe) in water. *Brassica napus* plants were grown in the same greenhouse under
142 identical conditions. *Plutella xylostella* (diamondback moth) eggs were obtained from
143 Dr. Patrick Hughes (Boyce Thompson Institute, Ithaca, NY, USA). The laboratory colony
144 was originally established at the New York State Agricultural Experiment Station
145 (Geneva, NY, USA) in 1994. The insects were reared on *B. napus* at 20 °C under a 16 h
146 light/8 h dark photoperiod in cages. Third instar *P. xylostella* larvae were used for the
147 insect feeding assay.

148 **Plasmids construction**

149 For stable transformation, the pJCV51 vector was employed to introduce the G-
150 type *LUP5* coding sequence and to assess tissue-specific promoter activity (native

151 *LUP5* and constitutive 35S), while pK7FWG2 was utilized for *CYP72A552* silencing
152 (Karimi, Inzé and Depicker, 2002). The *LUP5* coding sequence, its native promoter
153 (1988 bp), and the *CYP72A552* coding sequence were PCR-amplified from *B. vulgaris*
154 G-type cDNA and genomic DNA (Supplementary Table S1 and S2). G-type *LUP5*
155 expression was driven either by the 35S promoter or its native promoter, while
156 *CYP72A552* was driven by the 35S promoter.

157 The pJCV51-p35S::LUP5, pJCV51-p35S::eGFP, and pK7FWG2-CYP72A552 RNAi
158 plasmids were assembled using Gateway technology (Gateway™ BP Clonase™ II
159 Enzyme mix, Thermo Fisher Scientific, 11789020; Gateway™ LR Clonase™ II Enzyme
160 mix, Thermo Fisher Scientific, 11791020) (Supplementary Table S1). pDONR207 was
161 used as the entry clone vector for both constructs (Thermo Fisher Scientific, 117207-
162 021). Subsequently, the pJCV51-pLUP5::LUP5 and pJCV51-pLUP5::eGFP plasmids were
163 generated by replacing the 35S promoter in pJCV51-p35S::LUP5 or pJCV51-p35S::eGFP
164 with the G-type *LUP5* promoter using the SalI sites. Finally, the plasmids were
165 introduced into *Agrobacterium tumefaciens* strain AGL1 via electroporation (Hayta et
166 al., 2021).

167 **Stable transformation in wild crucifer *B. vulgaris***

168 Seed sterilization was achieved by immersing seeds in 70% ethanol for 30 s,
169 followed by a 12.5 min treatment in 3% sodium hypochlorite (Thermo Fisher Scientific)
170 with 0.1% Tween20 (Bio-RAD, 1706531) and thorough rinsing with sterile water. Seeds
171 were sown on germination medium in Magenta boxes and incubated in darkness for
172 6–8 days (Supplementary Table S3).

173 Hypocotyl explants (~3 mm) from 6–10-day-old seedlings were incubated for
174 15 min with *A. tumefaciens* strains (OD_{600} = 0.5 in 10 mM $MgCl_2$, 100 µM
175 acetosyringone) with gentle agitation and then co-cultivated in darkness for three
176 days (Supplementary Table S3). The pJCV51-pLUP5::LUP5 and pJCV51-35S::LUP5
177 constructs were transformed into both *B. vulgaris* ecotypes, while pJCV51-
178 p35S::eGFP, pJCV51-pLUP5::eGFP, and pK7FWG2-CYP72A552 RNAi were introduced
179 only into the G-type.

180 After rinsing in Milli-Q water containing 1 µL/mL timentin, the explants were
181 transferred sequentially to callus induction (one week), shoot induction (two rounds,
182 two weeks each), and root induction media (two rounds, two weeks each) at 25°C
183 under a 16:8 h light–dark cycle (Supplementary Table S3). Throughout these stages,
184 the explants were maintained in a climate chamber at 25°C under a 16:8 h light-dark
185 cycle. Kanamycin (50 ng/mL) was used for selecting transformed explants, whereas
186 wild-type explants were cultured without kanamycin supplementation in the medium.

187 CYP72A552 RNAi plants were confirmed by NPTII (neomycin phosphotransferase
188 II) ELISA Kit (Agdia, PSP 73000/0288). All the other transgenic plants were confirmed
189 by RFP fluorescence (for *LUP5* expression and promoter activity constructs).
190 Fluorescence was analyzed on a Leica M205FA fluorescence dissection microscope
191 (Leica Microsystems). RFP and eGFP were imaged using the dsRed plant (excitation
192 546/10 nm; emission 600/40 nm) or ET GFP filters (excitation 470/40 nm; emission
193 525/50 nm), respectively. Subsequently, the confirmed plants were transferred to soil
194 and grown in a greenhouse at 19°C under a 16:8 h light–dark cycle.

195 **Choice insect feeding assay**

196 Insect resistance was evaluated using at least three individual plants (biological
197 replicates) per transgenic line. For each transgenic plant, five transgenic and five wild-
198 type leaf discs (1.57 cm^2 each) were alternately arranged in a petri dish (9.4 x 1.6 cm)
199 (Greiner BIO-ONE, 633180) with filter paper moistened with 2 mL of water. After a 3–
200 5-hour starvation period, ten third-instar *P. xylostella* larvae were placed at the center
201 of the plate, following a modified protocol from Liu et al (Liu et al., 2019). Leaf
202 consumption was quantified when half of the wild-type leaves were consumed using
203 ImageJ (approx 7.5 h).

204 **Preparation of plant extracts and metabolite analysis**

205 The extraction of saponins and sapogenins was adapted from Khakimov et al.
206 (2016). *B. vulgaris* leaf samples were ground with liquid nitrogen using a mortar and
207 pestle. For saponin detection, 100 mg of leaf powder was weighed into 1.5 mL
208 Eppendorf tubes and extracted with 300 µL of 85% methanol (v/v) via sonication for

209 30 minutes, followed by centrifugation at 16,000 × g for 10 minutes. Supernatants
210 (200 µL) were filtered through a 0.22-µm filter before LC-qTOF-ESI-MS/MS analysis.

211 To detect total and free sapogenins, 200 mg of leaf powder was weighed into 2
212 mL Eppendorf tubes and extracted with 600 µL of 85% methanol (v/v) at 100°C while
213 mixing at 1,400 × rpm in a thermomixer (Eppendorf, Denmark). The samples were
214 cooled on ice and centrifuged at 16,000 × *g* for 3 min. After filtration through a 0.22-
215 µm filter, 100 µL of supernatant was collected for free sapogenin detection via LC-
216 qTOF-APCI-MS/MS. An additional 300 µL of supernatant was collected and dried under
217 nitrogen flow. To cleave off sugar residues, samples were treated with 500 µL of 2 M
218 hydrochloric acid at 100°C while mixing at 1,400 × rpm for 1.5 hours in a thermomixer
219 (Eppendorf, Denmark). After cooling on ice, a double volume of ethyl acetate was
220 added to the acid-water mixture, followed by vortexing and centrifugation at 3,500
221 × *g* for 5 minutes. The samples were extracted three times with ethyl acetate to collect
222 sapogenins. To remove residual acid, an equal volume of Milli-Q water was added, the
223 samples were subsequently vortexed and centrifugated at 3,500 × *g* for 5 minutes.
224 This washing step was repeated three times. The final extracts were dried under
225 nitrogen flow, re-solubilized in 300 µL of 100% methanol, and filtered through a 0.22-
226 µm filter before total sapogenin detection via LC-qTOF-APCI-MS/MS.

227 LC-qTOF-ESI-MS/MS was performed according to Liu et al. (2019). LC-qTOF-APCI-
228 MS/MS was performed as described by Patel et al. (2020). Untargeted metabolite data
229 were analyzed using XCMS (Tautenhahn et al., 2012), and targeted saponin and
230 sapogenin analyses were conducted with DataAnalysis software (v.4.3; Bruker,
231 Bremen, Germany) following Khakimov et al. (2016). Hederagenin cellobioside and
232 oleanolic acid cellobioside standards were used for saponin identification. Lupeol, β-
233 amyrin, oleanolic acid, and hederagenin standards were employed for sapogenin
234 identification.

235 **Gene expression analysis**

236 Total RNA was extracted using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich,
237 STRN50-1KT) and then synthesized to cDNA by iScript™ cDNA Synthesis Kit (Bio-Rad,
238 1708891). Relative expression of *LUP5* was quantified via qPCR using kapa SYBR® fast

239 kit (Merck-Sigma, KK4607) (Khakimov et al., 2015). *Tubulin* (GenBank accession no.
240 EU555191) was used as reference gene, and *LUP5* expression levels were normalized
241 to the wild-type P-type control (Liu TongJin et al., 2016). Primer sequences are
242 provided in Supplementary Table S1 (Khakimov et al., 2015).

243 Results

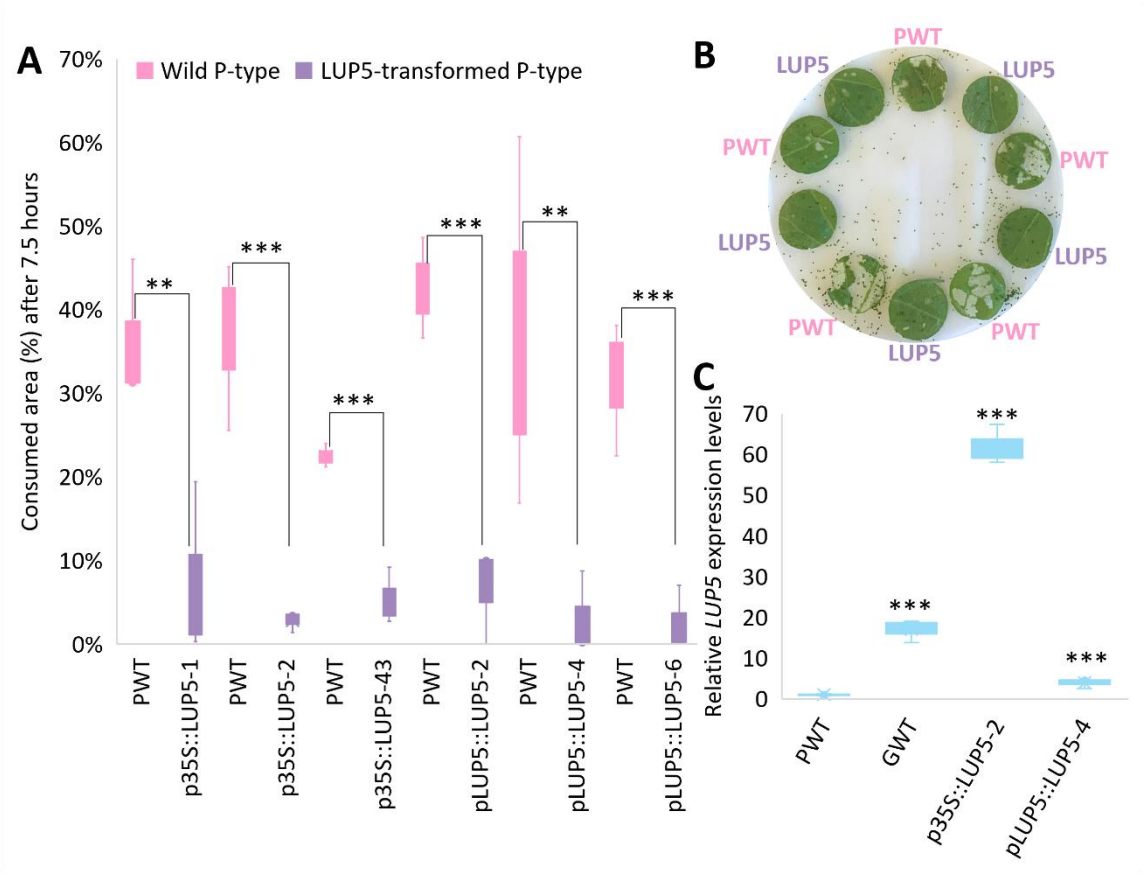
244 Establishment of a transformation and regeneration system in the wild crucifer *B.* 245 *vulgaris*

246 To enable functional studies of insect resistance in *B. vulgaris* as an ecological
247 model system, we developed a stable transformation and regeneration method for
248 both the G-type and P-type ecotypes (Fig.1, Supplementary Fig. S1, Supplementary
249 Table S3). To metabolically engineer the saponin compositions, we introduced the G-
250 type *LUP5* coding sequence into both ecotypes and attempted to silence *CYP72A552*
251 in the G-type. *LUP5* expression was driven either by the constitutive 35S promoter or
252 the native *LUP5* promoter. Transformation efficiency, calculated as the percentage of
253 positive transgenic plants to all obtained plants from tissue culture, reached up to 94%
254 across experiments (Table 1). The native *LUP5* promoter exhibited approximately two-
255 fold higher transformation efficiency than the 35S promoter in both the G- and P-type,
256 indicating a selection pressure against *LUP5* overexpression from the 35S promoter
257 (Table 1).

258 Stable expression of G-type *LUP5* in the P-type confers insect resistance

259 To evaluate the impact of G-type *LUP5* expression on insect resistance, a choice
260 insect feeding assay was conducted by presenting alternating wildtype and transgenic
261 *B. vulgaris* leaf discs to diamondback moth larvae. Leaf area consumed was quantified
262 when approximately half of the total leaf material of the wild type had been eaten
263 (approx. 7.5 hr). The best performing *LUP5*-transformed P-type plants exhibited a
264 significant reduction in larval consumption, ranging from 76 - 95% (Fig. 2, A and B).
265 QPCR on the best performing 35S and *LUP5* promoter lines, confirmed that the G-type
266 *LUP5* was transformed and expressed in the P-type both under the 35S promoter or
267 the native G-type *LUP5* promoter (Fig. 2C). Interestingly, the 35S promoter expressed

268 the *LUP5* at approximately sixteen-fold higher levels than that in the best pLUP5::LUP5
269 transformed P-type and the wildtype P- and G-type (Fig. 2C).



270

271 Fig. 2 Expression of the G-type *LUP5* in *B. vulgaris* P-type reduces feeding by
272 diamondback moth larvae. A) Consumption area (%) after 7.5 h (choice test, three
273 individual plants per transgenic line; each replicate represents one Petri dish
274 containing five transgenic and five wildtype leaf discs). B) Example of the insect choice
275 feeding assay, and the resulting leaf condition after 7.5 hr insect exposure. C) Relative
276 *LUP5* expression levels (three individual plants per transgenic line). PWT, wildtype P-
277 type *B. vulgaris*; GWT, wildtype G-type *B. vulgaris*; p35S::LUP5, p35S::LUP5
278 transformed P-type *B. vulgaris*; pLUP5::LUP5, pLUP5::LUP5 transformed P-type *B.*
279 *vulgaris*, LUP5, P-type *B. vulgaris* expressing the G-type *LUP5*. Statistical significance
280 was assessed using Student's *t*-test: *($p<0.5$), ** ($p < 0.05$) and *** ($p < 0.005$).

281 **Expression of G-type *LUP5* induced β-amyrin-derived saponins in P-type *B. vulgaris*.**

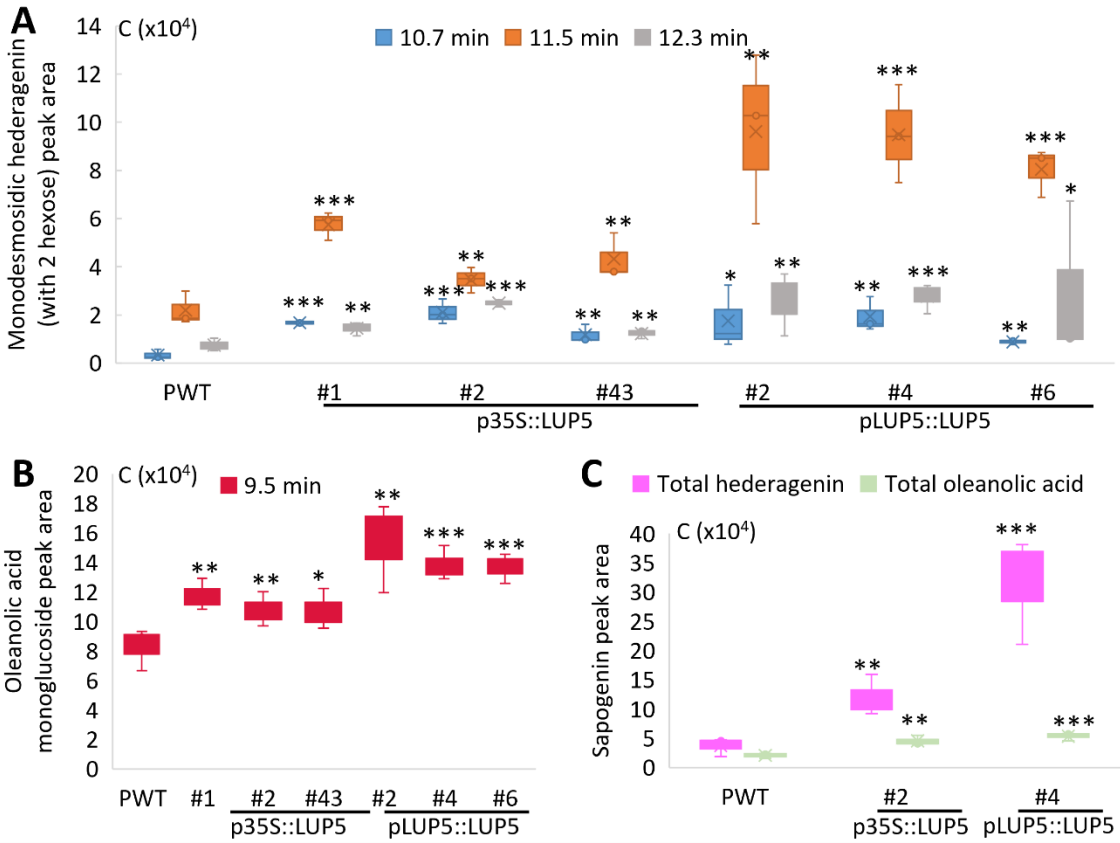
282 Based on the highest reduction of leaf area consumed, three transgenic lines

283 were selected for each promoter type (35S and native *LUP5*) and subjected to
284 untargeted and targeted LCMS analyses. The untargeted metabolite analysis revealed
285 approximately 9,500 distinct *m/z* features, each defined as a unique *m/z* and retention
286 time (RT) pair. Features differing between *LUP5*-expressing and wild-type plants (P <
287 0.5) were ranked by fold change, and the top 100 increased and 100 decreased
288 features were analyzed further. However, extraction of the *m/z* features using
289 DataAnalysis software revealed that the majority lacked clear MS^2 fragmentation
290 patterns, preventing structural annotation. Many of these features likely represent in-
291 source fragments, background ions, or uncharacterized metabolites with inconclusive
292 fragmentation.

293 Consequently, a targeted saponin analysis was conducted using DataAnalysis. *B.*
294 *vulgaris* saponins were detected by extracted ion chromatograms of our previously
295 tentatively identified aglycones and their structures were inferred from characteristic
296 fragmentation patterns. Based on 455 m/z, 469 *m/z,* and 471 m/z, the characteristic
297 fragmentation patterns of oleanolic acid, gypsogenin and hederagenin were
298 tentatively identified, respectively. This targeted approach identified ten putative
299 distinct saponins in P-type lines and thirty in G-type lines (Supplementary Table S4 and
300 S5). All detected saponins were classified as monodesmosidic, based on their
301 fragmentation patterns. Saponin abundance was subsequently quantified by
302 calculating peak areas from extracted ion chromatograms (Supplementary Table S4
303 and S5).

304 The LCMS analysis revealed three hederagenin-derived monodesmosidic
305 saponins (two hexoses; RT 10.7, 11.5, and 12.3 min) and one oleanolic acid
306 monoglucoside (one hexose; RT 9.5 min) that accumulated at significantly higher levels
307 in *LUP5*-transformed P-type compared to wild-type P-type (Fig. 3, A and B). The most
308 abundant of the three hederagenin-derived saponins eluted at 11.5 min and was
309 increased ~four-fold in the P-type expressing *LUP5* under the native *LUP5* promoter
310 compared to the wildtype and was ~two-fold more abundant than in P-type plants
311 transformed with *LUP5* under the 35S promoter (Fig. 3A). Hederagenin cellobioside
312 (eluted at 12.3 min) was also increased ~four-fold but remained at lower levels in
313 pLUP5::LUP5 transformed P-type (Fig. 3A). These three hederagenin-derived saponins

314 were also present in the control G-type plants, with hederagenin cellobioside (eluted
315 at 12.3 min) being the most abundant hederagenin-derived saponin (Supplementary
316 Table S5 and S8). The increased oleanolic acid monoglucoside (eluted at 9.5 min) was
317 not detected in G-type (Supplementary Table S5). The other detected saponins
318 remained unchanged (Supplementary Table S6).

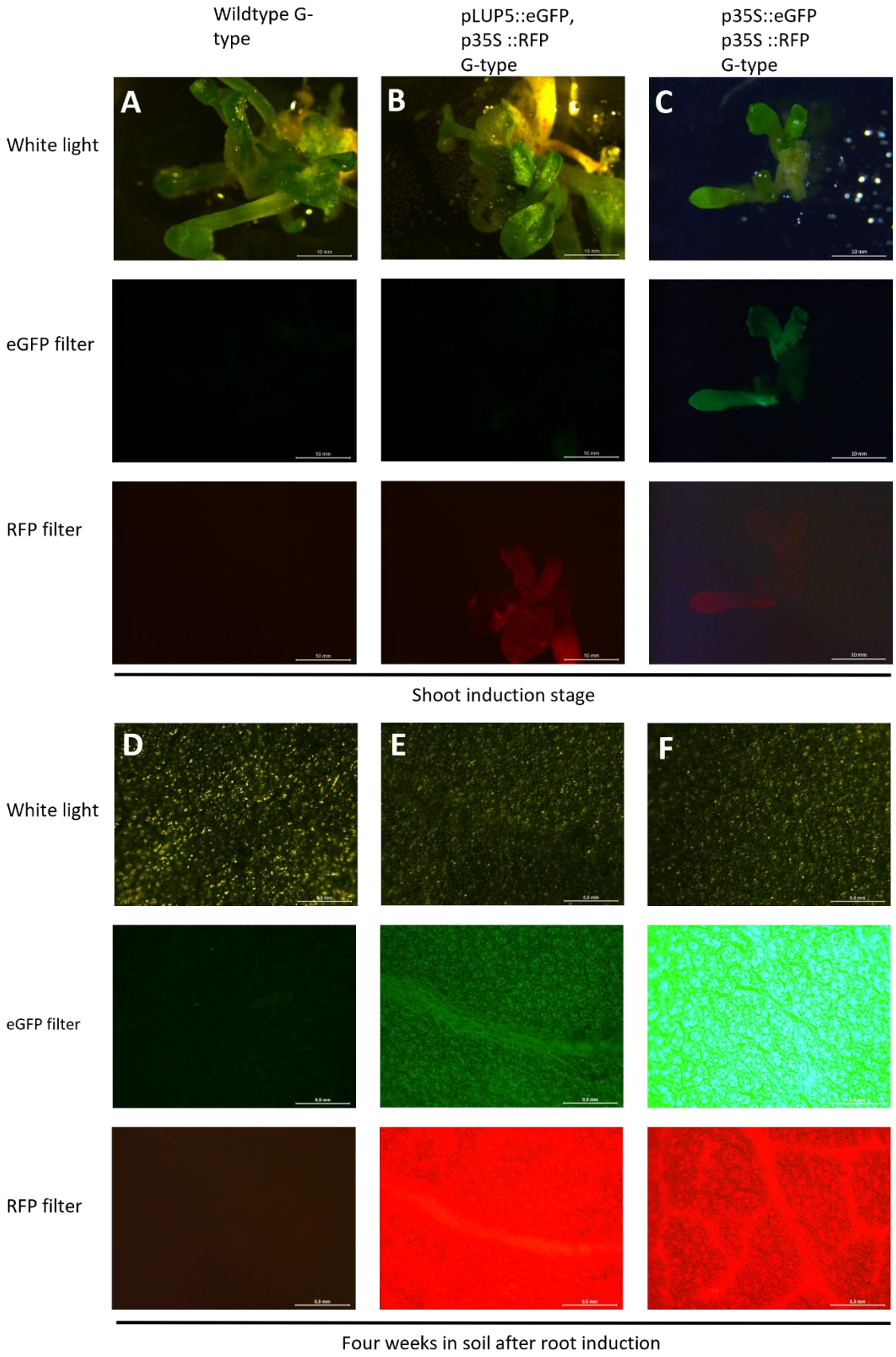


319

320 Fig. 3 Expression of the G-type *LUP5* increased glycosylated oleanolic acid and
321 hederagenin levels in the transgenic P-type *B. vulgaris*. A) Monodesmosidic
322 hederagenin (with 2 hexoses) peak area (EIC: *m/z* = 841.4591 ± 0.2). B) Oleanolic acid
323 monoglucoside peak area (EIC: *m/z* = 617.4059 ± 0.2). C) Peak area of sapogenins,
324 including total hederagenin (EIC: *m/z* = 437.3425 ± 0.2) and total oleanolic acid (EIC:
325 *m/z* = 439.3593± 0.2). PWT, wildtype P-type *B. vulgaris*; p35S::LUP5, p35S::LUP5
326 transformed P-type *B. vulgaris*; pLUP5::LUP5, pLUP5::LUP5 transformed P-type *B.*
327 *vulgaris*. Error bars represent the standard deviation of the mean from three
328 individual plants. Statistical significance was assessed using Student's *t*-test: *(p<0.5),
329 ** ($p < 0.05$) and *** ($p < 0.005$).

330 To evaluate the effect of G-type *LUP5* expression on P-type *B. vulgaris* saponin
331 aglycone biosynthesis and accumulation, levels of free and total sapogenins were
332 analyzed. Free sapogenins were measured from untreated samples, representing
333 aglycones that occur naturally in a non-glycosylated form. Total sapogenins were
334 defined as the sum of free and sugar-bound sapogenins, the latter released through
335 hydrochloric acid mediated hydrolysis of glycosylated saponins. A transgenic line (with
336 three vegetative clones) from both p35S::*LUP5* and pLUP5::*LUP5*-transformed P-type
337 *B. vulgaris* was selected based on reduced leaf area consumption and increased
338 accumulation of four saponins (Fig. 2A, Fig. 3, A and B). This analysis revealed a
339 significant increase in total hederagenin and oleanolic acid levels in *LUP5*-transformed
340 P-type plants (Fig. 3C), while total lupeol, β-amyrin, and free sapogenins remained
341 undetectable (Supplementary Table S7). Higher accumulation of total hederagenin
342 was observed in plants expressing *LUP5* under the native promoter compared to those
343 under the 35S promoter (Fig. 3C).

344 To investigate how promoter choice affects gene expression and metabolic
345 composition, we compared the previously described transgenic P-type lines
346 expressing *LUP5* under either the constitutive 35S or the native *LUP5* promoter.
347 Interestingly, although the 35S promoter drove much higher *LUP5* expression levels
348 compared to the native promoter (Fig. 2C), this did not result in proportionally
349 increased accumulation of β-amyrin–derived saponins (Fig. 3A–C). To investigate this,
350 we made eGFP lines with the *LUP5* and 35S promoter respectively. Analysis of these
351 lines the 35S promoter was already active at the shoot induction stage and
352 consistently drove stronger fluorescence signals, whereas the *LUP5* promoter only
353 became active only about four weeks after transfer to soil (Fig. 4A-F). These
354 observations indicate that the two promoters differ mainly in temporal activation and
355 relative expression strength, which likely contributes to the observed variation in
356 transcript levels, metabolite accumulation, and transformation efficiency.

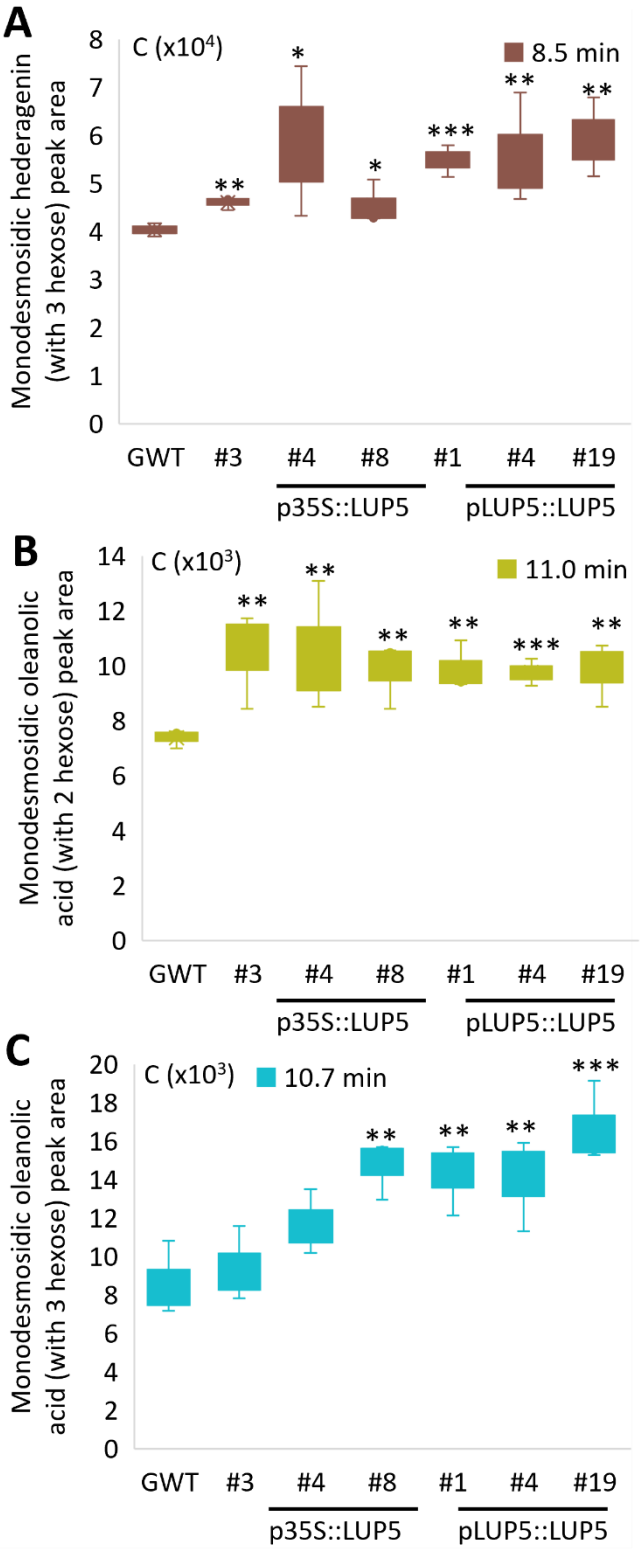


357

358 Fig. 4 Native *LUP5* promoter supports higher transformation efficiency via moderate

359 expression in *B. vulgaris*. During the shoot induction stage, no detectable eGFP
360 fluorescence was observed in the pLUP5::eGFP lines (B). Four weeks after root
361 induction, eGFP fluorescence could be detected in the pLUP5::GFP lines (D, E).
362 However, GFP expression driven by the 35S promoter was still much stronger, as
363 evident by the saturated image F, acquired with the same exposure time. RFP, driven
364 by the 35S promoter, was used for screening transgenic plants by microscopy. A, D:
365 Wild-type plants under white light, eGFP, and RFP filters; B, E: pLUP5::eGFP
366 transformed G-type plants under white light, eGFP, and RFP filters; C, F: p35S::eGFP
367 transformed G-type plants under white light, eGFP, and RFP filters; Scale bars: A–C, 10
368 mm, D–F, 0.5 mm.

369 Introduction of G-type *LUP5* in G-type increased a monodesmosidic hederagenin
370 (with three hexoses) and a oleanolic acid monoglucoside (with two hexoses) in both
371 p35S::LUP5 and pLUP5::LUP5 transformed G-type plants, retention times 8.5, and 11.0
372 min, respectively (Fig. 5, A to C). Monodesmosidic oleanolic acid (with 3 hexose,
373 retention time at 10.7 min) increased in all pLUP5::LUP5 transformed G-type lines (Fig.
374 5C). Twenty-six additional saponins remained unchanged (Supplementary Table S8),
375 and the three slightly elevated saponins remained at low abundance relative to other
376 saponins (Fig. 5A-C, Supplementary Table S8).



377

378 Fig. 5 Expression of G-type *LUP5* increased glycosylated oleanolic acid and hederagenin
379 levels in G-type *B. vulgaris*. A) Monodesmosidic hederagenin (with 3 hexose) peak area
380 (EIC: *m/z* = 1003.5115 ± 0.2). B) Monodesmosidic oleanolic acid (with two hexoses)
381 peak area (EIC: *m/z* = 617.4059 ± 0.2). C) Monodesmosidic oleanolic acid (with 3

382 hexose) peak area (EIC: *m/z* = 987.517 ± 0.2). GWT, wildtype G-type *B. vulgaris*;
383 p35S::LUP5, p35S::LUP5 transformed G-type *B. vulgaris*; pLUP5::LUP5, pLUP5::LUP5
384 transformed G-type *B. vulgaris*. Error bars represent the standard deviation of the
385 mean from three individual plants. Statistical significance was assessed using Student's
386 *t*-test: *(p<0.5), ** (p < 0.05) and *** (p < 0.005).

387 In conclusion, stable transformation of G-type *LUP5* in P-type significantly
388 increased one oleanolic acid monoglucoside (with one hexose) and three hederagenin-
389 derived monodesmosidic saponins (each with two hexoses) by approximately 50% to
390 fourfold, depending on the promoter and compound (Fig. 3, A and B). Notably, the
391 oleanolic acid monoglucoside, eluting at 9.5 minutes, was detected exclusively in the
392 transformed P-type (Fig. 3B, Supplementary Table S5). The most abundant
393 hederagenin-derived saponin in P-type eluted at 11.5 min, whereas hederagenin
394 cellobioside (eluted at 12.3 min) remained the dominant saponin in G-type (Fig. 3A,
395 Supplementary Table S6 and S8). Introduction of G-type *LUP5* in G-type slightly
396 increased saponin levels, but their concentrations remained low compared to other
397 oleanolic acid- and hederagenin-derived compounds (Fig. 5, A to C, Supplementary
398 Table S8). Expression of G-type *LUP5* by its native promoter resulted in higher
399 transformation efficiency and induced greater accumulation of β-amyrin-derived
400 sapogenins and their glycosylated saponins, despite lower LUP5 transcript levels
401 compared to the 35S promoter (Fig. 3, A to C, Fig. 5, A to C, Table 1, Supplementary
402 Table S7).

403 ***CYP72A552* silencing decreased to 40% hederagenin derived saponin accumulation**
404 **in G-type *B. vulgaris*.**

405 To evaluate the impact of G-type *CYP72A552* silencing on insect resistance, a
406 choice insect feeding assay was conducted as described above. However, neither the
407 transgenic G-type nor the wild-type G-type plants were consumed by diamondback
408 moth larvae as measure after 7.5 hours (Fig. 6G). To assess the impact of *CYP72A552*
409 downregulation on saponin and sapogenin levels in G-type *B. vulgaris*, a metabolomic
410 analysis was performed as described above. Silencing of *CYP72A552* reduced the
411 accumulation of six saponins in the G-type, including hederagenin monoglucoside and
412 three other hederagenin-derived monodesmosidic saponins (each with two hexoses),

413 monodesmosidic gypsogenin (with two hexoses), and one unknown saponin (Fig. 6, A
414 to E). MS/MS analysis indicated that the unknown saponin contains a sapogenin of
415 457 m/z with one hexose and two methylpentoses (Fig. 6E). Twenty-three other
416 saponins, including seven glycosylated oleanolic acids, remained unchanged
417 (Supplementary Table S9). When the sapogenin levels of *CYP72A552*-silenced line #13
418 and wild-type G-type *B. vulgaris* were compared, the analysis revealed a 40%
419 reduction in total hederagenin, while total and free oleanolic acid levels remained
420 unchanged, consistent with the saponin data (Fig. 6A–F).. Lupeol and β-amyrin
421 remained undetectable in G-type *B. vulgaris* (Supplementary Table S7). In conclusion,
422 silencing *CYP72A552* significantly reduced hederagenin biosynthesis and the
423 accumulation of its glycosylated derivatives in G-type *B. vulgaris*.

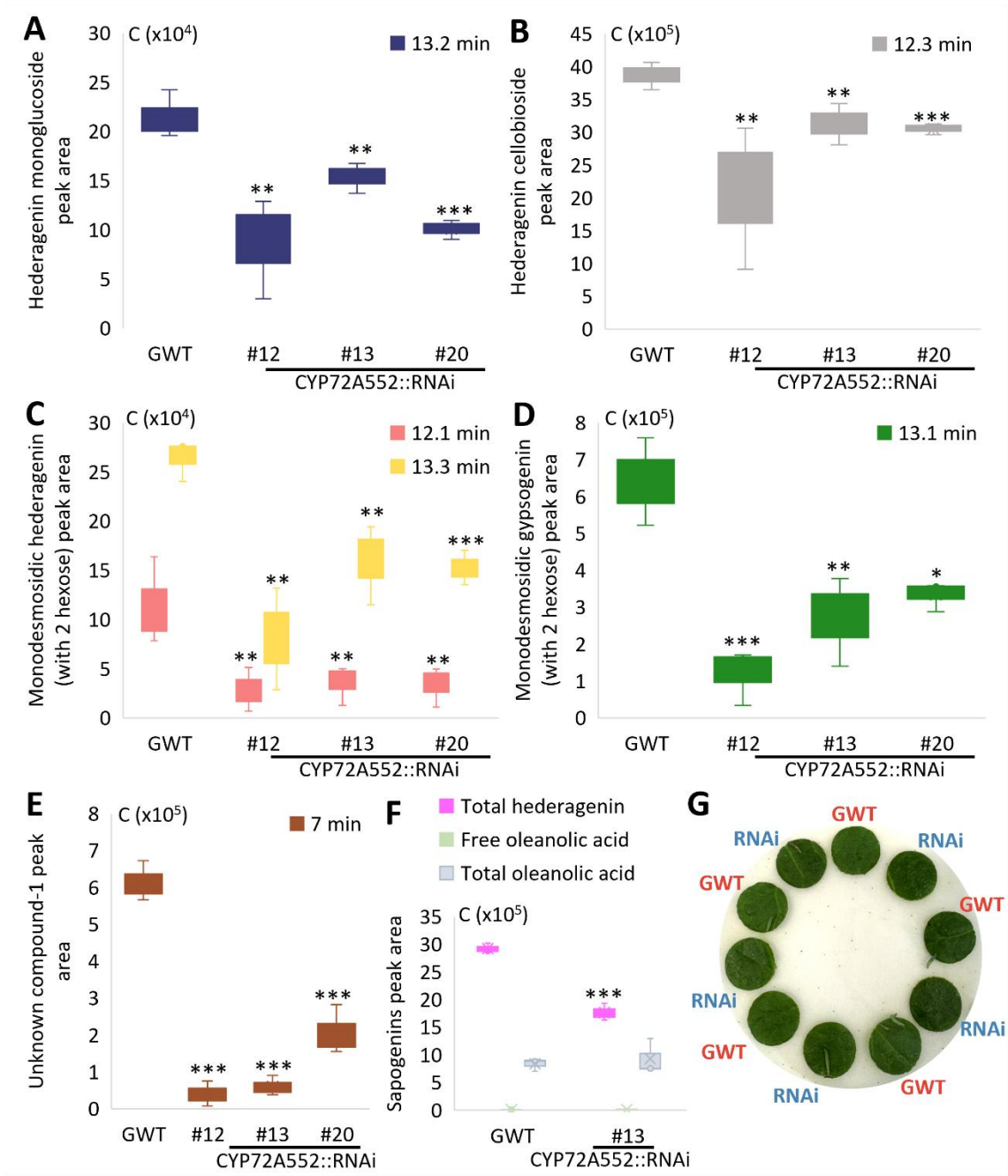


Fig. 6 Silencing *CYP72A552* decreased hederagenin and its derived saponins in G-type *B. vulgaris*. A) Hederagenin monoglucoside peak area (EIC: *m/z* = 679.4063 ± 0.2). B) Hederagenin cellobioside peak area (EIC: *m/z* = 841.4591 ± 0.2). C) Monodesmosidic hederagenin (with two hexoses) peak area (EIC: *m/z* = 841.4591 ± 0.2). D) Monodesmosidic gypsogenin (with two hexoses) peak area (EIC: *m/z* = 839.4409 ± 0.2). E) unknown compound-1 peak area (EIC: *m/z* = 901.2429 ± 0.2). F) Peak area of sapogenins, including total hederagenin (EIC: *m/z* = 437.3425 ± 0.2), free and oleanolic acid (EIC: *m/z* = 439.3593± 0.2). G) Example of the insect choice feeding assay, and the

433 resulting leaf condition after 7.5 hr insect exposure. GWT, wildtype P-type *B. vulgaris*;
434 CYP72A552::RNAi, *CYP72A552* silenced G-type *B. vulgaris*. Error bars represent the
435 standard deviation of the mean from three individual plants. Statistical significance
436 was assessed using Student's *t*-test: *(p<0.5), ** (p < 0.05) and *** (p < 0.005).

437 **Discussion**

438 Deciphering the genetic and biochemical bases of natural insect resistance is
439 essential for our basic understanding of chemical ecology and for addressing global
440 challenges in crop protection and sustainability. Our study contributes to this effort
441 by elucidating how structural diversification in plant specialized metabolites mediates
442 ecological defense and by identifying regulatory elements that can be harnessed for
443 future resistance breeding.

444 Triterpenoid saponins are key mediators of plant–insect interactions and have
445 attracted attention due to their high chemical diversity and structure-dependent
446 bioactivity and unique modes of action. Key factors influencing insect resistance
447 include the structure of the saponin, the composition of the sapogenin backbone, the
448 types of sugar moieties, the length and number of glycoside chains, and the sites of
449 glycosylation (Gao et al., 2011; Liu et al., 2019; Tian et al., 2021). Prior metabolic
450 engineering efforts have focused on modifying saponin composition by disrupting
451 biosynthetic genes in legumes (Confalonieri et al., 2021; Hodgins et al., 2024). For
452 example, Confalonieri et al. (2021) knocked out *CYP93E2*, which hydroxylates β-
453 amyrin at the C-24 position, in *Medicago truncatula*, resulting in transgenic lines that
454 no longer produced non-hemolytic soyasapogenols saponins but instead redirecting
455 metabolic flux toward the synthesis of hemolytic saponins. Hodgins et al.
456 (2024) knocked out β-amyrin synthase *BAS1* in pea (*Pisum sativum*), leading to a 99.8%
457 reduction in saponin content in the seeds. However, these studies did not address
458 how targeted saponin engineering influences ecological traits such as herbivore
459 resistance or consumption.

460 We have previously transiently expressed saponin biosynthetic genes in
461 *Nicotiana benthamiana* (Khakimov et al., 2015; Liu et al., 2019). An unexpected

462 drawback of this transient tobacco system was that hederagenin was further
463 metabolized by endogenous tobacco enzymes, which led to the accumulation of
464 hederagenin-3-*O*-monoglucoside to levels far below those naturally occurring in
465 *Barbarea vulgaris* (Liu et al., 2019). When oleanolic acid monoglucoside was
466 transiently produced in *N. benthamiana*, mainly non-toxic bidesmosidic saponins
467 accumulated, suggesting that the toxic monoglucosides were further glycosylated and
468 thereby detoxified by the plant (Khakimov et al., 2015). In addition, the diamondback
469 moth is a crucifer specialist and will not feed on *N. benthamiana*. To overcome these
470 major obstacles, we metabolically engineered the saponin biosynthetic pathway
471 directly in the wild crucifer *B. vulgaris* to test whether specific triterpenoid profiles
472 confer resistance to insect herbivory.

473 We hypothesized that differences in *LUP5* expression between G- and P-type *B.*
474 *vulgaris* underlie their distinct saponin profiles and insect resistance, based on
475 quantitative trait locus (QTL) analysis and insect feeding assays with saponin-treated
476 leaf disks (Augustin et al., 2012; Khakimov et al., 2015; Nielsen et al., 2010; Kuzina et
477 al., 2011). Our metabolic engineering strategy demonstrates that the susceptibility of
478 the P-type to diamondback moth feeding is primarily due to its lack of *LUP5*
479 expression. LUP5 encodes an oxidosqualene cyclase that produces β-amyrin, the
480 precursor of oleanolic acid and hederagenin-derived saponins, both implicated in
481 insect deterrence. Interestingly, our results further suggest that glycoside linkage
482 exerts less influence on feeding behavior than the sapogenin backbone itself.

483 Variation in oxidosqualene cyclase product specificity may be a key driver of
484 saponin-profile differences between the two *B. vulgaris* ecotypes. Insect-susceptible
485 P-type plants predominantly produce lupeol-derived saponins, reflecting their higher
486 expression of *LUP2*, while G-type plants express *LUP5* and accumulate β-amyrin-
487 derived saponins, including oleanolic acid and hederagenin cellobioside, both
488 contributing to insect resistance (Augustin et al., 2012; Khakimov et al., 2015; Liu et
489 al., 2019; Günther et al., 2021). By combining gene transfer with ecological assays, our
490 study provides a framework for linking specialized metabolism with ecological
491 function in a non-model plant system.

492 To test this hypothesis in an *in planta* system, the β-amyrin synthase *LUP5* was
493 expressed in both the insect resistant G-type and insect susceptible P-type *B. vulgaris*.
494 In the otherwise insect susceptible P-type, this resulted in up to a 95% reduction in
495 leaf consumption and increased levels of three hederagenin-derived monodesmosidic
496 saponins (each with two hexoses) and one oleanolic acid monoglucoside (Fig. 2A, Fig.
497 3, A and B). In wildtype *B. vulgaris* G-type, hederagenin cellobioside is the
498 predominant saponin, while only small amounts of the other two hederagenin-
499 derived saponins were detected at retention times of 10.7 and 11.5 min (Fig. 5, A and
500 B, Supplementary Table S8). In contrast, in the *LUP5*-transformed P-type, the
501 monodesmosidic hederagenin eluting at 11.5 min is more abundant than the other
502 two (Fig. 3A). This variation highlights differences in saponin glycosylation patterns
503 between P- and G-types, suggesting that these three hederagenin-derived
504 monodesmosidic saponins, particularly the one detected at 11.5 min, may contribute
505 to insect resistance in the transformed P-type plants. Hederagenin 3-*O*-glucoside
506 exhibits seven-fold stronger feeding reduction to diamondback moth and tobacco
507 hornworm compared to oleanolic acid 3-*O*-glucoside (Liu et al., 2019). This enhanced
508 efficacy has been linked to C-23 hydroxylation in hederagenin, which causes the C-3
509 glucose moiety to adopt a different orientation relative to the sapogenin backbone
510 compared to oleanolic acid (Cárdenas, Almeida and Bak, 2019; Liu et al., 2019).
511 Considering the bioactivity and the relatively low increase of oleanolic acid
512 monoglucoside compared to hederagenin-derived monodesmosidic saponins in *LUP5*-
513 transformed P-type suggest that there is a CYP72A552 ortholog in the P-type that
514 effectively transforms the oleanolic acid to hederagenin. This hederagenin is then
515 glycosylated by different UGTs than in the G-type *B. vulgaris* to other linkage types
516 than the 1→4 linkage cellobiosides that predominantly accumulate in the G-type.

517 Based on our findings, *LUP5* is identified as the key determinant for insect
518 resistance differences between G- and P-type *B. vulgaris*. Furthermore, three
519 hederagenin-derived monodesmosidic saponins (each with two hexoses) may play a
520 role in insect resistance in *LUP5*-transformed P-type, particularly the most abundant
521 saponin at retention time 11.5 min. These findings offer new mechanistic insight into
522 plant–insect coevolution and establish a model for dissecting genotype–phenotype

523 relationships in plant chemical defense. Future work comparing the insect resistance
524 differences between the elucidated structures of these three hederagenin-derived
525 monodesmosidic saponins could provide further insights into how glycosylation
526 patterns affect saponin bioactivity. Glycoside linkage impacts cytotoxicity of saponins
527 in *in vivo* assays (Chwalek et al., 2006). However, how glycoside linkage affects insect
528 resistance remains unknown. Unfortunately, due to their low concentration, isolating
529 and purifying these compounds for bioactivity testing remains challenging. Future
530 research should focus on identifying candidate glycosyltransferases that produces a
531 variety of glycoside linkage with the aim to produce defense-related hederagenin-
532 derived monodesmosidic saponins *in planta*.

533 Interestingly, when *LUP5* expression was driven by the constitutive 35S promoter,
534 *LUP5* was expressed at approximately sixteen-fold higher levels than in the best
535 *pLUP5::LUP5* transgenic line and also exceeded expression levels in both wild-type P-
536 and G-type plants (Fig. 2C). However, this high expression did not translate into
537 significantly increased saponin production (Fig. 2, A and B, Fig. 5, A to C). In contrast,
538 using the native *LUP5* promoter not only resulted in approximately two-fold higher
539 transformation efficiency (Table 1) but also led to greater accumulation of
540 hederagenin and its glycosides in the P-type background despite the overall much
541 lower expression profile (Fig. 3, A and C). These findings indicate that expression
542 strength alone does not determine metabolic output.

543 We propose that high and premature expression of *LUP5* under the 35S promoter
544 leads to early toxic β-amyrin accumulation, thereby reducing transformation
545 efficiency (Fig.2 and Fig. 3). In *B. vulgaris*, *LUP5* is expressed most strongly in leaves
546 compared with roots and petioles (Khakimov et al., 2015). To better understand the
547 observed differences in transformation efficiency, saponin and sapogenin
548 accumulation, and *LUP5* expression levels between the native and constitutive
549 promoters, we compared promoter activity and eGFP fluorescence intensity during
550 plant regeneration and early growth. Microscopy observations support this
551 interpretation: the 35S promoter was already active during shoot induction, whereas
552 the native *LUP5* promoter became detectable only four weeks after transfer to soil
553 following root induction (Fig. 2). Such premature activation likely causes β-amyrin to

554 accumulate before downstream pathway genes are expressed, disturbing metabolic
555 balance through substrate depletion or feedback inhibition. By contrast, delayed
556 activation under the native *LUP5* promoter allows coordinated expression with
557 downstream enzymes, promoting efficient accumulation of of saponins that are non-
558 toxic to the plant.

559 Thus, temporal coordination of gene expression emerges as a key determinant of
560 successful metabolic engineering. Although constitutive promoters such as 35S are
561 widely used to achieve strong expression, their untimed activity can impose metabolic
562 burdens and reduce transformation efficiency. Similar effects have been reported in
563 other plant systems: in cotton (*Gossypium hirsutum*), continuous 35S activity
564 throughout development caused unintended metabolic load (Sunilkumar et al., 2002),
565 while in maize (*Zea mays*), the *ZmUbi1* promoter achieved higher transformation
566 efficiency and better tissue-culture survival than 35S despite lower transcript levels
567 (Beringer et al., 2017). These findings reinforce that expression timing, rather than
568 transcriptional magnitude, is crucial for maintaining metabolic homeostasis and
569 efficient pathway flux.

570 Taken together, although the 35S promoter can drive high levels of transgene
571 expression, its continuous activity often disrupts metabolic balance, decreases
572 transformation efficiency, and induces pleiotropic effects on plant development. In
573 contrast, the native LUP5 promoter confines expression to appropriate developmental
574 stages, providing precise metabolic control and coordinated biosynthetic regulation.
575 These results highlight that promoter choice governs metabolic outcomes not merely
576 through expression strength but through temporal alignment with downstream
577 biosynthetic processes, thereby ensuring efficient *in planta* production of defense-
578 related hederagenin-derived monodesmosidic saponins.

579 We had anticipated that silencing of *CYP72A552* in G-type would reduce insect
580 consumption as hederagenin derived saponins are more bioactive than those derived
581 from oleanolic acid (Liu et al., 2019) . However, our results show that while silencing
582 of *CYP72A552* decreased hederagenin biosynthesis by up to 40%, leading to a
583 significant reduction in hederagenin- and gypsogenin (which is hederagening that has

584 been further oxidized at C-23 to the ketone)- derived monodesmosidic saponins.
585 Specifically, the bioactive hederagenin monoglucoside, hederagenin cellobioside and
586 monodesmosidic gypsogenin (with two hexoses) were decreased (Fig. 6, A to D). The
587 decrease in monodesmosidic gypsogenin and unknown compound 1 may have been
588 caused by the reduced hederagenin biosynthesis or off-target effects (Fig. 6, D and E,
589 Supplementary Fig. S2), possibly resulting from the high nucleotide sequence identity
590 of eight tandemly repeated *CYP72A* homologs (Liu et al., 2019). Previous studies
591 indicate that cellobiosides of oleanolic acid, hederagenin, and gypsogenin both
592 correlate with insect resistance in *B. vulgaris* in a concentration-dependent manner
593 (Shinoda et al., 2002; Agerbirk et al., 2003; Kuzina et al., 2009; Liu et al., 2019). The
594 hederagenin cellobioside concentration in G-type *B. vulgaris* leaves is ~eleven-fold
595 higher than the ED_{50} (50% effective dose) for *P. xylostella* (Shinoda et al., 2002; Liu et
596 al., 2019). Thus, even with a 40% reduction, the remaining saponin levels still exceed
597 the toxicity threshold required to deter feeding. Likewise, *LUP5*-transformed P-type
598 plants did not reach the full resistance potential of G-type plants, likely due to
599 incomplete accumulation of the most potent hederagenin saponins (Fig. 2, A and B,
600 Fig. 3, A to C, Fig. 6, A to G). Overexpression of *LUP5* in G-type plants produced only
601 marginal increases in minor oleanolic acid and hederagenin-derived saponins,
602 suggesting that G-type plants may already accumulate saponin at or close the
603 acceptable threshold for accumulation (Fig. 6, A to C, Supplementary Table S8).

604 **Conclusion**

605 In summary, this study demonstrates that targeted metabolic engineering of
606 triterpenoid saponins in the wild crucifer *Barbarea vulgaris* is a robust approach to
607 dissect the ecological roles of structurally diverse saponins and study insect resistance
608 mediated by specialized metabolism. Expression of the β-amyrin synthase *LUP5* under
609 its native promoter was more effective than constitutive expression, leading to higher
610 transformation efficiency and increased accumulation of insect-deterring
611 hederagenin-derived monodesmosidic saponins. This underscores the importance of
612 promoter choice and precise transcriptional regulation in optimizing specialized
613 metabolite production.

614 Our findings establish stable metabolic engineering of saponins not only as a tool
615 for biotechnological enhancement of resistance traits, but also as a powerful approach
616 to understanding the role of plant defense compounds in chemical ecology. In contrast
617 to transient expression systems, which often result in unstable metabolite
618 accumulation and further modification by host enzymes, stable engineering enables
619 consistent expression of pathway genes and reliable assessment of metabolite
620 function. We identified *LUP5* as a crucial determinant of insect resistance in *Barbarea*
621 *vulgaris* and demonstrated that three specific hederagenin-derived monodesmosidic
622 saponins, each containing two hexose units, are associated with this defense. Taken
623 together, this work introduces a novel application of metabolic engineering to unravel
624 the ecological functions of plant metabolites and provides promising strategies for
625 crop protection through fine-tuned manipulation of biosynthetic pathways.

626 Table 1. A stable transformation system in *B. vulgaris* enabled up to 94% efficiency.

Construct for transformation	Ecotype	Total numbers of plants	Positive/total
pJCV51-pLUP5::LUP5	G-Type	31	94%
	P-Type	35	60%
pJCV51-p35S::LUP5	G-Type	52	40%
	P-Type	66	26%
pJCV51-pLUP5::eGFP	G-Type	24	58%
pJCV51-p35S::eGFP	G-Type	20	55%
pK7FWG2-CYP72A552 RNAi	G-Type	46	28%

627

628 **Acknowledgements**

629 We thank Thure Hauser for providing Plutella *xylostella*, Niels Agerbirk for
630 providing wildtype *Barbarea vulgaris* G- and P-type seeds, Barry Cohen and Hanne
631 Volpin for their help in setting up the transformation system, and Jack Olsen and
632 Mariela Alejandra González Ramírez for their assistance in LCMS. This work was
633 financially supported by Novo Nordisk Foundation (EcoSap NNF20OC0060298 to SB),
634 the China Scholarship Council (CSC, 202108330041 to JS) and Marie Sklodowska-Curie
635 Individual Fellowship (MSCA-IF 752437 to PC).

636 **Conflicts of interest**

637 Authors declare that there are no conflicts of interest.

638 **Author contributions**

639 JS, PC, and SB designed the research; JS, PC, JG, SKN and SB analysed the data and
640 wrote the paper; JS, PC and SKN performed the research; JG provided the data analysis
641 strategy and standards.

642 **References**

643 **Agerbirk N, Hansen CC, Olsen CE, Kiefer C, Hauser TP, Christensen S, Jensen KR,**
644 **Orgaard M, Pattison DI, Lange CBA, Cipollini D, Koch MA** (2021) Glucosinolate
645 profiles and phylogeny in *Barbarea* compared to other tribe Cardamineae
646 (Brassicaceae) and *Reseda* (Resedaceae), based on a library of ion trap HPLC-
647 MS/MS data of reference desulfoglucosinolates. Phytochemistry **185:** 112658

648 **Agerbirk N, Olsen CE, Bibby BM, Frandsen HO, Brown LD, Nielsen JK, Renwick JAA**
649 (2003) A saponin correlated with variable resistance of *Barbarea vulgaris* to
650 the diamondback moth *Plutella xylostella*. Journal of Chemical Ecology **29:**
651 1417-1433

652 **Agerbirk N, Olsen CE, Nielsen JK** (2001) Seasonal variation in leaf glucosinolates and
653 insect resistance in two types of *Barbarea vulgaris ssp. arcuata*.
654 Phytochemistry **58:** 91-100

655 **Augustin JM, Drok S, Shinoda T, Sanmiya K, Nielsen JK, Khakimov B, Olsen CE, Hansen**
656 **EH, Kuzina V, Ekstrøm CT, Hauser T, Bak S** (2012) UDP-glycosyltransferases
657 from the UGT73C subfamily in *Barbarea vulgaris* catalyze sapogenin 3-*O*-
658 glucosylation in saponin-mediated insect resistance. Plant Physiology **160:**
659 1881-1895

660 **Augustin JM, Kuzina V, Andersen SB, Bak S** (2011) Molecular activities, biosynthesis
661 and evolution of triterpenoid saponins. Phytochemistry **72:** 435-457

662 **Beran F, Petschenka G** (2022) Sequestration of plant defense compounds by insects:
663 from mechanisms to insect–plant coevolution. Annual Review of Entomology
664 **67:** 163-180

665 **Beringer J, Chen W, Garton R, Sardesai N, Wang PH, Zhou N, Gupta MJ, Wu HX** (2017)
666 Comparison of the impact of viral and plant-derived promoters regulating
667 selectable marker gene on maize transformation and transgene expression.
668 Plant Cell Reports **36:** 519-528

669 **Byrne SL, Erthmann PO, Agerbirk N, Bak S, Hauser TP, Nagy I, Paina C, Asp T** (2017)
670 The genome sequence of *Barbarea vulgaris* facilitates the study of ecological
671 biochemistry. Scientific Reports **7:** 40728

672 **Cárdenas PD, Almeida A, Bak S** (2019) Evolution of structural diversity of
673 triterpenoids. Frontiers in Plant Science **10:** 1523

674 **Chwalek M, Lalun N, Bobichon H, Plé K, Voutquenne-Nazabadioko L** (2006)
675 Structure-activity relationships of some hederagenin diglycosides::
676 Haemolysis, cytotoxicity and apoptosis induction. Biochimica Et Biophysica
677 Acta-General Subjects **1760:** 1418-1427

678 **Confalonieri M, Carelli M, Gianoglio S, Moglia A, Biazzi E, Tava A** (2021) CRISPR/Cas9-
679 mediated targeted mutagenesis of *CYP93E2* modulates the triterpene saponin
680 biosynthesis in *Medicago truncatula*. Frontiers in Plant Science **12:** 690231

681 **Cui CJ, Yang YQ, Zhao TY, Zou KK, Peng CY, Cai HM, Wan XC, Hou RY** (2019)
682 Insecticidal activity and insecticidal mechanism of total saponins from *Camellia*
683 *oleifera*. Molecules **24:** 4518

684 **Dervishi M**, **Schmitt FM**, **Günther J**, **Cedergreen N**, **Bak** S (2025) Structure–activity
685 relationship of triterpenoid saponins based on toxicity towards organisms of
686 different phylogenetic origin. Journal of Chemical Ecology. In press.

687 **Dolma SK, Suresh PS, Singh PP, Sharma U, Reddy SGE** (2021) Insecticidal activity of
688 the extract, fractions, and pure steroidal saponins of *Trillium govanianum* Wall.
689 ex D. Don for the control of diamondback moth (*Plutella xylostella* L.) and aphid
690 (*Aphis craccivora* Koch). Pest Management Science **77:** 956-962

691 **Erthmann PO, Agerbirk N, Bak S** (2018) A tandem array of UDP-glycosyltransferases
692 from the UGT73C subfamily glycosylate sapogenins, forming a spectrum of
693 mono- and bisdesmosidic saponins. Plant Molecular Biology **97:** 37-55

694 **Fürstenberg-Hägg J, Zagrobelny M, Bak S** (2013) Plant defense against insect
695 herbivores. International Journal of Molecular Sciences **14:** 10242-10297

696 **Gao GC, Lu ZX, Tao SH, Zhang S, Wang FZ** (2011) Triterpenoid saponins with
697 antifeedant activities from stem bark of *Catunaregam spinosa* (Rubiaceae)
698 against *Plutella xylostella* (Plutellidae). Carbohydrate Research **346:** 2200-2205

699 **Günther J, Erthmann PØ, Khakimov B, Bak S** (2021) Reciprocal mutations of two
700 multifunctional β-amyrin synthases from *Barbarea vulgaris* shift α/β-amyrin
701 ratios. Plant Physiology **188:** 1483-1495

702 **Hauser TP, Toneatto F, Nielsen JK** (2012) Genetic and geographic structure of an
703 insect resistant and a susceptible type of *Barbarea vulgaris* in western Europe.
704 Evolutionary Ecology **26:** 611-624

705 **Hayta S, Smedley MA, Clarke M, Forner M, Harwood WA** (2021) An efficient
706 *Agrobacterium*-mediated transformation protocol for hexaploid and
707 tetraploid wheat. Current Protocols **1:** e58

708 **Hodgins CL, Salama EM, Kumar R, Zhao Y, Roth SA, Cheung IZ, Chen JY, Arganosa GC,**
709 **Warkentin TD, Bhowmik P, Ham BK, Ro DK** (2024) Creating saponin-free
710 yellow pea seeds by CRISPR/Cas9-enabled mutagenesis on β-amyrin synthase.
711 Plant Direct **8:** e563

712 **Huang FQ, Dong XS, Yin XJ, Fan Y, Fan YM, Mao CC, Zhou W** (2020) A mass
713 spectrometry database for identification of saponins in plants. Journal of
714 Chromatography A **1625:** 461296

715 **Hussain M, Debnath B, Qasim M, Bamisile BS, Islam W, Hameed MS, Wang L, Qiu D**
716 (2019) Role of saponins in plant defense against specialist herbivores.
717 Molecules **24:** 2067

718 **Karimi M, Inzé D, Depicker A** (2002) GATEWAY™ vectors for *Agrobacterium*-mediated
719 plant transformation. Trends in Plant Science **7:** 193-195

720 **Khakimov B, Amigo JM, Bak S, Engelsen SB** (2012) Plant metabolomics: resolution and
721 quantification of elusive peaks in liquid chromatography–mass spectrometry

722 profiles of complex plant extracts using multi-way decomposition methods.
723 Journal of Chromatography A **1266:** 84-94

724 **Khakimov B, Kuzina V, Erthmann PØ, Fukushima EO, Augustin JM, Olsen CE,**
725 **Scholtalbers J, Volpin H, Andersen SB, Hauser TP, Muranaka T, Bak S** (2015)
726 Identification and genome organization of saponin pathway genes from a wild
727 crucifer, and their use for transient production of saponins in *Nicotiana*
728 *benthamiana*. The Plant Journal **84:** 478-490

729 **Khakimov B, Tseng LH, Godejohann M, Bak S, Engelsen SB** (2016) Screening for
730 triterpenoid saponins in plants using hyphenated analytical platforms.
731 Molecules **21:** 1614

732 **Kuzina V, Ekstrøm CT, Andersen SB, Nielsen JK, Olsen CE, Bak S** (2009) Identification
733 of defense compounds in *Barbarea vulgaris* against the herbivore Phyllotreta
734 nemorum by an ecometabolomic approach. Plant Physiology **151:** 1977-1990

735 **Kuzina V, Nielsen JK, Augustin JM, Torp AM, Bak S, Andersen SB** (2011) *Barbarea*
736 *vulgaris* linkage map and quantitative trait loci for saponins, glucosinolates,
737 hairiness and resistance to the herbivore *Phyllotreta nemorum*.
738 Phytochemistry **72:** 188-198

739 **Liu Q, Khakimov B, Cárdenas PD, Cozzi F, Olsen CE, Jensen KR, Hauser TP, Bak S** (2019)
740 The cytochrome P450 CYP72A552 is key to production of hederagenin-based
741 saponins that mediate plant defense against herbivores. New Phytologist **222:**
742 1599-1609

743 **Liu TongJin LT, Zhang XiaoHui ZX, Yang HaoHui YH, Agerbirk N, Qiu Yang QY, Wang**
744 **HaiPing WH, Shen Di SD, Song JiangPing SJ, Li XiXiang LX** (2016) Aromatic
745 glucosinolate biosynthesis pathway in *Barbarea vulgaris* and its response to
746 Plutella xylostella infestation. Frontiers in Plant Science **7:** 83

747 **Moses T, Papadopoulou KK, Osbourn A** (2014) Metabolic and functional diversity of
748 saponins, biosynthetic intermediates and semi-synthetic derivatives. Critical
749 Reviews in Biochemistry and Molecular Biology **49:** 439-462

750 **Nielsen JK, Nagao T, Okabe H, Shinoda T** (2010) Resistance in the plant, *Barbarea*
751 *vulgaris*, and counter-adaptations in flea beetles mediated by saponins.
752 Journal of Chemical Ecology **36:** 277-285

753 **Patel VS, Chhalotiya UK, Patel SB, Nuruddin J** (2020) Simultaneous quantification of
754 betulinic acid, lupeol, and β-sitosterol in *Madhuca longifolia* methanolic
755 extract of bark by liquid chromatography-tandem mass spectrometric method.
756 Journal of AOAC International **104:** 498-505

757 **Shinoda T, Nagao T, Nakayama M, Serizawa H, Koshioka M, Okabe H, Kawai A** (2002)
758 Identification of a triterpenoid saponin from a crucifer, *Barbarea vulgaris*, as a
759 feeding deterrent to the diamondback moth, *Plutella xylostella*. Journal of
760 Chemical Ecology **28:** 587-599

761 **Stevenson PC, Nyirenda SP, Veitch NC** (2010) Highly glycosylated flavonoids from the
762 pods of *Bobgunnia madagascariensis*. Tetrahedron Letters **51:** 4727-4730

763 **Sunilkumar G, Mohr L, Lopata-Finch E, Emani C, Rathore KS** (2002) Developmental
764 and tissue-specific expression of CaMV 35S promoter in cotton as revealed by
765 GFP. Plant Molecular Biology **50:** 463-479

766 **Tautenhahn R, Patti GJ, Rinehart D, Siuzdak G** (2012) XCMS online: a web-based
767 platform to process untargeted metabolomic data. Analytical Chemistry **84:**
768 5035-5039

769 **Tian X, Li Y, Hao N, Su X, Du J, Hu J, Tian X** (2021) The antifeedant, insecticidal and
770 insect growth inhibitory activities of triterpenoid saponins from *Clematis*
771 *aethusifolia* Turcz against *Plutella xylostella* (L.). Pest Management Science **77:**
772 455-463

773 **Xu R, Fazio GC, Matsuda SPT** (2004) On the origins of triterpenoid skeletal diversity.
774 Phytochemistry **65:** 261-291

775 **Yactayo-Chang JP, Tang HV, Mendoza J, Christensen SA, Block AK** (2020) Plant
776 defense chemicals against insect pests. Agronomy **10:** 1156

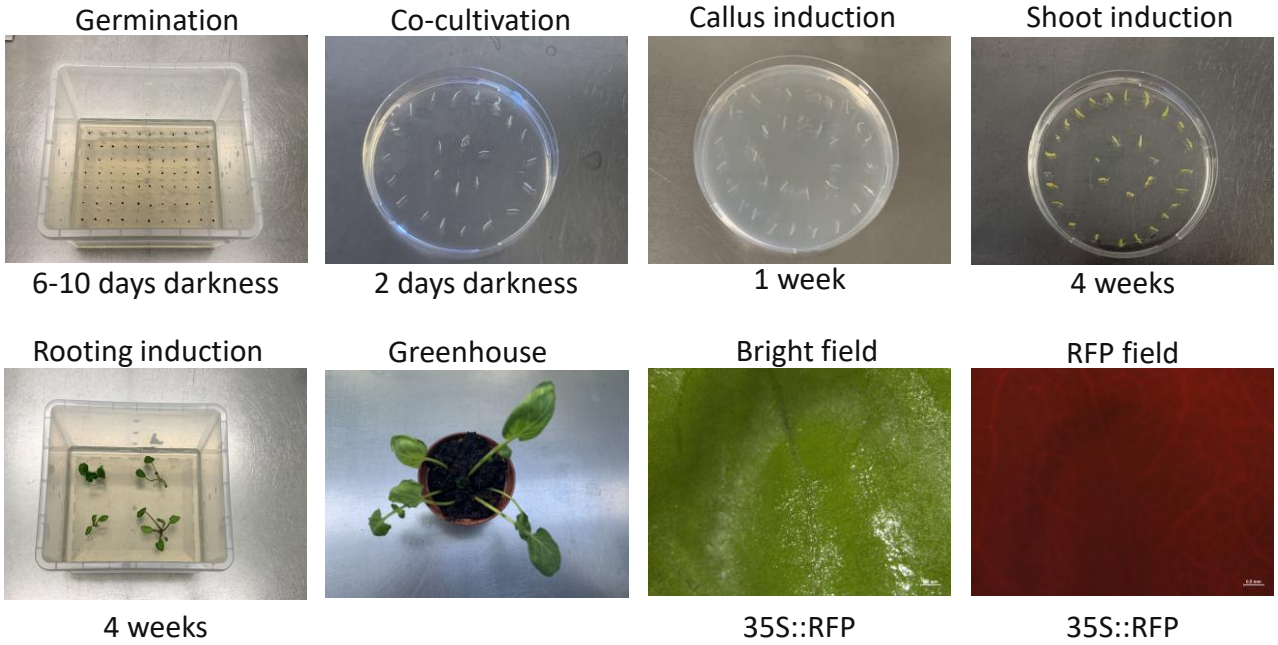
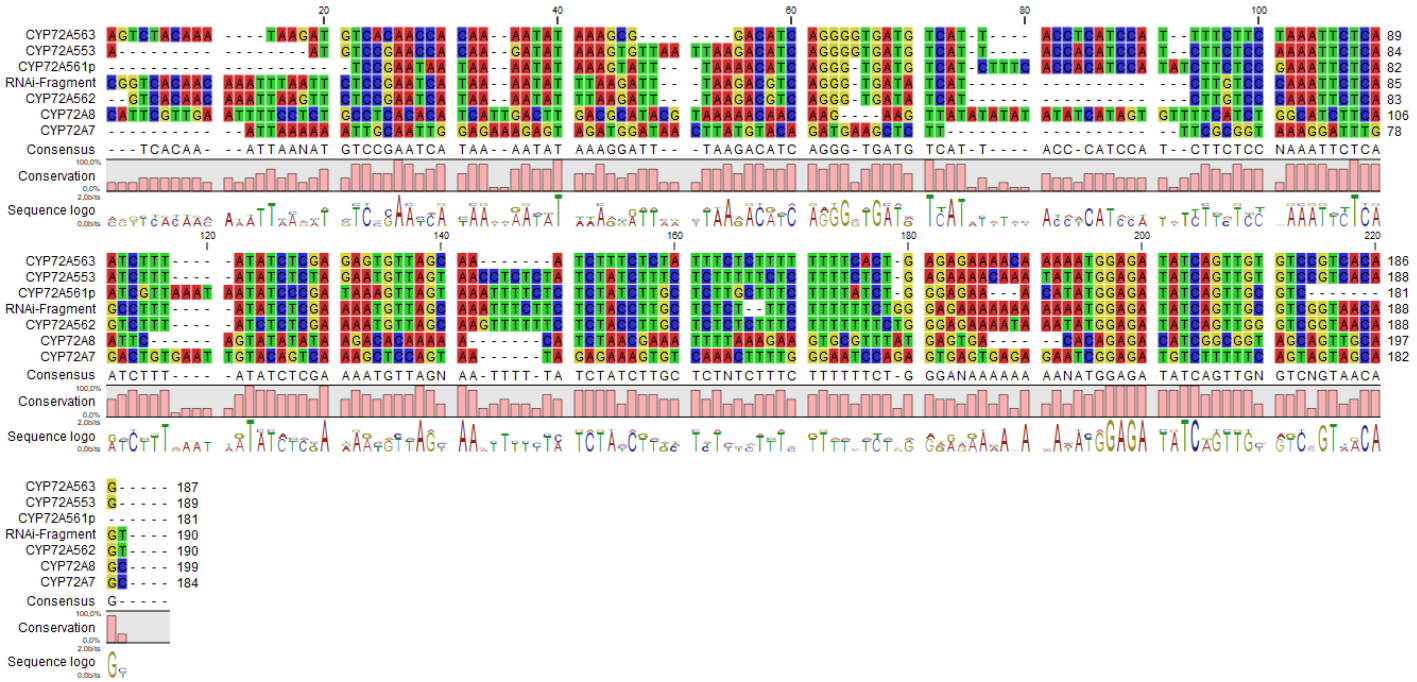


Fig. S1 *Agrobacterium*-mediated stable plant transformation process.

Fig. S2 Nucleotide sequence alignment between the RNAi fragment and eight tandemly repeated *CYP72A* genes in *B. vulgaris*. Accession numbers of the eight CYP72A genes range from MH252567 to MH252574.



Tab. S1 Primers used in this study

	Oligo name	Sequence 5'-3'
Amplication	LUP5_Fw	ATGTGGAGGTTGAAGTTAGGAGAGG
	LUP5_Rv	GCTGCGTTCGTCACTCATAAACATTAA
	LUP5_Promoter_Fw	GGATACGGATACGGATACGGAT
	LUP5_Promoter_Rv	GTTGTACTCTACATATAGATCATCATT
	eGFP-Fw	ATGGTGAGCAAGGGCGA
	eGFP-Rv	TTACTTGTACAGCTCGTCCATG
Gateway	LUP5_attB1_Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGTGGAGGTTGAAGTTAGGAGAGG
	LUP5_attB2_Rv	GCTGCGTTCGTCACTCATAAACATTAAACCCAGCTTTCTTGTACAAAGTGGTCCCC
	CYP72A552_attB1_Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGGTCACAACAAATTTAATTCTCCGAATC
	CYP72A552_attB2_Rv	GAGATATCAGTTGCGTCGGTAACAGTACCCAGCTTTCTTGTACAAAGTGGTCCCC
	eGFP_attB1_Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGGTGAGCAAGGGCGA
	eGFP_attB2_Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTTTACTTGTACAGCTCGTCCATG
qPCR	Tubulin qPCR Fw	GGAGATGTTTAGGCGTGTG
	Tubulin qPCR Rv	GCGTCTTGGTATTGCTGGT
	LUP5 qPCR Fw	CTGAGCGTGAGTACGTGGAA
	LUP5 qPCR Rv	GATCTGGATAGAGTTGCTGGA

Tab. S2 Nucleotide sequence of the G-type LUP5 promoter

GGATACGGATACGGATACGGATACGGATAGTAGAATAACGGATAATACGAATACGGATACAGATACGGATATC
CTGAAAAATCCGGATATCCGTATCCGTAACAGGGCTATTACTGACCCCACCCGTCCGGCAGTAGAGTCCGTC
TACGTCCTGCGTCCCCTCATTCATTGTACCTATGACAAAGACGCATAGTTATTAGAGCTACACCGCCGCGACT
TTGCTGCTCCTCCACTCATACATTATCATAGCTAAGTTCTCAAAAAGATTCAATCAAAGCCCTTCTCTTTGTATGG
GACCATTTTACACGTGATTATGTATCAACGTTACTTGCTCATTGACAAAATTGAGTTAGGTCTCTCCTCTTTTGCT
CTATTTTACTTCTATATTTGTCAATTTACTCTACTATGCATCTAATTATGACGGTACATGAAGTCCTTATGGTTAGGTT
TGCCACCGGTTGTGATTACAAATTTATTAGAATTTATCCCATGGTTAGAGAATTTATAAGTTTTTTTTCATTAGTTATA
TTATAATATTTTCTAAGTATAAGGTTGAATATGAGTATTAAAAAAAAAAAGAGAGAGAACTAACAATGGATAAAATAT
CATATGACTGAAATTTTTTTATACATTCCGTAAGGAACTAAGGATGATTATTAGTATAAAGATTCATTGCATCAAAC
AATTATGGTTAAGTAGTTAATTTACATCAATTTGGTATAATCAACATCCAATTTAAAAAATACGCGTCACACAAAGA
TTAAACCCAAATACACATGTGATAAAGAGAAGAACAAAGATGACCATCAATTACGCTGGTTGTCTTTGCATGAAT
GGAACACCTAATCGCGGGAAACTTTTTGAATCAGCGACTCAAACATACTTCGTTTCTACATAGGTGTCATTCAA
TATAATTACATATGTAGGTAGCTAGTTTTAACGGACTGATCATAATTTTTGTAGGTTTTTTTTTTTTAATAACTAAAAG
GTTTTAGGTCTAAGCCTGTACATCCTTTTAGACCCAGGACCAAAGCAGGTTATATGGTCTCGTCCCATGAGCC
ACTCGAACCAACGACCTCTAACCTTTGCCTGAGATATTTACCAGTTGAGTTATTAGCTCTTGGTTAATTTTTATAG
GTTTAAATCAAATGTATCTCATAATGATCGATATAACAAACAATAAACAAAAAAAGTTTGAAGGTGAATTTTGCAC
CCAAAAAAAAAAAAAAACTTTTTATTATCTTTTTCATTTTTTTTATAAATCACGGGAGTTCTAAACCTTTGACCCGTA
ATCTTTTTCGGTATCCGGAACCAGCCTTTACTCTATTTCTATTTTAAAAATGTCTCTCAGTACCAATGGAGCCGG
GTTTGAACCCGTGAAACAAAAATATTTATTTATTATTTATTATTATATTTTGGAAAGAACGTAAAAGTTATTTCAACGA
AAAAAATGTGGGAATATAAATAAAAAAAAAATTCAATATAATATTAAATTATTTACCAAACTGTACTCTTATAAAGTTA
TTTGATTCGTCGTTACCTCTATATAAGCGTACGTACATGAAACTCTTCCAAAGTATTGGTGATAAAGGCAAACAT
CGACCCAGCACAAGTACTCTCTCTTTCTTTCTCTTATCTCACTTTTCCTCTGTTTCACATATATGTATTCGTTACAT
GAGGGCTCTTATAATATATAAAGACATGCATATAAAATCATATATATATATATATATGTATGTAAATATAGTGTTCTATAT
ATAATGTTAATCTATGTATACGGTATACCGTATGAATGTACACTCACTACATCATATATCATCATGCGGTTATGAAAT
TCTCAATTCAGTGAGTATAATTTGGACATATACGTATAGATAGCGTCTAAATTTTTGTAATTTTTTAAAAAAAAATAAA
TGAATGACGGTTTGAAACGAACATAACAATAACATATATCGATATTTTGTTGTACTCTACATATAGATCATCATT

Tab. S3 Meida for tissue culture

	Germination		Co-cultivation		Callus induction		Shoot induction		Root induction	
Media (800 mL)	MS (M0222)	3.52 g	MS (M0222)	3.52 g	Gamborg B5	3.16 g	Gamborg B5	3.16 g	Gamborg B5	3.16 g
	Sucrose (1%)	8 g	Sucrose (1%)	8 g	MES	400 mg	MES	400 mg	MES	400 mg
	pH 5.8 (with KOH)		pH 5.8 (with KOH)		Glucose	16 g	Glucose	16 g	Glucose	16 g
	Agar (0.8%)	6.4 g	Agar (0.8%)	6.4 g	pH 5.8 (with KOH)		pH 5.8 (with KOH)		pH 5.8 (with KOH)	
					Phytagel	2.4 g	Phytagel	2.4 g	Phytagel	2.4 g
Adding before using (800 mL)			AS	800 µL	Kinetin	80 µl	Zeatin	800 µl	NAA	108 µl
					2,4-D	400 µl	NAA	80 µl	IAA	50 µl
					Timentin	800 µl	Timentin	800 µl	Timentin	800 µl
					Kanamycin	800 µl	Kanamycin	800 µl	Kanamycin	800 µl

Note: Chemical information: MS, Murashige & Skoog medium including vitamins (Duchefa, M0222.0050); Gamborg B5 (Duchefa, G0210.0050). Stocks concentration: Kanamycin: 50 mg/mL, Timentin: 200 mg/mL, Kinetin: 1 mg/mL, Zetin: 1 mg/mL , NAA (naphthalene acetic acid): 0.93 mg/mL, IAA (indole-3-aceti cacid): 1 mg/mL, 2,4-D (2,4-dichlorophenoxyacetic acid): 1 mg/mL , AS (Acetosyringone): 19.62 mg/mL.

Tab. S4 Saponins detected in P-type *B. vulgaris*. EIC, extract ion chromatogram

Sapogenin	Saponin fragment	Sugar moieties	Formic acid	Fragmentation pattern in Mass Spectra	EIC±0.2	Retention time (min)		
457	457,619,781,943, 989	3 x hexose	Yes	[M_{989}-46-162-162-162=457]	989.5404	8.6	8.9	
447	447,609,755,901	1 x hexose; 2 x methylpentose	No	[M_{901}-146-146-162=447]	901.2429	6.3		
473	473,635,797,959, 1005	3 x hexose	Yes	[M_{1005}-46-162-162-162=473]	1005.5337	8.1	8.4	
473	473,635,797,843	2 x hexose	Yes	[M_{843}-46-162-162=473]	843.4798	8.6		
471	471,633,795,841	2 x hexose	Yes	[M_{841}-46-162-162=471]	841.4591	10.7	11.5	12.3
455	455,617	1 x hexose	No	[M_{617}-162=455]	617.4059	9.5		

Tab. S5 Saponins detected in G-type *B. vulgaris*. EIC, extract ion chromatogram

Sapo genin	Saponin fragment	Sugar moieties	Formic acid	Fragmentation pattern in Mass Spectra	EIC±0.2	Retention time (min)				
457	457,619,781,943, 989	3 x hexose	Yes	[M_{989}-46-162-162-162=457]	989.5404	8.6	8.9	9.1		
447	447,609,771,917, 1079,1225	3 x hexose ; 2 x methylpentose	No	[M_{1225}-146-162-146-162-162=447]	1225.348	5.3				
447	447,609,755,901	1 x hexose; 2 x methylpentose	No	[M_{901}-146-146-162=447]	901.2429	7.0				
649	649,811,973,1019	2 x hexose	Yes	[M_{1019}-46-162-162=649]	1019.5076	7.9				
649	649,811,857	1 x hexose	Yes	[M_{857}-46-162=649]	857.4543	8.4				
469	469,631,793,839	2 x hexose	Yes	[M_{839}-46-162-162=469]	839.4409	13.1				
473	473,635,797,959, 1005	3 x hexose	Yes	[M_{1005}-46-162-162-162=473]	1005.5337	8.4				
473	473,635,797,843	2 x hexose	Yes	[M_{843}-46-162-162=473]	843.4798	8.0				
471	471,633,795,841	2 x hexose	Yes	[M_{841}-46-162-162=471]	841.4591	10.7	11.5	12.1	12.3	13.3
471	471,633,795	2 x hexose	No	[M_{795}-162-162=471]	795.4536	9.5	11.7			
471	471,603,749,795	1 x pentos; 1 x methylpentose	Yes	[M_{795}-46-146-132=471]	795.4536	13.8				
471	471,633,679	1 x hexose	Yes	[M_{679}-46-162=471]	679.4063	13.2				
471	471,633,795,957, 1003	3 x hexose	Yes	[M_{1003}-46-162-162-162=471]	1003.5115	8.5	9.3			
455	455,617	1 x hexose	No	[M_{617}-162=455]	617.4059	15.9				
455	455,617,779,941	3 x hexose	No	[M_{941}-162-162-162=455]	941.5115	10.2				
455	455,617,779,941, 987	3 x hexose	Yes	[M_{987}-46-162-162-162=455]	987.517	10.7				
455	455,617,779	2 x hexose	No	[M_{779}-162-162=455]	779.4587	11.0	11.5	14.2		
455	455,617,779,825	2 x hexose	Yes	[M_{825}-46-162-162=455]	825.4642	11.3	13,3	14.5		

Tab. S6 Non significantly peak area changed saponins in LUP5 transformed P-type *B. vulgaris* compared to wildtype P-type. EIC, extract ion chromatogram; PWT, wildtype P-type *B. vulgaris*; p35S::LUP5, p35S::LUP5 transformed P-type *B. vulgaris*.

Fragments	Sugar moieties	EIC ± 0.2	Retention time (min)	Peak area						
				PWT	p35S::LUP5-1	p35S::LUP5-2	p35S::LUP5-43	pLUP5::LUP5-2	pLUP5::LUP5-4	pLUP5::LUP5-6
457,619,781,943,989	3 x hexose	989.5404	8.6	104397 ± 14484	76740 ± 6720	78388 ± 5652	116101 ± 7003	105183 ± 12223	96111 ± 4826	96408 ± 6982
457,619,781,943,989	3 x hexose	989.5404	8.9	103949 ± 22256	160286 ± 11490	164287 ± 6733	90535 ± 2995	93870 ± 16808	120504 ± 23440	112588 ± 6319
447,609,755,901	1 x hexose + 2 x methylpentose	901.2429	6.3	514487 ± 15989	515685 ± 14393	681218 ± 36231	456469 ± 39764	460209 ± 82384	491958 ± 19291	399510 ± 15858
473,635,797,843	2 x hexose	843.4798	8.6	379938 ± 129721	767152 ± 82948	793761 ± 38062	373927 ± 25068	283426 ± 26935	497737 ± 137792	413195 ± 21207
473,635,797,959,1005	3 x hexose	1005.5337	8.1	731408 ± 68292	700227 ± 1622	678699 ± 28074	647943 ± 3678	615053 ± 23641	618340 ± 7003	584503 ± 9686
473,635,797,959,1005	3 x hexose	1005.5337	8.4	248947 ± 33719	309143 ± 5811	328133 ± 1430	235073 ± 10365	213887 ± 21875	256559 ± 16436	209736 ± 7098

Tab. S7 Identification and relative quantification of total and free sapogenins in *B. vulgaris*. EIC, extract ion chromatogram

Sapogenins	EIC±0.2	Retention time (min)	PWT	p35S::LUP5-2	pLUP5::LUP5-4	GWT	CYP72A552::RNAi-13
Total lupeol	409.3851	31.1	N/A	N/A	N/A	N/A	N/A
Free lupeol	409.3851	31.1	N/A	N/A	N/A	N/A	N/A
Total β-amyrin	409.3851	32.5	N/A	N/A	N/A	N/A	N/A
Free β-amyrin	409.3851	32.5	N/A	N/A	N/A	N/A	N/A
Total oleanolic acid	439.3593	21.5	20884 ± 3963	45589 ± 6970	53763 ± 5506	837805 ± 94973	930214 ± 259536
Free oleanolic acid	439.3593	21.5	N/A	N/A	N/A	7694 ± 2187	12102 ± 3978
Total hederagenin	437.3425	16.7	37387 ± 13280	119157 ± 28689	316704 ± 75506	2923104 ± 69572	1763484 ± 124066
Free hederagenin	437.3425	16.7	N/A	N/A	N/A	N/A	N/A

Tab. S8 Non significantly peak area changed saponins in LUP5 transformed G-type *B. vulgaris* compared to wildtype G-type. EIC, extract ion chromatogram; GWT, wildtype G-type *B. vulgaris*; p35S::LUP5, p35S::LUP5 transformed G-type *B. vulgaris*. Hederagenin cellbioside (EIC: m/z 841.4591± 0.2, retention time 12.3 min) was quantified by after diluting sample 60 times

Fragments	Sugar moieties	EIC ± 0.2	Retention time (min)	Peak area						
				GWT	p35S::LUP5-3	p35S::LUP5-4	p35S::LUP5-8	pLUP5::LUP5-1	pLUP5::LUP5-4	pLUP5::LUP5-19
457,619,781, 943,989	3 x hexose	989.5404	8.6 min	20482 ± 2118	38987 ± 1098	51472 ± 15255	21690 ± 3768	19634 ± 865	20222 ± 1111	31016 ± 5656
457,619,781, 943,989	3 x hexose	989.5404	8.9 min	14821 ± 1198	18031 ± 2051	17787 ± 1628	12320 ± 1638	18273 ± 2536	18420 ± 1669	21301 ± 3786
457,619,781, 943,989	3 x hexose	989.5404	9.1 min	38031 ± 3104	52061 ± 5462	65168 ± 12087	35694 ± 2021	48842 ± 4730	49031 ± 4808	51257 ± 2596
447,609,771, 917,1079,1225	3 x hexose 2 x methylpentose	1225.3480	5.3 min	13663 ± 2246	18448 ± 3368	22842 ± 4336	15106 ± 3645	25363 ± 6389	25538 ± 6666	20553 ± 5429
447,609,755, 901	1 x hexose 2 x methylpentose	901.2429	7.0 min	120237 ± 16382	234354 ± 60723	265639 ± 33723	153354 ± 46292	103209 ± 18904	158676 ± 26195	69202 ± 27745
649,811,973, 1019	2 x hexose	1019.5076	7.9 min	21216 ± 6337	46666 ± 8855	55221 ± 9747	32606 ± 7696	24344 ± 7109	19506 ± 5435	24221 ± 3354
649,811,857	1 x hexose	857.4543	8.4 min	12711 ± 5812	54358 ± 10282	39625 ± 11515	11513 ± 1677	8841 ± 1206	6367 ± 1110	12549 ± 4138
469,631,793, 839.4409	2 x hexose	839.4409	13.1 min	184062 ± 9520	177262 ± 6220	188763 ± 36759	162025 ± 11433	163858 ± 9362	158627 ± 16412	169893 ± 1543
473,635,797, 959,1005	3 x hexose	1005.5337	8.4 min	26331 ± 1839	30415 ± 1107	32803 ± 6153	18898 ± 1143	22765 ± 1235	23008 ± 3232	29006 ± 6414
473,635,797, 843.4798	2 x hexose	843.4798	8.0 min	17071 ± 5223	10868 ± 3119	14373 ± 6021	10972 ± 1123	21960 ± 1914	25028 ± 5756	11849 ± 1801
471,633,795, 841	2 x hexose	841.4591	12.3 min	100456 ± 5418	106334 ± 12545	110089 ± 8301	95604 ± 2647	107654 ± 11391	102554 ± 10010	96034 ± 5295
471,633,795, 841	2 x hexose	841.4591	10.7 min	33658 ± 5344	37416 ± 8271	38169 ± 6400	31962 ± 9915	47190 ± 3606	45222 ± 2232	27295 ± 7916
471,633,795, 841	2 x hexose	841.4591	11.5 min	25386 ± 3902	24225 ± 3487	29513 ± 3269	18135 ± 1956	32274 ± 3485	28554 ± 3328	22019 ± 5539
471,633,795, 841	2 x hexose	841.4591	12.1 min	47036 ± 16885	32613 ± 17910	29637 ± 9302	19750 ± 14769	47282 ± 7261	45142 ± 2657	25284 ± 6002
471,633,795, 841	2 x hexose	841.4591	13.3 min	88619 ± 2761	90928 ± 7073	95672 ± 18388	90196 ± 5713	86544 ± 4723	83258 ± 8292	92525 ± 5391
471,633,795	2 x hexose	795.4536	9.5 min	5117 ± 699	6368 ± 987	6909 ± 2907	6047 ± 640	9009 ± 1110	7197 ± 861	8855 ± 1485
471,633,795	2 x hexose	795.4536	11.7 min	218249 ± 23129	216482 ± 35533	230016 ± 16668	210209 ± 23914	309077 ± 21009	301037 ± 23688	200395 ± 82158
471,603,749, 795	1 x pentos 1 x methylpentose	795.4536	13.8 min	17769 ± 2552	22284 ± 3550	65568 ± 57745	19566 ± 2099	25555 ± 1657	26195 ± 1627	18956 ± 4763
471,633,679	1 x hexose	679.4063	13.2 min	212610 ± 36755	233137 ± 4815	207951 ± 34240	207624 ± 16534	167168 ± 11718	179878 ± 11956	208232 ± 22129
471,633,795, 957,1003	3 x hexose	1003.5115	9.3 min	11634 ± 2468	11052 ± 1548	10676 ± 2002	20561 ± 6247	18977 ± 2852	16839 ± 3467	23814 ± 3476
455,617	1 x hexose	617.4059	15.9 min	10294 ± 1439	13237 ± 2571	13947 ± 1427	15815 ± 829	17539 ± 283	17940 ± 1969	12755 ± 3767
455,617,779, 941	3 x hexose	941.5115	10.2 min	2146 ± 144	2830 ± 851	3396 ± 249	2760 ± 899	3733 ± 575	3265 ± 921	2592 ± 375
455,617,779	2 x hexose	779.4587	11.5 min	5830 ± 286	7906 ± 539	5088 ± 2615	7406 ± 635	5930 ± 366	8004 ± 812	6783 ± 1732
455,617,779	2 x hexose	779.4587	14.2 min	2662 ± 1048	4302 ± 1240	5675 ± 814	2779 ± 669	5970 ± 1384	5353 ± 735	3217 ± 1321
455,617,779, 825	2 x hexose	825.4642	11.3 min	9091 ± 1213	7393 ± 793	10092 ± 3281	9520 ± 1552	10141 ± 1139	9067 ± 2291	9807 ± 1897
455,617,779, 825	2 x hexose	825.4642	13,3 min	71245 ± 20814	92414 ± 31302	113829 ± 25789	98847 ± 34763	176841 ± 21237	178891 ± 7069	89426 ± 54408
455,617,779, 825	2 x hexose	825.4642	14.5 min	243257 ± 17534	274932 ± 25755	308652 ± 55083	298511 ± 15815	308259 ± 6512	499973 ± 154132	303445 ± 31768

Tab. S9 Non significantly peak area changed saponins in CYP72A552 RNAi G-type *B. vulgaris* compared to wildtype G-type. EIC, extract ion chromatogram; GWT, wildtype P-type *B. vulgaris*; CYP72A552::RNAi, CYP72A552 silenced G-type *B. vulgaris.*

Fragments	Sugar moieties	EIC ± 0.2	Retention time (min)	Peak area			
				GWT	CYP72A552::RNAi -12	CYP72A552::RNAi -13	CYP72A552::RNAi -20
457,619,781,943,989	3 x hexose	989.5404	8.6	59779 ± 11070	22819 ± 12186	50571 ± 16040	114054 ± 20836
457,619,781,943,989	3 x hexose	989.5404	8.9	138182 ± 17852	73616 ± 40884	128750 ± 25492	191028 ± 39924
457,619,781,943,989	3 x hexose	989.5404	9.1	114945 ± 21459	58662 ± 32186	103841 ± 16971	193364 ± 49342
447,609,771,917,1079,1225	3 x hexose 2 x methylpentose	1225.3480	5.3	46154 ± 7366	21556 ± 14434	71688 ± 59579	113983 ± 69871
649,811,973,1019	2 x hexose	1019.5076	7.9	135335 ± 19808	55449 ± 29797	88059 ± 16763	65890 ± 20234
649,811,857	1 x hexose	857.4543	8.4	77952 ± 10900	25061 ± 12578	50488 ± 8478	49935 ± 17760
473,635,797,959,1005	3 x hexose	1005.5337	8.4	113195 ± 20812	29081 ± 13620	56670 ± 9168	145066 ± 50308
473,635,797,843.4798	2 x hexose	843.4798	8.0	31446 ± 7784	12140 ± 5535	18386 ± 2755	19709 ± 5620
471,633,795,841	2 x hexose	841.4591	10.7	136220 ± 38770	76534 ± 45911	99014 ± 25908	82659 ± 26899
471,633,795,841	2 x hexose	841.4591	11.5	35229 ± 5872	11376 ± 3350	18245 ± 5585	25617 ± 5052
471,633,795	2 x hexose	795.4536	9.5	90519 ± 8592	29836 ± 11870	58144 ± 15541	66536 ± 14249
471,633,795	2 x hexose	795.4536	11.7	728536 ± 151774	470860 ± 306512	665595 ± 140211	582805 ± 161723
471,603,749,795	1 x pentos 1 x methylpentose	795.4536	13.8	550036 ± 42464	380767 ± 129601	494529 ± 57332	285280 ± 150650
471,633,795,957,1003	3 x hexose	1003.5115	8.5	60301 ± 7106	41708 ± 24516	60618 ± 18649	36844 ± 8278
471,633,795,957,1003	3 x hexose	1003.5115	9.3	47445 ± 9326	13747 ± 6648	32703 ± 11770	87323 ± 26478
455,617	1 x hexose	617.4059	15.9	143159 ± 29068	132769 ± 98284	175934 ± 61931	52051 ± 22730
455,617,779,941	3 x hexose	941.5115	10.2	22899 ± 4464	14295 ± 7815	25146 ± 5922	23357 ± 5447
455,617,779,941,987	3 x hexose	987.5170	10.7	21236 ± 2477	16389 ± 7742	27686 ± 5895	38374 ± 8331
455,617,779	2 x hexose	779.4587	11.0	99602 ± 24999	62016 ± 54628	68706 ± 46239	86853 ± 28604
455,617,779	2 x hexose	779.4587	11.5	71824 ± 16651	52094 ± 28782	68902 ± 18317	75586 ± 28924
455,617,779	2 x hexose	779.4587	14.2	29037 ± 6717	15961 ± 11135	23933 ± 1307	19142 ± 2578
455,617,779,825	2 x hexose	825.4642	11.3	29293 ± 1026	20550 ± 10861	26234 ± 4644	30276 ± 7026
455,617,779,825	2 x hexose	825.4642	13.3	321419 ± 118351	201969 ± 167251	291024 ± 109151	266926 ± 75364
455,617,779,825	2 x hexose	825.4642	14.5	2161276 ± 255225	1722183 ± 777331	2351852 ± 301176	1599796 ± 531805