Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

Molecular Plant
Research Article

CellPress
PARTNER JOURNAL

CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis

Elisa Biazzi[1,5], Maria Carelli[1,5], Aldo Tava[1], Pamela Abbruscato[2], Ilaria Losini[2], Pinarosa Avato[3], Carla Scotti[1,*] and Ornella Calderini[4]

[1]Consiglio per la ricerca in agricoltura e l'analisi dell'economia agraria (CRA), Centro di Ricerche per le Produzioni Foraggere e Lattiero Casearie, 26900 Lodi, Italy

[2]Parco Tecnologico Padano, 26900 Lodi, Italy

[3]Dipartimento di Farmacia-Scienze del Farmaco, Università degli Studi di Bari Aldo Moro, 70125 Bari, Italy

[4]Consiglio Nazionale delle Ricerche (CNR), Istituto di Bioscienze e Biorisorse, 06128 Perugia, Italy

[5]These authors contributed equally to this article.

***Correspondence: Carla Scotti (carla.scotti@entecra.it)**

http://dx.doi.org/10.1016/j.molp.2015.06.003

ABSTRACT

In the *Medicago* genus, triterpenic saponins are bioactive secondary metabolites constitutively synthesized in the aerial and subterranean parts of plants via the isoprenoid pathway. Exploitation of saponins as pharmaceutics, agrochemicals and in the food and cosmetic industries has raised interest in identifying the enzymes involved in their synthesis. We have identified a cytochrome P450 (CYP72A67) involved in hemolytic sapogenin biosynthesis by a reverse genetic TILLING approach in a *Medicago truncatula* ethylmethanesulfonate (EMS) mutagenized collection. Genetic and biochemical analyses, mutant complementation, and expression of the gene in a microsome yeast system showed that CYP72A67 is responsible for hydroxylation at the C-2 position downstream of oleanolic acid synthesis. The affinity of CYP72A67 for substrates with different substitutions at multiple carbon positions was investigated in the same *in vitro* yeast system, and in relation to two other CYP450s (CYP72A68) responsible for the production of medicagenic acid, the main sapogenin in *M. truncatula* leaves and roots. Full sib mutant and wild-type plants were compared for their sapogenin profile, expression patterns of the genes involved in sapogenin synthesis, and response to inoculation with *Sinorhizobium meliloti*. The results obtained allowed us to revise the hemolytic sapogenin pathway in *M. truncatula* and contribute to highlighting the tissue specificities (leaves/roots) of sapogenin synthesis.

Key words: *Medicago truncatula*, saponin pathway, cytochrome P450, CYP72A67, CYP72A68, TILLING analysis

Biazzi E., Carelli M., Tava A., Abbruscato P., Losini I., Avato P., Scotti C., and Calderini O. (2015). CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis. Mol. Plant. ■ ■, 1–14.

INTRODUCTION

Triterpenic saponins are secondary metabolites constitutively synthesized in the aerial (leaves, stems, flowers, pods) and subterranean parts (roots, nodules) of the plants in the *Medicago* genus. Moreover, saponin synthesis is induced in response to herbivore insect damage and abiotic stresses (Agrell et al., 2004) and elicited by methyl jasmonate (MJ), a signal component in plant defense responses (Suzuki et al., 2005). Triterpenic saponins are synthesized via the isoprenoid pathway through which important plant hormones (sterols, cytokinins, gibberellins, abscissic acid) are also formed (Naoumkina et al., 2010). Cyclization of 2,3-oxidosqualene into β-amyrin is responsible for the synthesis of the common skeleton of all the sapogenins (aglycone moieties of saponins) reported in the genus. The β-amyrin scaffold is then transformed by means of oxidative modifications at particular positions of the A, D, and E rings mediated by cytochrome P450 (CYP450) oxygenases. Finally, sugar chains are attached in precise positions on the A and E rings by several glycosyltransferases (GTs) to form mono- and bidesmoside saponins.

Hemolytic saponins, showing hemolytic activity due to their affinity for membrane sterols, and non-hemolytic saponins are both

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and IPPE, SIBS, CAS.

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

present in *Medicago* genus and are synthesized through different pathways branching immediately downstream of β-amyrin formation (Tava and Avato, 2006; Tava et al., 2011b). The biosynthetic independence of the two pathways has been demonstrated by *lha* (lacking hemolytic activity) mutants in barrel medic (*Medicago truncatula* Gaertn.) (Carelli et al., 2011). In fact, knockdown of cytochrome CYP716A12, catalyzing the first step of the hemolytic sapogenin pathway, blocked the synthesis of all hemolytic sapogenins while soyasapogenol (non-hemolytic) biosynthesis persisted. Several CYP450s involved in sapogenin biosynthesis in *Medicago* genus have been identified and their role in the pathway has been characterized both *in planta* (Carelli et al., 2011) and in *in vivo* heterologous systems (e.g. yeast) (Fukushima et al., 2011, 2013; Seki et al., 2011). Some of these CYP450s are present in multiple copies sharing high sequence homology and a conserved exon-intron structure in the *Medicago truncatula* genome. GTs involved in sapogenin glycosylation in *Medicago* genus are likely to represent a larger class than CYP450s due to the variety of sugar units and the different positions of attachment of the sugar moieties. However, only a few of them have been identified and characterized up to now (Achnine et al., 2005; Naoumkina et al., 2010). In addition, a post-translational negative control triggered by the ectopic accumulation of bioactive saponins has been reported in *M. truncatula* based on the RING membrane anchor E3 ubiquitin ligase (MKB1), which targets the activity of 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR), the rate-limiting enzyme in the isoprenoid pathway (Pollier et al., 2013). The increasing knowledge on saponin biosynthesis and regulation suggests that the saponin biosynthetic pathway is highly integrated in plant developmental machinery and its maintenance and correct working are subject to complex control systems. The identification and characterization of the entire set of CYP450s involved in sapogenin biosynthesis is therefore of great importance not only as a prerequisite for their heterologous expression and manipulation but also as a tool to thoroughly investigate the range of biological roles of saponins in *Medicago* genus.

Here, we report the identification of a CYP450 gene involved in the biosynthetic pathway of hemolytic sapogenins by a reverse genetic TILLING approach (McCallum et al., 2000) in a *M. truncatula* ethylmethanesulfonate (EMS) mutagenized collection (Porceddu et al., 2008; Carelli et al., 2013) and its functional characterization by expression in an *in vitro* yeast system, also in relation to two other CYP450s involved in the same pathway. The consequences on the biosynthetic pathway of hemolytic sapogenins in different plant tissues (leaves and roots) are discussed. Mutant and wild-type full sib *M. truncatula* plants are compared for their sapogenin profile, expression patterns of genes involved in sapogenin synthesis and control, and response to inoculation with *Sinorhizobium meliloti* to improve the understanding of the roles of saponins in plant development.

RESULTS

TILLING Analysis of a *M. truncatula* Mutant Collection Leads to the Identification of a Mutant Line with Altered Hemolytic Sapogenin Pattern

The CYP72A67 gene was co-expressed with β-amiryn synthase and induced by MJ (Naoumkina et al., 2010). It was differentially expressed in genome-wide transcript profiling of the *Medicago truncatula lha-1* mutant lacking hemolytic saponins (Carelli et al., 2011) and consistently co-expressed with CYP716A12 and CYP72A68 (Fukushima et al., 2013), both involved in hemolytic sapogenin biosynthesis, in hybrid derivatives between *M. sativa* and *M. arborea* (Carelli et al., 2015).

Seven mutants for the CYP72A67 gene were identified in the TILLING analysis of the EMS mutagenized collection (Porceddu et al., 2008), two of which, Mt1903 and Mt2961, presented amino acid substitution in the coding region (Supplemental Figure 1 online). In the Mt1903 M2 plant, a Pro to Leu (Pro439/Leu) change was identified in the heterozygous state; this substitution was 26 amino acids before the highly conserved heme Cys ligand. In the Mt2961 M2 plant, a mutation in the homozygous state caused an Arg to Trp (Arg434/Trp) change; this substitution was 31 amino acids before the highly conserved heme Cys ligand. The genotyping of the segregating Mt1903 progeny (M4 generation, 69 plants) by means of high-resolution melt (HMR) analysis allowed 22 mutant, 23 heterozygous, and 24 wild-type plants to be identified. All three genotypic classes were analyzed at the pod setting stage; because the sapogenin composition in heterozygous plants was similar to that in wild-type and mutant plants (Supplemental Figure 2 online), this class was not further investigated. Thus, the M5 progeny obtained from the genotyped M4 mutant and wild-type plants was analyzed at the vegetative stage. Leaves and roots from mutant and wild-type plants were collected, pooled, and used for chemical analyses. The chemical composition of the crude saponin mixtures of Mt1903 mutant and wild-type plants was first evaluated by thin-layer chromatography (TLC) (Supplemental Figure 3 online). Compositional differences between mutant and wild-type plants (roots and leaves) were evident, in particular for hemolytic saponins, while soyasaponin B, the major non-hemolytic saponin, was present in both (Supplemental Figure 3 online). Total sapogenins were then evaluated by gas chromatography (GC) after acid hydrolysis of the saponins (Figure 1). In Mt1903 mutant, no bayogenin, medicagenic, or zanhic acid was detected in the leaves and roots; moreover, three additional peaks were found in the mutant (Figure 1A, peaks 7, 9, and 10).

The M4 generation (four plants) of the homozygous line, Mt2961, was grown together with the M4 progeny of a Jemalong 2HA10-9-3 control line that underwent the same treatment as the mutagenized lines but with 0% EMS. The chemical analysis of the leaves and roots showed no differences in saponin and sapogenin content and composition in the Mt2961 mutant line with respect to the control; thus, this line was not further studied.

Identification of the Mutant-Specific Sapogenins

The mass spectra of the three additional compounds detected in the Mt1903 mutant plants (Figure 1B) were consistent with the triterpene pentacyclic structure reported for sapogenins of the *Medicago* genus. In fact, they showed the typical retro Diels-Alder fragmentation pattern observed in all oleanane-type compounds possessing a Δ^{12} double bond.

The molecular weight (MW) of compound **7** (analyzed as a methyl-silyl derivative) was *m/z* = 586, consistent with the

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

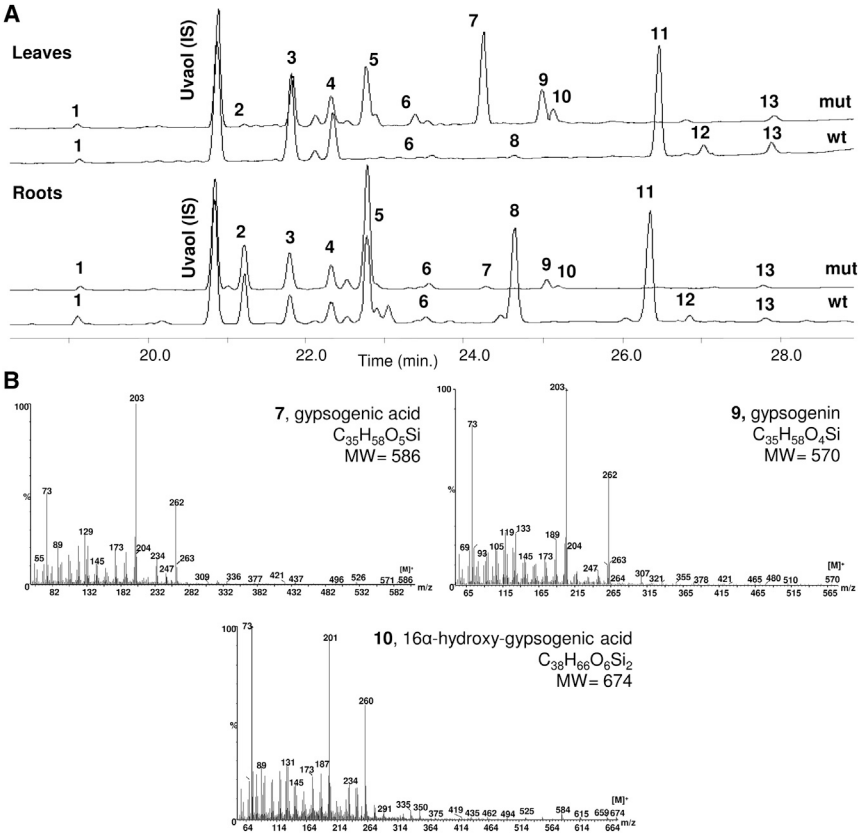

Figure 1. Analysis of Sapogenins in Leaves and Roots of Mt1903 Mutant and Wild-Type Plants.

(A) GC-MS chromatograms of sapogenins in leaves and roots: Mt1903 mutant (mut) and wild-type (wt) plants (M5 generation). 1, soyasapogenol C; **2**, oleanolic acid; **3**, soyasapogenol D; **4**, soyasapogenol F; **5**, hederagenin; **6**, soyasapogenol B; **7**, gypsogenic acid; **8**, bayogenin; **9**, gypsogenin; **10**, 16α-hydroxy gypsogenic acid; **11**, medicagenic acid; **12**, zanhic acid; **13**, soyasapogenol A. Uvaol is used as internal standard (IS).

(B) Mass spectra of methyl-trimethylsilyl derivatives of the mutant-specific sapogenins (**7**, **9**, and **10**).

molecular formula $C_{35}H_{58}O_5Si$, while compound **9** had a MW of *m/z* = 570, consistent with the molecular formula $C_{35}H_{58}O_4Si$ (Figure 1B). Consequently, compounds **7** and **9** were assumed to be gypsogenic acid and its aldehyde precursor gypsogenin, respectively. The MW of compound **10**, *m/z* = 674, matched the molecular formula $C_{38}H_{66}O_6Si_2$ corresponding to 3β,16α-dihydroxyolean-12-en-23,28-dioic acid (Figure 1B). To confirm our hypotheses, gypsogenin, gypsogenic acid, and 3β,16α-dihydroxyolean-12-en-23,28-dioic acid standards were produced by fractionating the aglycone mixture obtained after acid hydrolysis of *Saponaria vaccaria* (Caryophyllaceae) purified saponins. Gypsogenin and gypsogenic acid are in fact among the most abundant aglycones found in this species (Meesapyodsuk et al., 2007); both are derived from β-amyrin and characterized by the presence of a carboxylic group at the C-28 position and an aldehydic (gypsogenin) or a carboxylic (gypsogenic acid) group at the C-23 position. The aglycone 3β,16α-dihydroxyolean-12-en-23,28-dioic acid, also detected in *S. vaccaria* (Meesapyodsuk et al., 2007) and in *S. officinalis* (Moniuszko-Szajwaj et al., 2013), is characterized by the presence of two carboxylic groups at the C-28 and C-23 positions and a hydroxylic group at the C-16 position. This compound is referred to from now on as 16α-hydroxy gypsogenic acid, the same denomination used in *Saponaria* spp.

Comparison of the behavior of compounds **7**, **9**, and **10** on GC and GC–mass spectrometry (GC-MS) with the reference standards obtained from *S. vaccaria* confirmed their identification.

Interestingly, Fukushima et al. (2013) recently found that yeast strains transformed with two *M. truncatula* cytochromes, CYP716A12 and CYP72A68v2, responsible for carboxylation at the C-28 and C-23 positions of β-amyrin, respectively, produced gypsogenic acid.

Gas Chromatographic Analysis of Mutant and Wild-Type Plants Reveals that the Mt1903 Mutant Carries a CYP72A67 Loss-of-Function Allele

After identification by GC-MS and electrospray ionization-tandem mass spectrometry (ESI-MS/MS), aglycones of mutant and wild-type plants were quantified by GC–flame ionization detection (FID) using the internal standard method. In the leaves of Mt1903 mutants, hemolytic sapogenins significantly differed from wild-type plants both in their total content and composition (Figure 2 and Supplemental Table 1 online); in particular, bayogenin, medicagenic acid, and zanhic acid were completely absent, while oleanolic acid and hederagenin showed a significant increase. In addition, the mutant-specific sapogenins gypsogenic acid and gypsogenin were the most abundant aglycones, accounting for 53% and 11%, respectively, of the leaf hemolytic sapogenins. The leaf content of non-hemolytic sapogenins also decreased significantly in Mt1903 mutants compared with wild-type plants. In roots, no differences in both hemolytic and non-hemolytic sapogenin content were found between mutant and wild-type plants. However, the composition of the hemolytic fraction differed remarkably in mutant and wild-type plants because of the absence of bayogenin, medicagenic acid, and zanhic acid and a significant increase in oleanolic acid (21% and 10% of the root hemolytic sapogenins in mutant and wild-type, respectively) and hederagenin content (74% and 23% of the root hemolytic sapogenins in mutant and wild-type, respectively) in mutant roots. Notably, the content of the mutant-specific sapogenins gypsogenic acid and gypsogenin in roots was lower than in leaves, accounting for only 1% and 3% of the root hemolytic sapogenins, respectively. 16α-Hydroxy gypsogenic acid represented a minor aglycone, in particular in roots, where it accounted for <1% of the hemolytic sapogenins.

In conclusion, all the hemolytic sapogenins carrying a hydroxylic group at the C-2 position (bayogenin, medicagenic acid, and

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

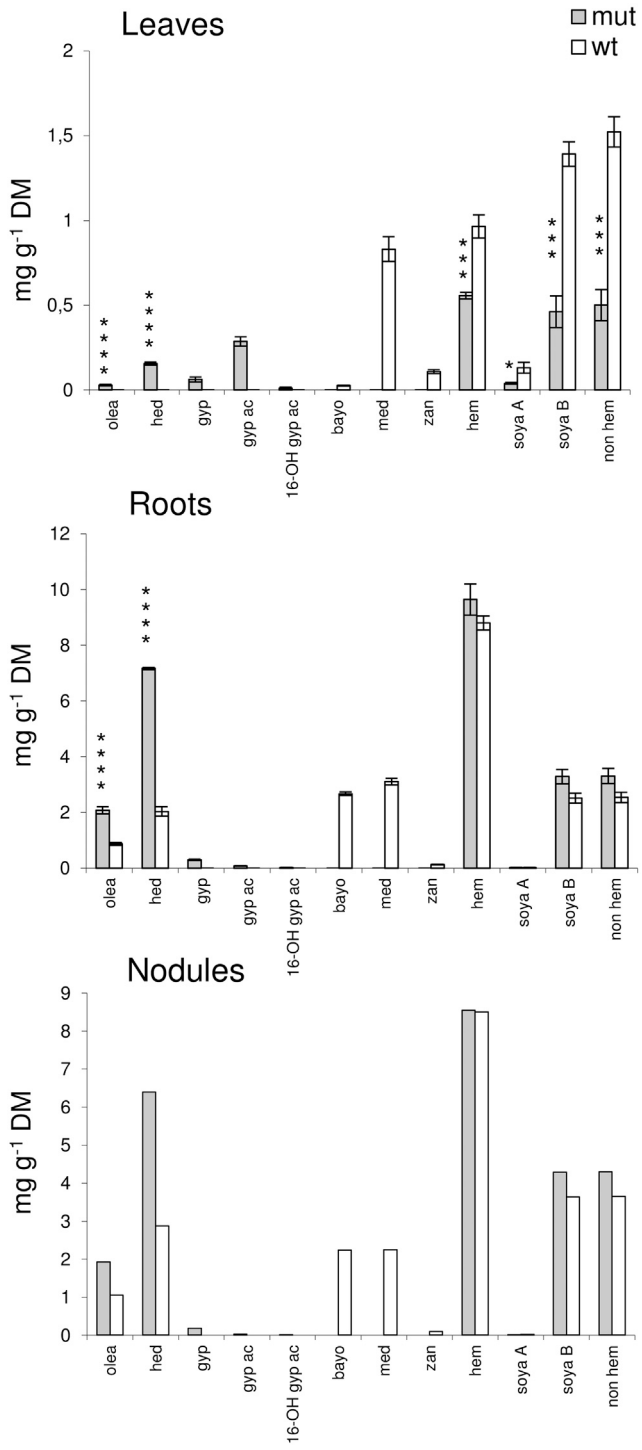

Figure 2. Analysis of Sapogenin Content in Leaves, Roots, and Nodules of Mt1903 Mutant and Wild-Type Plants.

Sapogenin content obtained by GC-FID analysis in leaves, roots, and nodules of Mt1903 M5 generation; values for mutant (mut, gray bars) and wild-type (wt, white bars) are means ± SE of three plants. Identification of sayasapogenol B was achieved considering all the artifact compounds detected (soyasapogenols C, D, and F). *$P < 0.05$, ***$P < 0.005$, ****$P < 0.001$ with the F-test. DM, dry matter; olea, oleanolic acid; hed, hederagenin; gyp, gypsogenin; gyp ac, gypsogenic acid; 16-OH gyp ac, 16α-hydroxy gypsogenic acid; bayo, bayogenin; med, medicagenic acid; zan, zanhic acid; hem, total hemolytic sapogenins; soya A, soyasapogenol A; soya B, soyasapogenol B; non hem, total non-hemolytic sapogenins.

zanhic acid) were absent both in leaves and in roots of Mt1903 mutant plants; consequently, we assume CYP72A67 mainly catalyzes oxidation at the C-2 position.

Genetic Complementation of the Mt1903 Mutant

Transgenic hairy roots were obtained for the construct pCambia35S-CYP72A67 and pCambia35S in Mt1903 mutant plants. Two independent pools of transgenic hairy root pCambia35S-CYP72A67 and one of pCambia35S were examined for sapogenin content by GC analysis. The pCambia35S-CYP72A67 transformants showed detectable levels of bayogenin and medicagenic acid, confirming that disruption of the CYP72A67 gene was responsible for the lack of these sapogenins in Mt1903 mutant plants (Figure 3).

Functional Expression of CYP72A67 in a Yeast *In Vitro* System Confirms the Gene Is a C-2 Oxidase

To elucidate the function of CYP72A67 in the hemolytic sapogenin pathway, the gene was expressed in a yeast (*Saccharomyces cerevisiae*) system employing microsomes from the GAL-induced yeast strain WAT11 (Pompon et al., 1996) transformed with the pESC-HIS expression vector containing the CYP72A67 coding sequence. The same strain transformed with an empty vector was used as control. WAT11 strain, overexpressing a yeast P450 reductase, was used to optimize electron transfer during catalysis. The lack of bayogenin, medicagenic, and zanhic acid in the mutant Mt1903 suggested that a block in oxidation at the C-2 position was present. Enzymatic activities in microsomes were then tested by supplying oleanolic acid, hederagenin, gypsogenic acid, and gypsogenin. GC-MS analysis of the reaction products (Figure 4) showed the CYP72A67-dependent formation of 2β-hydroxy oleanolic acid from oleanolic acid (Figure 4A), bayogenin from hederagenin (Figure 4B), medicagenic acid from gypsogenic acid (Figure 4C), and 2β,3β-dihydroxyolean-12-en-23-oxo-28-oic acid from gypsogenin (Figure 4D). The retention times and mass spectra of the products matched well with those of the authentic standards. No enzymatic activity on the tested substrates was detected in microsomes from the control strains.

These results confirmed that CYP72A67 mainly catalyzed oxidation at the C-2 position of different intermediates of the hemolytic sapogenin biosynthesis.

Functional Characterization of CYP72A67 with Different Sapogenin Intermediates Reveals Differences in Efficiency Related to the Chemical Structure of the Substrates

Differences in C-2 oxidation efficiency of CYP72A67 with the different substrates supplied were evident from the relative (percentage) comparison of the peak areas of the different substrates and corresponding products as determined by GC analyses (Figure 4). In fact, in the yeast system and under the conditions used, CYP72A67 transformed on average 46.06% of oleanolic acid into 2β-hydroxy oleanolic acid, 34.62% of hederagenin into bayogenin, and 10.22% of gypsogenic acid

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

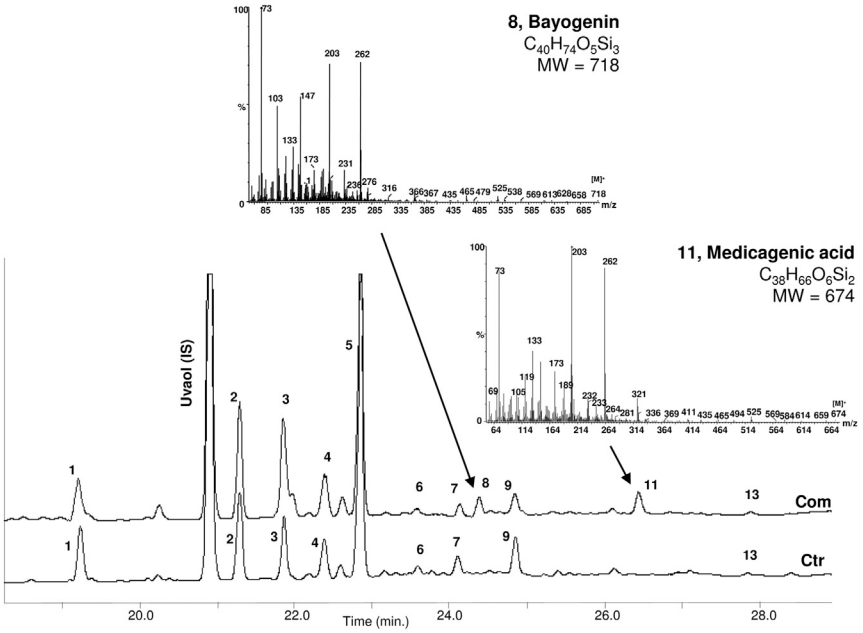

Figure 3. Analysis of Sapogenins in Retro-Transformed Hairy Roots from the Mt1903 Mutant Line.

GC-MS chromatograms of sapogenins from the retro-transformed Mt1903 mutant line: complemented (Com) and control (Ctr) hairy roots. Mass spectra of methyl-trimethylsilyl derivatives of the restored sapogenins (**8**, **11**) are reported. **1**, soyasapogenol C; **2**, oleanolic acid; **3**, soyasapogenol D; **4**, soyasapogenol F; **5**, hederagenin; **6**, soyasapogenol B; **7**, gypsogenic acid; **8**, bayogenin; **9**, gypsogenin; **11**, medicagenic acid; **13**, soyasapogenol A. Uvaol is used as internal standard (IS).

into medicagenic acid (Table 1). All these substrates carried a carboxylic group at the C-28 position; that is, they were all reaction products of CYP716A12-dependent oxidation of β-amyrin, the first step in the hemolytic sapogenin pathway (Carelli et al., 2011). In order to verify this sequential oxidation specificity of CYP72A67, β-amyrin and erythrodiol, a β-amyrin derivative carrying a hydroxylic group at the C-28 position, were used as substrates. No transformation of these compounds was found indicating that CYP72A67 was able to accept and oxidize only the products of CYP716A12. The same was true when α-boswellic acid, similar to gypsogenic acid but lacking the C-28 carboxylic group, was used as substrate (Table 1).

Conversely, many C-28 carboxylated substrates were oxidized by CYP72A67 in the *in vitro* yeast system although with different efficiencies; in particular, the addition of a C-16 hydroxylic group decreased the enzymatic activity of CYP72A67 as indicated by comparison of the relative conversion rates of echinocystic acid versus oleanolic acid and caulophyllogenin versus hederagenin (Table 1). Considering substitution at C-23, mediated by CYP72A68, the presence of a carboxylic group (highest oxidation degree) reduced CYP72A67 enzymatic efficiency as shown by the relative conversion values found for gypsogenic acid with respect to gypsogenin and hederagenin (Table 1). The absence of the geminal methyl groups at C-20 in ursolic acid also decreased CYP72A67 efficiency compared with oleanolic acid (Table 1), suggesting that the entire E-ring was involved in the substrate–enzyme relationship.

Functional Characterization of CYP72A68 Shows Differences in Efficiency for Substrates Subjected or Not Subjected to CYP72A67-Dependent C-2 Hydroxylation

Differences in the efficiency of the C-23 multifunctional oxidase CYP72A68 for substrates carrying or not carrying the hydroxylic group in the C-2 position mediated by CYP72A67 are likely to influence the biosynthetic steps through which hemolytic sapogenins downstream of oleanolic acid (hederagenin, bayogenin, medicagenic, and zanhic acid) are formed. To test these hypotheses, two CYP72A68 genes, CYP72A68-430 and CYP72A68-470 (Medtr2g055430 and Metr2g055470, respectively), showing 97% identity in nucleotide sequence and 96% identity in amino acid sequence, were separately expressed in the same yeast system used for CYP72A67. Both sapogenins lacking C-2 oxidation (oleanolic acid, hederagenin) and the corresponding aglycones carrying the C-2 hydroxylic group (2β-hydroxy oleanolic acid and bayogenin) were supplied as substrates. CYP72A68-470 showed a higher efficiency, mediated on the four substrates, compared with CYP72A68-430; in addition, the C-2 hydroxylated substrates 2β-hydroxy oleanolic acid and bayogenin were more effectively transformed by both cytochromes compared with oleanolic acid (Table 2). Notably, in the presence of the C-2 hydroxylated substrates, both CYP72A68s produced medicagenic acid by means of three oxidative steps in the case of 2β-hydroxy oleanolic acid or two oxidative steps in the case of bayogenin. On the contrary, both CYP72A68s performed a single oxidative step when supplied with oleanolic acid (Table 2).

Quantitative RT–PCR Analysis of the Genes Involved in the Sapogenin Pathway Reveals Some Differential Gene Expression between Leaves and Roots of Mutant and Wild-Type Plants

In order to investigate the effect of the change in sapogenin/saponin composition on the expression of genes directly involved in the pathway or implicated in its control, quantitative real-time PCR analyses were performed on leaves and roots of the same M5 mutant and wild-type plants at the vegetative stage, analyzed for sapogenin content. In particular, the BAS1 gene, responsible for the cyclization of 2,3-oxidosqualene to β-amyrin (Suzuki et al., 2002) and the CYP450 genes identified in the *M. truncatula* sapogenin pathway were examined. With regard to the hemolytic sapogenin pathway, we studied CYP716A12, catalyzing the conversion of β-amyrin to oleanolic acid, the first biosynthetic step of the pathway (Carelli et al., 2011); two CYP72A68 genes, one of which is similar to CYP72A68v2 oxidizing oleanolic acid at the C-23 position in yeast (Fukushima et al., 2013); and CYP72A67, which we propose to be responsible for C-2 oxidation of CYP716A12 products. For the non-hemolytic sapogenin pathway, we investigated

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

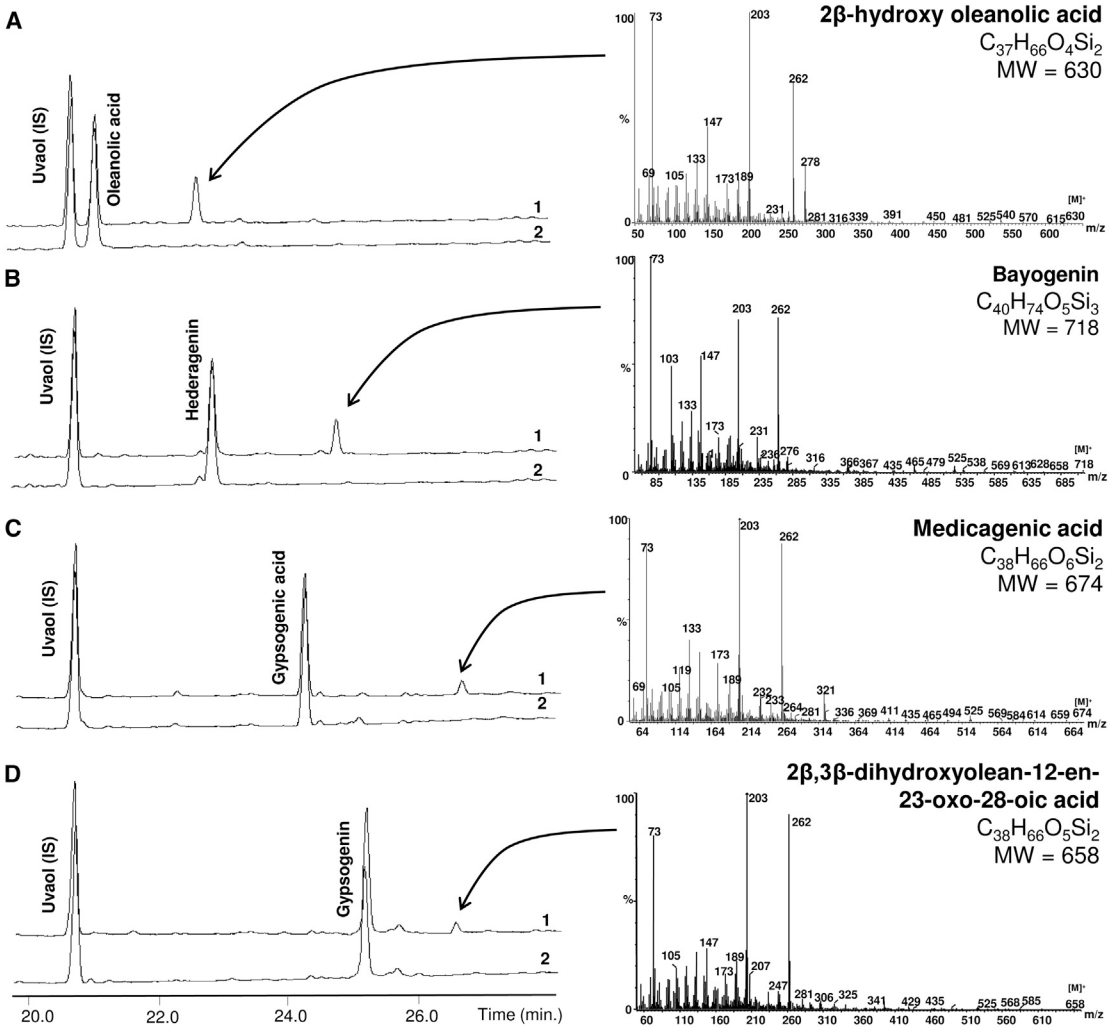

Figure 4. *In Vitro* Oxidation of Different Substrates by CYP72A67 in Microsomes of the Yeast Strain WAT11.
GC-MS analyses of the reaction products resulting from *in vitro* assays containing oleanolic acid **(A)**, hederagenin **(B)**, gypsogenic acid **(C)**, and gypsogenin **(D)** as substrates on strains expressing CYP72A67 (1) and control (2). Mass spectra of methyl-trimethylsilyl derivatives of the additional (product) peaks indicated by the black arrows in **(A)**, **(B)**, **(C)**, and **(D)** are shown. Uvaol is used as internal standard (IS).

CYP93E2, involved in C-24 oxidation of the β-amyrin nucleus (Fukushima et al., 2011), and CYP72A61, catalyzing the C-22 oxidation of the 24-OH-β-amyrin in yeast (Fukushima et al., 2013), both required for soyasapogenol B synthesis. In addition, the RING membrane anchor E3 ubiquitin ligase Makibishi 1 gene (MKB1), the first regulatory factor proposed to control the activity of HMGR, a very early gene of the terpene biosynthetic pathway (Pollier et al., 2013), was also analyzed.

Gene expression levels were calculated using the comparative Ct method with the wild-type as reference to highlight differential gene expression in mutant compared with wild-type in the two tissues (Figure 5A). Expression values were also calculated with the root as reference to investigate tissue differential gene expression in mutant and wild-type plants (Figure 5B). In general, mutant plants did not show significant differences in expression levels with respect to wild-type plants for the genes involved in the hemolytic or non-hemolytic sapogenin pathway both in leaves and in roots (Figure 5A). MKB1 gene was more expressed in mutant leaves compared with wild-type although not significantly (Figure 5A). When comparing gene expression levels in leaves and roots, CYP72A68-470 was significantly more expressed in leaves in both mutant and wild-type plants (Figure 5B).

These results were confirmed by expression analysis of the same genes performed on seedlings grown under axenic conditions for the nodulation assay (see below) grouped in three batches of 18 seedlings for mutant and wild-type (Supplemental Figure 4 online).

The Nodulation Assay Indicates Differences in Nodule Distribution between Mutant and Wild-Type Plants

The biological activity of root saponins could be in relation to the control of nodulation by the symbiotic microorganism *S. meliloti* (Delis et al., 2011). To estimate the biological activity of the root saponins in mutant and wild-type Mt1903 plants, a test with the

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

C-23	C-28			
	None	–OH	–COOH	–COOH and C16-αOH
None	β-Amyrin	Erythrodiol	Oleanolic acid[a]	Echinocystic acid[f]
	0	0	46.06 ± 5.80	3.72 ± 1.63
			Ursolic acid[b]	
			18.79 ± 5.11	
–OH			Hederagenin[c]	Caulophyllogenin[g]
			34.62 ± 3.75	1.76 ± 1.14
–CHO			Gypsogenin[d]	
			32.20 ± 1.75	
–COOH	Boswellic acid		Gypsogenic acid[e]	
	0		10.22 ± 0.96	

Table 1. CYP72A67 Enzymatic Efficiencies in Yeast, Expressed as Product Peak Area versus the Sum of the Substrate and Product Peak Areas (Percentage Values) with Substrates Characterized by Different Substitutions at C-28, C-23, and C16 Positions of the β-Amyrin Nucleus.

Means ± standard errors from one to three independent experiments each in duplicate.

Products of the different substrates are:

[a]2β-Hydroxy oleanolic acid.

[b]2β-Hydroxy ursolic acid.

[c]Bayogenin.

[d]2β,3β-Dihydroxyolean-12-en-23-oxo-28-oic acid.

[e]Medicagenic acid.

[f]2β-Hydroxy echinocystic acid.

[g]2β-Hydroxy caulophyllogenin.

brine shrimp *Artemia salina* L. (Meyer et al., 1982; Tava and Pecetti, 2012) was used. The root saponin mixture of mutant showed higher toxicity, estimated by the median lethal dose (LD_{50}) after 24 h, than that of wild-type plants (96.6 μg/ml, 95% fiducial limits 38.2–132.3 and 274.7 μg/ml, 95% fiducial limits 239.5–310.0, respectively). Mutant and wild-type Mt1903 seedlings were then inoculated with *S. meliloti* strain 1021 under axenic conditions to assay the nodulation pattern. No significant differences were found between mutant and wild-type for any of the traits examined (days from inoculum to detectable nodule formation, total plant fresh weight, root fresh weight, nodule number expressed as the absolute value and relative to the unit of root weight) with the exception of a lower number of nodules per unit of root weight in mutant in the second experiment (Supplemental Table 2 online). To estimate nodule distribution in root, the relationship between absolute nodule number and the number of nodules per unit of root weight (relative nodule number) was used (Figure 6). In wild-type, the seedlings maintained a rather constant number of nodules per unit of root fresh weight for the range of absolute number of nodules observed (Figure 6), indicating a tendency for even nodule distribution along the root. In mutant seedlings, absolute and relative numbers of nodules tended to increase together as indicated by the positive and significant correlation coefficients between the two parameters (Figure 6). As root fresh weight did not significantly differ in mutant and wild-type seedlings (Supplemental Table 2 online), this suggests that nodules in mutant seedlings formed closer to each other with respect to wild-type. A similar behavior was found in the case of adult mutant and wild-type plants grown in the presence of the natural population of *S. meliloti* in soil (Supplemental Figure 5 online).

DISCUSSION

CYP72A67 Functional Characterization Allows a New Hypothesis on the Hemolytic Sapogenin Pathway in *M. truncatula* to Be Proposed

The mutation present in the Mt1903 line caused a similar altered pattern of hemolytic sapogenin in all plant tissues studied (leaves, roots, nodules) and in the different plant developmental stages investigated. In addition, the change observed was consistent over generations (M4−M7) and under different plant growth conditions. This alteration is characterized by the absence of the C-2 hydroxylated sapogenins (bayogenin, medicagenic, and zanhic acids) and the presence of mutant-specific sapogenins (gypsogenin, gypsogenic acid, and 16α-hydroxy gypsogenic acid) not previously reported at a comparable level in the *Medicago* genus. Gypsogenic acid, however, has been detected in a yeast system heterologously expressing the *M. truncatula* CYP716A12 gene, responsible for C-28 carboxylation, and the CYP72A68v2 gene, responsible for C-23 carboxylation (Fukushima et al., 2013). The complementation of mutant hairy roots with a construct containing a full-length coding sequence of CYP72A67 restored the synthesis of sapogenins carrying the hydroxylic group at the C-2 position. We concluded that CYP72A67 catalyzes oxidation at the C-2 position in the *M. truncatula* hemolytic sapogenin pathway.

Gypsogenin and gypsogenic acid have already been reported in the Fabaceae family (*Swartzia* spp, Borel et al., 1987; *Castanospermum* spp, Ahmed et al., 1992) and more recently putatively identified gypsogenin and gypsogenic acid saponins have been detected in *M. truncatula* root border cells (Watson et al., 2015). However, in the Mt1903 line, the high content of the

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

	C2-OH absent	C2-OH present
CYP72A68-430 (Medtr2g055430)	Oleanolic acid[a]	2β-Hydroxy oleanolic acid[b]
	5.09 ± 1.58	17.51 ± 2.94
	Hederagenin[c]	Bayogenin[b]
	4.92 ± 1.33	13.74 ± 1.24
CYP72A68-470 (Medtr2g055470)	Oleanolic acid[a]	2β-Hydroxy oleanolic acid[b]
	14.66 ± 1.86	37.42 ± 7.29
	Hederagenin[c]	Bayogenin[b]
	28.24 ± 3.70	28.25 ± 8.11

Table 2. Enzymatic Efficiencies in Yeast, Expressed as Product Peak Area versus the Sum of the Substrate and Product Peak Areas (Percentage Values) of CYP72A68 Genes with Substrates Characterized by the Absence or Presence of the C-2 Hydroxylic Group.

Means ± standard errors from one to two independent experiments in triplicate (first experiment) or in duplicate (second experiment).
Products of the different substrates are:
[a]Hederagenin.
[b]Medicagenic acid.
[c]Gypsogenic acid.

mutant-specific sapogenins and the absence of the corresponding C-2 hydroxylated aglycones clearly indicated CYP72A67 loss of function was responsible for the synthesis of gypsogenin, gypsogenic acid, and 16α-hydroxy gypsogenic acid.

Sequential specificity of CYP72A67 for the products of CYP716A12-dependent carboxylation at the C-28 position was demonstrated in the *in vitro* yeast system used by the absence of transformation of substrates lacking the carboxylic group at the C-28 position (β-amyrin, erythrodiol, α-boswellic acid) (Table 1). Moreover, no detectable C-2 hydroxylated sapogenin was observed in *M. truncatula lha* mutant plants with disruption in the CYP716A12 gene (Carelli et al., 2011), indicating this sequential specificity was present also *in planta*. Thus, C-2 oxidations mediated by CYP72A67 occur downstream of oleanolic acid synthesis producing in *M. truncatula* C-2 hydroxylated sapogenins (bayogenin, medicagenic acid, zanhic acid) with different C-23 and C-16 substitutions. In the *Medicago* genus, other C-2 hydroxylated sapogenins have been reported: 2β-hydroxy oleanolic acid (methyl group at the C-23 position) in *M. arabica*, *littoralis*, *polymorpha*, *rigidula*, and *tornata* (Tava and Pecetti, 2012); 2β-hydroxy queretaroic acid (methyl group at C-23 and OH substitution at the C-30 position) in *M. arabica* (Tava et al., 2009); 2β,3β-dihydroxyolean-12-en-23-oxo-28-oic acid (CHO substitution at C-23) in *M. arborea* and *hybrida* (Tava et al., 2005; Bialy et al., 2006). Therefore, in the *Medicago* genus, CYP72A67 seems able to oxidize *in planta* a set of derivatives of oleanolic acid mainly differing in the level of oxidation at the C-23 position, from none (methyl group) to the three oxidative steps necessary to introduce a hydroxylic, aldehydic, and carboxylic group in this position. Differences in CYP72A67 efficiency in relation to the degree of substrate oxidation at C-23 and, conversely, differences in CYP72A68 efficiency toward C-2 hydroxylated or not hydroxylated substrates, are likely to have a role in shaping

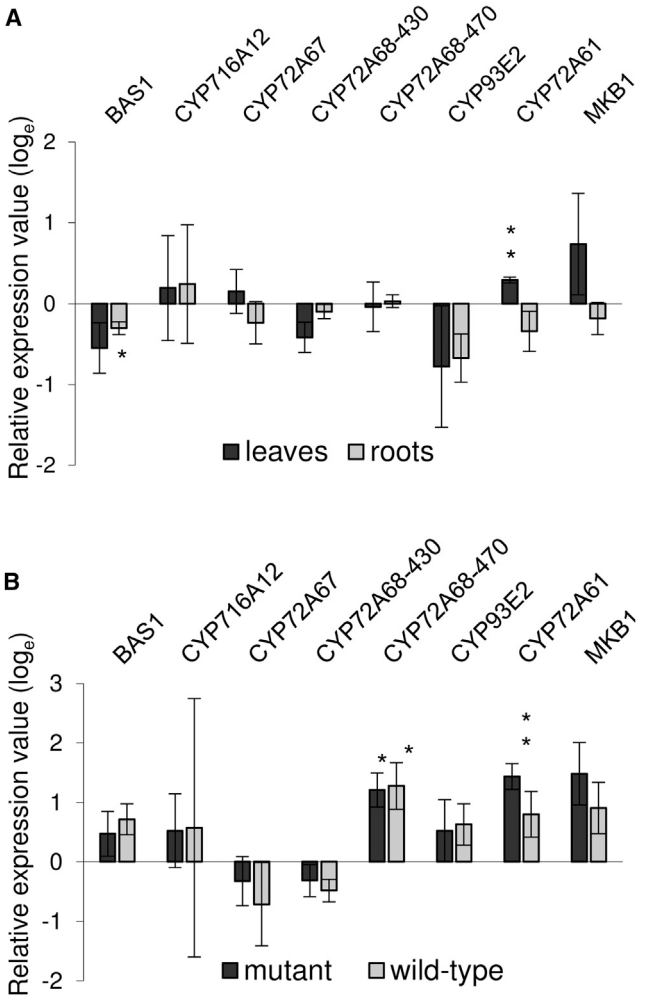

Figure 5. Expression Analysis of BAS1, CYP450 Genes Involved in Hemolytic (CYP716A12, CYP72A67, CYP72A68-430, CYP72A68-470), and Non-hemolytic (CYP92E3, CYP72A61) Sapogenin Synthesis and MKB1 in Mutant and Wild-Type Mt1903 Plants.

Values ($\log_e$ transformation) are calculated by the comparative Ct method using wild-type **(A)** and root **(B)** as reference. Values are means ± SE of three plants. *$P < 0.05$; **$P < 0.01$ in a *t*-test assessing the difference of the averages from 0, representing the null hypothesis (absence of differences between mutant and wild-type **(A)** and between leaves and roots **(B)**).

the main biosynthetic pathways through which hemolytic sapogenins downstream of oleanolic acid are formed.

In the *in vitro* system and under the conditions used, the efficiency of C-2 hydroxylation mediated by CYP72A67 was reduced by the presence of a carboxylic group (highest degree of oxidation) at the C-23 position, while it was highest with the substrate oleanolic acid, carrying no substitution at C-23 (Table 1). On the other hand, C-2 hydroxylation drove the catalytic activity of both CYP72A68 genes to the formation of medicagenic acid via three, in the case of 2β-hydroxy oleanolic acid, or two, in the case of bayogenin, oxidative steps (Table 2). These data suggest that a main pathway for the biosynthesis of medicagenic acid, the chief *M. truncatula* hemolytic sapogenin both in leaves and in roots, is through 2β-hydroxy oleanolic

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

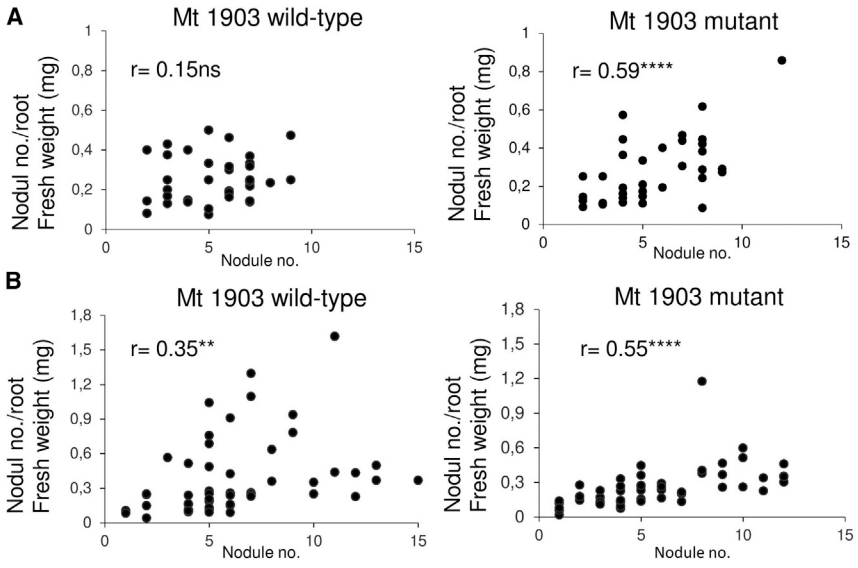

Figure 6. Nodulation Assay in Axenic Conditions in Mt1903 Wild-Type and Mutant Seedlings Inoculated with *S. meliloti* Strain 1021.
Plots of number of nodules versus number of nodules/unit of root fresh weight in two independent experiments **(A and B)**. Significance of the correlation coefficient r: $^{**}P < 0.01$; $^{****}P < 0.001$; ns, not significant.

acid and bayogenin. In this hypothetical sapogenin biosynthetic pathway (Figure 7), partially modifying that proposed by Carelli et al. (2011), the biosynthesis of hederagenin could be the residual result of the hydroxylation activity of CYP72A68 genes on the substrate oleanolic acid. Instead, the main flux of C-2 hydroxylated intermediates, either 2β-hydroxy oleanolic acid or bayogenin, would be catalyzed to medicagenic acid through multi-step oxidation by the same CYP72A68 genes (Figure 7). This hypothesis is different from that proposed by Fukushima et al. (2013); they proposed biosynthesis of medicagenic acid from gypsogenic acid based on the production of this compound by the yeast strain expressing BAS/CPR/CYP716A12 transformed with CYP72A68v2. The Mt1903 mutant line proves that the knockout of CYP72A67 actually leads to the production of gypsogenin and gypsogenic acid in plants too; however, none of these compounds have been detected consistently in *Medicago* genus, while 2β-hydroxy oleanolic acid was reported in some annual medics (Tava and Pecetti, 2012). In addition, in the Mt1903 mutant line, the presence of a compound (gypsogenin) carrying an aldehydic substitution at C-23 never reported in *M. truncatula*, suggests a mismatch between the substrates lacking the C-2 hydroxylic group and the multi-step oxidative function of CYP72A68 genes.

We investigated here for the first time two CYP450s annotated with the same name, CYP72A68-430 and CYP72A68-470, both expressed in plants and both catalyzing the same substrates in the *in vitro* system used though with different efficiencies. Notably, CYP72A68-430 is more similar to CYP72A68v2 of Fukushima et al. (2013) than CYP72A68-470 (99% and 97% identity of coding sequence, respectively) but not identical. This indicates that a variation in sequence and possibly in enzymatic efficiency in CYP450 involved in sapogenin synthesis is present not only among CYP450 orthologs in different genera (Moses et al., 2014b) but also within the *M. truncatula* species and between different forms of the same gene in individual lineages. Due to their physical proximity on chromosome 2, it is possible that CYP72A68-430 and CYP72A68-470 genes represent paralogs whose function has not substantially diverged.

The drop in CYP72A67 efficiency when a C-16 hydroxylic group is added to a substrate (Table 1) suggests that C-16 hydroxylation is the final step in the hemolytic sapogenin pathway. The enzyme responsible for this step in *M. truncatula* is likely to prefer substrates with the highest degree of oxidation (COOH group) at the C-23 position as only 16α-hydroxy gypsogenic acid has been detected in Mt1903 mutant and not quillaic acid, corresponding to gypsogenin with a C-16 hydroxylic group, commonly found in *Saponaria* spp. Interestingly, C-16 oxidase activity in triterpenoid saponin biosynthesis has recently been reported for two CYP450s belonging to different families, CYP87D16 and CYP716Y1 (Moses et al., 2014a, 2015), suggesting the possibility that this activity could have evolved independently and distinctly in different plant species/lineages.

How the Proposed Hemolytic Sapogenin Pathway Matches Hemolytic Sapogenin Synthesis in Leaves

In leaves of Mt1903 mutant and wild-type plants, the most abundant hemolytic sapogenins were those carrying a carboxylic group (maximum degree of oxidation) at the C-23 position: medicagenic and zanhic acid (66%–86% and 31%–11%, respectively, of the leaf hemolytic sapogenin) in wild-type; gypsogenic acid and 16α-hydroxy gypsogenic acid (57%–53% and 12%–2%, respectively, of the leaf hemolytic sapogenin) in mutant plants (Figure 2 and Supplemental Figure 2 online). The synthesis of gypsogenin (17%–11% of the leaf hemolytic sapogenin), with an aldehydic group at the C-23 position, could be due to the reduced affinity of both CYP72A68 genes for substrates lacking C-2 hydroxylation or could be the result of a mechanism controlling gypsogenic acid accumulation. A similar composition of hemolytic sapogenins was found in the leaves of the *M. truncatula* R108 line E113 where medicagenic and zanhic acid accounted for more than 95% of the hemolytic sapogenins in the three phenologic stages examined (Carelli et al., 2011). Thus, sapogenin biosynthesis in *M. truncatula* leaves, after the CYP716A12-dependent oxidation of β-amyrin, is mainly targeted to the production of C-23 carboxylated sapogenins, while the aglycones with the lowest degree of oxidation (hydroxylic group) at the C-23 position (hederagenin, bayogenin) represent minor components of the leaf saponin mixture. CYP72A68-470 gene (Metr2g055470), more efficient than CYP72A68-430 in transforming 2β-hydroxy oleanolic acid and bayogenin into medicagenic acid in the *in vitro* system used, was more expressed in leaves than roots in both mutant and wild-type Mt1903 plants (Figure 5B). This supports the

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

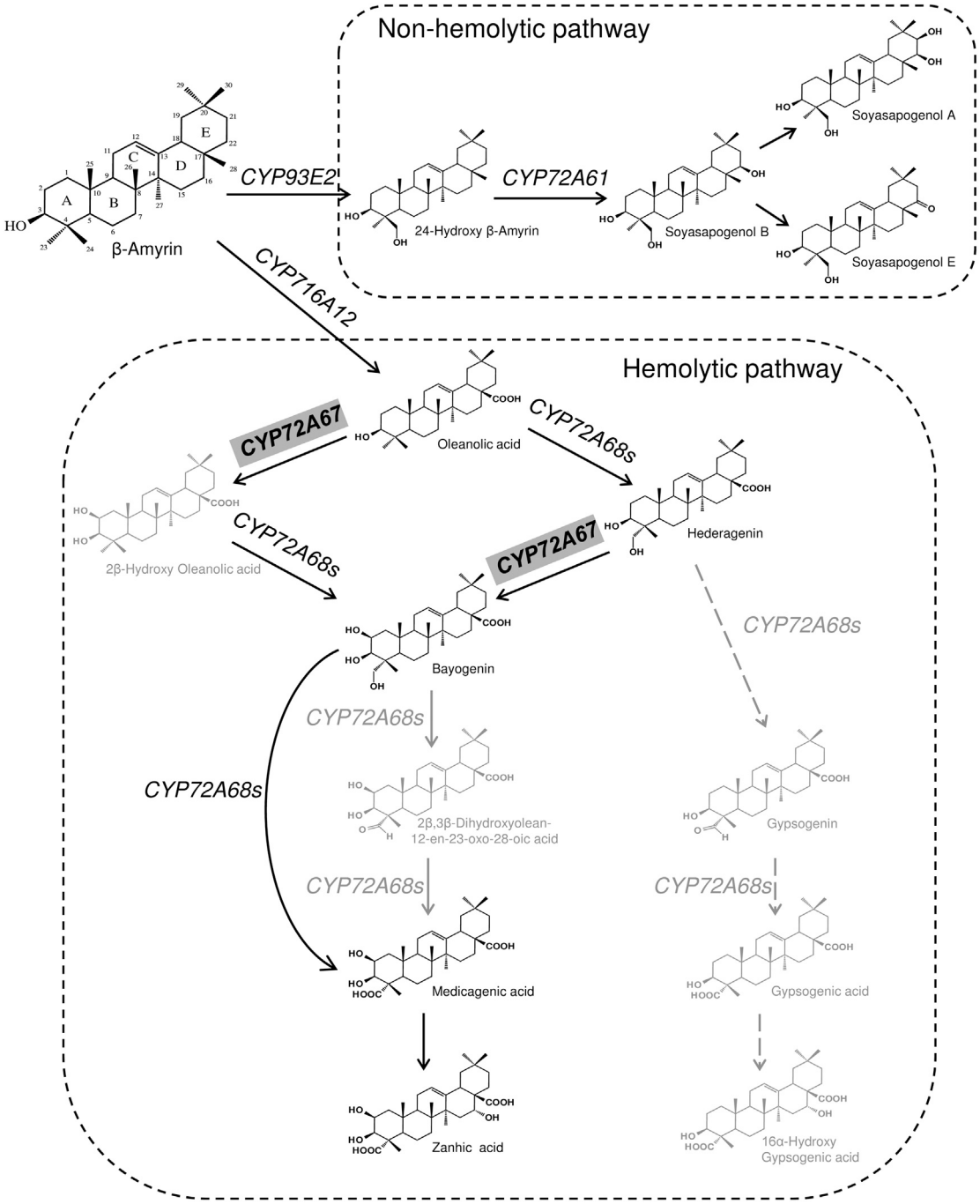

Figure 7. Hypothetical Sapogenin Biosynthetic Pathway in *M. truncatula*.
The aglycones actually detected in wild-type *M. truncatula* plants are indicated in black; aglycones not detected, probably because they are completely transformed, are indicated in gray; broken gray arrows represent the Mt1903 mutant-specific part of the hemolytic sapogenin pathway giving the aglycones lacking a C-2 hydroxylic group. Cytochromes P450 identified to date are reported.

hypothesis that the CYP72A68-470 gene is mainly involved in the constitutive synthesis of medicagenic acid in leaves through the multi-step oxidation of the C-2 hydroxylated substrates 2β-hydroxy oleanolic acid and bayogenin. Despite their physical proximity on chromosome 2 (distance between the two genes of 14 kb), CYP72A68-430 and CYP72A68-470 did not show similar expression patterns in leaves, suggesting they did not act as a paralog gene cluster. On the contrary, consistent expression in different tissues and developmental stages has been found for genes (an oxido-squalene cyclase and two CYP450s) belonging to a putative metabolic cluster involved in triterpene biosynthesis in *Lotus japonicus* (Krokida et al., 2013). This suggests that large diversification is present in the mechanisms of triterpene pathway assembly and pathway regulation in legumes.

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

Considering the key role of medicagenic and zanhic acid in leaf hemolytic sapogenin, it was surprising that their substitution with aglycones lacking C-2 hydroxylation was largely tolerated by mutant plants. In fact, the total hemolytic sapogenin content in mutant leaves was similar (Supplemental Figure 2 online) or lower (Figure 2) compared with wild-type; the ratio between hemolytic and non-hemolytic sapogenins did not differ significantly in mutant and wild-type and the genes involved in both pathways did not show significant and consistent alterations in expression levels (Figure 5A and Supplemental Figure 4 online). Moreover, shoot growth, estimated by stem length, and plant reproduction ability, estimated by pod number and pod fertility, were similar in mutant and wild-type plants (Supplemental Table 2 online). Two hypotheses can be put forward to explain this behavior: either the maintenance of a hemolytic sapogenin pool (C-28 and C-23 carboxylated) in leaves was sufficient to match most of the biological requirements of the plants independently from the presence/absence of the hydroxylic group at the C-2 position or mutant plants displayed a finely tuned resilient response to changes in leaf hemolytic sapogenin composition. The higher expression level of the regulatory gene MKB1 in mutant leaves, where mutant-specific aglycones are concentrated, seems to support the second hypothesis. In the future, a thorough comparison of leaf saponins in mutant and wild-type plants is likely to provide further indications as GTs linking sugar chains at the C-3 hydroxylic group could be affected in substrate recognition by the absence of hydroxylation at the contiguous C-2 position.

Hemolytic Sapogenin Synthesis and Nodulation Pattern in Roots

Differently from leaves, the composition of hemolytic sapogenin in roots was more equilibrated among aglycones with a lower degree of oxidation at the C-23 position (hederagenin and bayogenin) and C-23 carboxylated sapogenins mainly represented in root by medicagenic acid. Interestingly, the two CYP72A68 genes did not showed significant differences in expression levels in roots. In mutants, the content of gypsogenin and gypsogenic acid (<5% of the root hemolytic sapogenins) was far from that of medicagenic acid in wild-type plants and the sapogenin biosynthetic pathway was markedly shifted toward the synthesis of oleanolic acid and hederagenin, accounting for more than 95% of the root hemolytic sapogenin (Figure 2 and Supplemental Figure 2 online). Notably, the biological activity of root saponins, estimated by the test with *Artemia salina*, was higher for mutant than for wild-type plants; this indicates that saponins with high biological activity are required in roots more than in leaves. Medicagenic acid is known for being one of the most biologically active aglycones in the *Medicago* genus (Tava and Avato, 2006); however, its synthesis in Mt1903 mutant is prevented by CYP72A67 loss of function. Thus, sapogenin and saponin synthesis in mutant roots is redirected toward aglycones which ensure a comparable level of biological activity. Mutant and wild-type Mt1903 plants did not significantly and consistently differ for the traits estimating symbiotic aptitude, with the exception of the number of nodules per unit of root weight, which was lower in mutants in two cases (Supplemental Table 2 online). This suggests that the altered root saponin composition in mutants is functional for the maintenance of a nodulation pattern similar to wild-type. According to Kevei et al. (2007), the intracellular domain of the receptor-like kinase NORK, required for Nod factor signaling and nodule development, specifically interacted in epidermal cells of the nodulation-competent root zone with the catalytic domain of HMGR1, an isoform of the rate-limiting enzyme of the mevalonate pathway HMGR. Thus, HMGR1 was identified as a potential downstream element of NORK signaling in legume roots (Kevei et al., 2007). HMGR1 was also one of the HMGR isoforms involved in the control system based on the RING membrane anchor E3 ubiquitin ligase MKB1 managing the production of triterpene saponins (Pollier et al., 2013). Therefore, it is possible to assume that the accumulation of bioactive saponins in the mutant roots could be used to control the nodulation pattern by means of post-translational regulation of HMGR1 mediated by MKB1.

The positive and significant relationship between absolute number of nodules and number of nodules per unit of root weight found in mutants suggests an altered regulation of nodulation is present in mutant compared with wild-type plants. Autoregulation of nodulation (AON) implies a root and shoot signaling cross-talk controlled by leaf receptor kinases whereby early nodulation events inhibit subsequent nodule development (Reid et al., 2011). In the case of the Mt1903 mutant, the altered sapogenin/saponin profile in leaves with the absence of medicagenic and zanhic acids could be implied in the alteration of AON signal transmission, resulting in a different nodule distribution in mutant roots. Further investigations are necessary to confirm these data and to precisely dissect the regulatory cues possibly involved in AON signaling transmission in the leaves of *M. truncatula*.

MATERIALS AND METHODS

Identification of Mutant Lines Using TILLING Collection

The DNA of the M2 generation (2300 plants) of an EMS mutagenized collection obtained in Jemalong genotype 2HA10-9-3 (Porceddu et al., 2008) was screened by the TILLING technique (McCallum et al., 2000) to identify mutants in the CYP72A67 gene using specific primers (Supplemental Table 3 online). The DNA of the mutant plants identified was sequenced to confirm the mutation. The M4 and M5 progenies of the mutant plants identified were used for phenotypic, genetic and chemical characterizations.

Plant Growth Conditions

Seeds were germinated in Petri dishes with mechanical scarification to ensure contemporary germination. Typically, the plant cycle was from January to July; all the viable seedlings were transplanted in plug trays with cells 5-cm diameter at the end of January and grown for about 2 months in a greenhouse (minimum temperature set at 8°C). In the second half of March, plugs were transferred to tube plots 5 cm diameter × 80 cm height with a layer of expanded clay in the bottom and filled with a mixture by volume of 46% soil/46% peat/8% expanded clay sand grade, P- and K-supplied; no mineral N was used. The tube plots were put the outside under a rain shelter structure and provided with individual automatic irrigators. Alternatively, seedlings were directly transplanted in the tube plots in September and grown in the greenhouse until May.

Progeny Genotyping

Progeny genotyping of the heterozygous Mt1903 mutant line (69 plants) was performed by high-resolution melt (HRM) analysis using a Rotor-Gene 6000 (Corbett). DNA samples (50 ng) were amplified in a 10-μl reaction containing 5 μl of SSoFast EvaGreen supermix (Bio-Rad) and

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

specific primers (Supplemental Table 3 online) designed to amplify 120 bp surrounding the polymorphic site. Thermal cycling conditions were 10 min of initial denaturation at 95°C, 50 cycles of denaturation (95°C, 25 s), annealing, and extension (55°C, 30 s). Melting analysis was performed using an HRM-specific channel with the following conditions: denaturation from 65°C to 85°C, ramp of 0.1°C with a hold of 2 s at each step. The results were analyzed with Rotor-Gene 6000 series Software 1.7 and the visual calling of melting curves allowed the progeny individuals to be classified into the three allelic classes. Two samples of each class were sequenced to confirm the correct allelic call.

Saponin/Sapogenin Content Analysis

Saponins were obtained from leaves and roots as reported in Pecetti et al. (2010) and evaluated by TLC (Tava et al., 1993). The reference saponin used as standard in TLC had been previously characterized (Tava et al., 2009).

Total sapogenins were obtained after acid hydrolysis of the ground tissues (leaves, roots, nodules, and hairy roots), identified by GC-MS and quantified by GC-FID using uvaol as internal standard (Tava and Pecetti, 1998; Pecetti et al., 2006). The identities of all sapogenins were confirmed by ESI-MS/MS analysis (Bialy et al., 1999; Tava et al., 2005, 2009, 2011a; Tava and Avato, 2006).

Saponin/Sapogenin Isolation from *Saponaria vaccaria*

Saponaria vaccaria seeds were obtained from Chiltern Seeds (Chiltern Seeds Ltd., Wallingford, UK); plants were grown from September to March in a greenhouse under the same conditions described above for *M. truncatula* mutants but in tube plots 8 cm diameter × 80 cm height. Leaves (180 g) from 30-week-old plants (flowering stage) were ground with liquid nitrogen, dried at 50°C, and defatted with chloroform. Defatted material (150 g) was extracted with MeOH under reflux for 48 h, the solvent was removed under reduced pressure, and the residue was suspended in MeOH 30%. The solution was applied onto a 100 × 60 mm, 40–63 μm Lichroprep RP-18 column (VWR international PBI) preconditioned with MeOH 30%; elution was carried out with MeOH 40%. Total saponins were eluted with MeOH 90% and dried under a vacuum: 3.4 g of crude saponins was obtained (2.2% of defatted material).

Purified saponins (2.0 g) were hydrolyzed in 500 ml of MeOH/H_2O (1:1) 2 N HCl under reflux for 8 h to obtain the corresponding sapogenins. After cooling, MeOH was eliminated under a vacuum, 250 ml of water was added, and aglycones were extracted with ethyl acetate (3 × 200 ml). The organic solution was dried under anhydrous Na_2SO_4 and solvent was removed under a vacuum to give 875 mg of crude sapogenins (43.7% of saponins).

The sapogenin mixture was subjected to silica gel column chromatography eluting with petroleum Et/$CHCl_3$/MeOH (7:2:1). Several fractions were obtained and checked by TLC. Fractions containing the same compounds were combined and used for successive analyses. All fractions were subjected to GC-MS and GC-FID analyses. Pure gypsogenin (45 mg) and gypsogenic acid (104 mg) were analyzed by nuclear magnetic resonance (Tava et al., 2005, 2009) while 16α-hydroxy gypsogenic acid was identified based on GC-MS analysis and the results were compared with those reported in the literature (Jia et al., 1998; Meesapyodsuk et al., 2007; Moniuszko-Szajwaj et al., 2013).

Toxicity Assay with *Artemia salina*

The toxicity of saponin extracts was determined on a population of the brachiopod microcrustacean *Artemia salina* as reported in Tava and Pecetti (2012). Fifteen larvae (at nauplium I stage) of *A. salina* were put in a Petri dish 5-cm in diameter containing artificial sea water (38 g/l sea salt in distilled water) and saponin solutions (total volume of 2 ml). For each saponin extract, five concentrations were tested in triplicate to determine the dose–response relationship. Saponin extracts were dissolved in a 25% EtOH solution; the same solution, without saponins, was used as control. After 24 and 48 h, the mortality of the larvae was checked. The LD_{50} values were estimated using the PROBIT procedure in SAS Software.

Cloning of CYP450 Genes

The full-length CYP72A67 (Medtr2g023680.1), CYP72A68-430 (Metr2g055430), and CYP72A68-470 (Medtr2g055470) coding sequences were amplified with appropriate primers (Supplemental Table 3 online) from cDNA obtained from total leaf RNA of *M. truncatula*. The PCR products were cloned into the pGEM-T Easy Vector (Promega) and sequenced. The correct clones were used for complementation (pGEM-TCYP72A67) and yeast expression (pGEM-TCYP72A67, pGEM-TCYP72A68-430, pGEM-TCYP72A68-470).

Complementation of the Mt1903 Mutant Line

For complementation analysis, the pGEM-TCYP72A67 clone was re-amplified with primers CYP72A67BamHI/SalI fw and rev (Supplemental Table 3 online) and sequenced. The correct clone was digested with BamHI and SalI restriction enzymes to excise the CYP72A67 fragment, which was then ligated to the vector pCambia35S (O. Calderini, unpublished results) cut with the same restriction enzymes to obtain pCambia35S-CYP72A67. The vector pCambia35S, derived from pCAMBIA2300 (Cambia), harbors a 35S cassette from the vector pDHA (Tabe et al., 1995).

The binary vectors pCambia35S-CYP72A67 and pCambia35S were electroporated into Arqua1 *Agrobacterium rhizogenes* strain (kindly provided by Mireille Chabaud, Laboratory of Plant-Microbe Interactions, INRA-CNRS, Auzeville-Tolosan, France). The Arqua1-CYP72A67 strain and Arqua1-pCambia35S empty vector strain (control) were used to generate retro-transformed hairy roots from seedlings of the Mt1903 mutant line as described by Boisson-Dernier et al. (2001). After 28 days of seedling growth on Fabreus modified medium with 25 mg/l kanamycin, the emerging hairy roots were cut and sub-cultured in Murashige & Skoog medium with vitamins (Duchefa) supplemented with 100 mg/l of cefotaxime and 1% saccharose; the roots were cultured in darkness at 24°C with shaking (130 rpm). When hairy roots reached a fresh weight of 2.5 g, they were dried at 50°C for analysis of the sapogenin content as previously described.

Expression of CYP450 Genes in *Saccharomyces cerevisiae*

The CYP72A67, CYP72A68-430, and CYP72A68-470 fragments were excised from the corresponding pGEM-TCYP vectors with the restriction enzymes EcoRI and ClaI and subcloned into the pESC-HIS vector (Agilent Technologies). The expression vector was transformed into yeast (*S. cerevisiae*) strain WAT11 (Pompon et al., 1996) by the lithium acetate procedure (Gietz et al., 1992). For CYP450 gene expression, the recombinant strains were cultured according to a low-density procedure (Pompon et al., 1996) with the addition of 13 mg/l hemin (Sigma-Aldrich) to the medium in the last induction step. WAT11 clones transformed with pESC-HIS empty vector were used as control. Microsome preparation was performed as described by Pompon et al. (1996) except that ultracentrifugation at 100 000 *g* was performed for 60 min.

In Vitro Enzymatic Activity Assay

The activity of the CYP72A67, CYP72A68-430, CYP72A68-470 proteins was tested in a 500-μl reaction mixture consisting of 100 mM potassium phosphate buffer (pH 7.4) containing 20 mM Glc-6-phosphate, 2.5 U of Glc-6-phosphate dehydrogenase, 50 μg of substrate, and 2 mg of microsomal protein fraction. After incubating the reaction mixture for 5 min at 30°C, the reaction was started by adding NADPH to a final concentration of 2 mM. After 4 h, the internal standard uvaol (50 μg per reaction tube) was added and the reaction products were then extracted with ethyl acetate; the sapogenin content was evaluated as described above. CYP72A67 enzymatic activity was tested on the following

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

substrates: β-amirin, erythrodiol, oleanolic acid, hederagenin (Extrasynthese), gypsogenin, gypsogenic acid (from *S. vaccaria*), echinocystic acid, caullophyllogenin, α-boswellic acid, and ursolic acid (Extrasynthese). CYP72A68-430 and CYP72A68-470 enzymatic activities were tested on the following substrates: oleanolic acid, hederagenin, 2β-hydroxy oleanolic acid, and bayogenin.

Quantitative RT–PCR Analysis

Total RNA was extracted from leaves and roots of Mt1903 mutant and wild-type plants (M5 generation) at the vegetative stage; three biological replicates, represented by three individual plants, were analyzed. A Nucleospin RNA plant kit with DNase (Sigma-Aldrich) was used according to the manufacturer's protocol for RNA extraction. cDNA was obtained by iScript cDNA Synthesis Kit (Bio-Rad) starting from 1 μg of RNA. cDNA was diluted 1:5 and 3 μl was used as template in a 10-μl reaction containing 5 μl of SSoFast EvaGreen supermix (Bio-Rad) and specific primers for P450 genes (Supplemental Table 3 online). Thermal cycling conditions were 3 min of initial denaturation at 95°C, 50 cycles of denaturation (95°C, 25 s), annealing and extension (59.5°C, 30 s), and a final melting analyses from 55°C to 95°C with 1°C increase each step. The Msc27 (Pay et al., 1992) and actin control genes were amplified under the same conditions using specific primers. All PCR reactions were carried out in three replicates each in a Rotor-Gene 6000 (Corbett). Data analysis was performed with Rotor-Gene 6000 series Software 1.7 (Corbett). To compare data from different PCR runs and different cDNA samples, the Ct value of the CYP450 genes were normalized against the geometric mean of the Ct values of the two reference genes, Msc27 and actin. Expression levels of the CYP450 genes were calculated by the comparative Ct method using the equation $E = 2^{-\Delta\Delta Ct}$ (where ΔΔCt is the differences in ΔCt between mutant and wild-type or leaves and roots). PCR conditions for each primer combination were optimized for efficiency = 1 and PCR products were verified by melting curve analysis.

Nodulation Assay

Mt1903 mutant and wild-type seeds were scarified by immersion in H_2SO_4 97% for 10 min, rinsed with distilled water, and then sterilized by immersion in a solution of PPM (Plant Preservative Mixture; Micropoli, Italy) 50% v/v for 10 min. Seeds were germinated in sterile water–agar Petri plates at 20°C in the dark. Seedlings with 1–2 cm radicle were transferred on nitrogen-free Jensen agar slants in glass tubes 2 × 20 cm closed with air-permeable plugs. The tubes were put in a growth chamber under controlled conditions (16 h light at 24°C, 8 h darkness at 22°C). After 3 days, the tubes were inoculated with 1 ml of a fresh *S. meliloti* (strain 1021) culture grown in YEM medium and diluted to 2.8×10^8 bacteria/ml.

Statistical Analyses

Analyses of variance were performed using the General Linear Model procedure of SAS Software version 8c (SAS Institute Inc.); plant within-genotype variation was used as the error term. Phenotypic correlations were analyzed using the CORR procedure of the SAS Software.

In the case of gene expression data, the comparative Ct method generated expression values >1 and <1, whose variances were proportional to the squares of the means; consequently, a logarithmic transformation was applied to the expression values to equalize the variances.

SUPPLEMENTAL INFORMATION

Supplemental Information is available at *Molecular Plant Online*.

FUNDING

This research activity has been funded by Regione Lombardia, Fondo per la promozione di Accordi Istituzionali, Project BIOGESTECA 15083/RCC.

AUTHOR CONTRIBUTIONS

E.B. and M.C. performed the research; P. Abbruscato and I.L. performed the TILLING analysis; A.T. and P. Avato contributed chemical analytical tools; O.C. contributed molecular tools; C.S. analyzed the data and wrote the paper; and C.S., A.T., E.B., M.C., and O.C. designed the research.

ACKNOWLEDGMENTS

Technical support from Patrizia Gaudenzi, Agata Ursino (laboratory analyses), Roberto Beghi and Annalisa Carenzi (plant management) of CRA-FLC Lodi, Italy, is gratefully acknowledged. No conflict of interest declared.

Received: March 3, 2015
Revised: June 5, 2015
Accepted: June 7, 2015
Published: June 12, 2015

REFERENCES

Achnine, L., Huhman, D.V., Farag, M.A., Sumner, L.W., Blount, J.W., and Dixon, R.A. (2005). Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. Plant J. **41**:875–887.

Agrell, J., Anderson, P., Oleszek, W., Stochmal, A., and Agrell, C. (2004). Combined effects of elevated CO_2 and herbivore damage on alfalfa and cotton. J. Chem. Ecol. **30**:2309–2324.

Ahmed, V.U., Ahmed, W., and Usmanghanik, K. (1992). Triterpenoid saponins from leaves of *Castanospermum australe*. Phytochemistry **31**:2805–2807.

Bialy, Z., Jurzysta, M., Oleszek, W., Piacente, S., and Pizza, C. (1999). Saponins in alfalfa (*Medicago sativa* L.) root and their structural elucidation. J. Agric. Food Chem. **47**:3185–3192.

Bialy, Z., Jurzysta, M., Mella, M., and Tava, A. (2006). Triterpene saponins from the roots of *Medicago hybrida*. J. Agric. Food Chem. **54**:2520–2526.

Boisson-Dernier, A., Chabaud, M., Garcia, F., Bécard, G., Rosenberg, C., and Barker, D.G. (2001). Hairy roots of *Medicago truncatula* as tools for studying nitrogen-fixing and endomycorrhizal symbioses. Mol. Plant Microbe Interact. **14**:693–700.

Borel, C., Gupta, M.P., and Hostettman, K. (1987). Molluscicidal saponins from *Swartzia simplex*. Phytochemistry **26**:2685–2689.

Carelli, M., Biazzi, E., Panara, F., Tava, A., Scaramelli, L., Porceddu, A., Graham, N., Odoardi, M., Piano, E., Arcioni, S., et al. (2011). *Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins. Plant Cell **23**:3070–3081.

Carelli, M., Calderini, O., Panara, F., Porceddu, A., Losini, I., Piffanelli, P., Arcioni, S., and Scotti, C. (2013). Reverse genetics in *Medicago truncatula* using a Tilling mutant collection. Methods Mol. Biol. **1069**:101–118.

Carelli, M., Biazzi, E., Tava, A., Losini, I., Abbruscato, P., Depedro, C., and Scotti, C. (2015). Sapogenin content variation in *Medicago* interspecific hybrid derivatives highlights some aspects of saponin synthesis and control. New Phytol. **206**:303–314.

Delis, C., Krokida, A., Georgiou, S., Peña-Rodriguez, L., Kavroulakis, N., Ioannou, E., Roussis, V., Osbourn, A.E., and Papadopoulou, K. (2011). Role of lupeol synthase in *Lotus japonicus* nodule formation. New Phytol. **189**:335–346.

Fukushima, E.O., Seki, H., Ohyama, K., Ono, E., Umemoto, N., Mizutani, M., Saito, K., and Muranaka, T. (2011). CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis. Plant Cell Physiol. **52**:2050–2061.

Fukushima, E.O., Seki, H., Sawai, S., Suzuki, M., Ohyama, K., Saito, K., and Muranaka, T. (2013). Combinatorial biosynthesis of legume natural and rare triterpenoids in engineered yeast. Plant Cell Physiol. **54**:740–749.

Please cite this article in press as: Biazzi et al., CYP72A67 Catalyzes a Key Oxidative Step in *Medicago truncatula* Hemolytic Saponin Biosynthesis, Molecular Plant (2015), http://dx.doi.org/10.1016/j.molp.2015.06.003

Gietz, D., St Jean, A., Woods, R.A., and Schiestl, R.H. (1992). Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res. **20**:1425.

Jia, Z., Koike, K., Kudo, M., Li, H., and Nikaido, T. (1998). Triterpenoid saponins and sapogenins from *Vaccaria segetalis*. Phytochemistry **48**:529–536.

Kevei, Z., Lougnon, G., Mergaert, P., Horvath, G.V., Kereszt, A., Jayaraman, D., Zaman, N., Marcel, F., Regulskj, K., Kiss, G.B., et al. (2007). 3-Hydroxy-3-methylglutaryl coenzyme A reductase1 interacts with NORK and is crucial for nodulation in *Medicago truncatula*. Plant Cell **19**:3974–3989.

Krokida, A., Delis, C., Geisler, K., Garagounis, C., Tsikou, D., Peña-Rodriguez, L.M., Katsarou, D., Field, B., Osbourn, A.E., and Papadopoulou, K. (2013). A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. New Phytol. **200**:675–690.

McCallum, C.M., Comai, L., Greene, E.A., and Henikoff, S. (2000). Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. Plant Physiol. **123**:439–442.

Meesapyodsuk, D., Balsevich, J., Reed, D.W., and Covello, P.S. (2007). Saponin biosynthesis in *Saponaria vaccaria* cDNAs encoding β-Amyrin synthase and a triterpene carboxylic acid glucosyltransferase. Plant Physiol. **143**:959–969.

Meyer, B.N., Ferrigni, N.R., Putnam, J.E., Jacobsen, L.B., Nichols, D.E., and Mc Laughlin, J.L. (1982). Brine shrimp: a convenient general bioassay for active plant constituents. Planta Med. **45**:31–34.

Moniuszko-Szajwaj, B., Pecio, L., Kowalczyk, M., Simonet, A.M., Macias, F.A., Szumacher-Strabel, M., Cieslak, A., Oleszek, W., and Stochmal, A. (2013). New triterpenoid saponins from the roots of *Saponaria officinalis*. Nat. Prod. Commun. **8**:1687–1690.

Moses, T., Pollier, J., Almagro, L., Buyst, D., Van Montagu, M., Pedreño, M.A., Martins, J.C., Thevelein, J.M., and Goossens, A. (2014a). Combinatorial biosynthesis of sapogenins and saponins in *Saccharomyces cerevisiae* using a C-16α hydroxylase from *Buplerum falcatum*. Proc. Natl. Acad. Sci. USA **111**:1634–1639.

Moses, T., Pollier, J., Faizal, A., Apers, S., Pieters, L., Thevelein, J.M., Geelen, D., and Goossens, A. (2015). Unravelling the triterpenoid saponin biosynthesis of the African shrub *Maesa lanceolata*. Mol. Plant **8**:122–135.

Moses, T., Thevelein, J.M., Goossens, A., and Pollier, J. (2014b). Comparative analysis of CYP93E proteins for improved microbial synthesis of plant triterpenoids. Phytochemistry **108**:47–56.

Naoumkina, M.A., Modolo, L.V., Huhman, D.V., Urbanczyk-Wochniak, E., Tang, Y., Sumner, L.W., and Dixon, R.A. (2010). Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. Plant Cell **22**:850–866.

Pay, A., Heberle-Bors, E., and Hirt, H. (1992). An alfalfa cDNA encodes a protein with homology to translationally controlled human tumor protein. Plant Mol. Biol. **19**:501–503.

Pecetti, L., Tava, A., Romani, M., De Benedetto, M.G., and Corsi, P. (2006). Variety and environment effects on the dynamics of saponins in lucerne (*Medicago sativa* L.). Eur. J. Agron. **25**:187–192.

Pecetti, L., Biazzi, E., and Tava, A. (2010). Variation in saponin content during the growing season of spotted medic [*Medicago arabica* (L.) Huds]. J. Sci. Food Agric. **90**:2405–2410.

Pollier, J., Moses, T., González-Guzmán, M., De Geyter, N., Lippens, S., Vanden Bossche, R., Marhavý, P., Kremer, A., Morreel, K., Guérin, C.J., et al. (2013). The protein quality control system manages plant defence compound synthesis. Nature **504**:148–152.

Pompon, D., Louerat, B., Bronine, A., and Urban, P. (1996). Yeast expression of animal and plant P450s in optimized redox environments. Methods Enzymol. **272**:51–64.

Porceddu, A., Panara, F., Calderini, O., Molinari, R., Taviani, P., Lanfaloni, L., Scotti, C., Carelli, M., Scaramelli, L., Bruschi, G., et al. (2008). An Italian functional genomic resource for *Medicago truncatula*. BMC Res. Notes **1**:129.

Reid, D.E., Ferguson, B.J., Hayashi, S., Lin, Y.H., and Gresshoff, P.M. (2011). Molecular mechanisms controlling legume autoregulation of nodulation. Ann. Bot. **108**:789–795.

Seki, H., Sawai, S., Ohyama, K., Mizutani, M., Ohnishi, T., Sudo, H., Fukushima, E.O., Akashi, T., Aoki, T., Saito, K., et al. (2011). Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. Plant Cell **23**:4112–4123.

Suzuki, H., Achnine, L., Xu, R., Matsuda, S.P., and Dixon, R.A. (2002). A genomics approach to the early stages of triterpene saponin biosynthesis in *Medicago truncatula*. Plant J. **32**:1033–1048.

Suzuki, H., Reddy, M.S.S., Naoumkina, M.A., Aziz, N., May, G.D., Huhman, D.V., Sumner, L.W., Blount, J.W., Mendes, P., and Dixon, R.A. (2005). Methyl jasmonate and yeast elicitor induce differential transcriptional and metabolic re-programming in cell suspension cultures of the model legume *Medicago truncatula*. Planta **220**:696–707.

Tabe, L.M., Wardley-Richardson, T., Ceriotti, A., Aryan, A., McNabb, W., Moore, A., and Higgins, T.J. (1995). A biotechnological approach to improving the nutritive value of alfalfa. J. Anim. Sci. **73**:2752–2759.

Tava, A., and Pecetti, L. (1998). Hemolytic activity and saponin content in lucerne (*Medicago sativa* complex) genotypes. J. Genet. Breed. **52**:33–37.

Tava, A., and Avato, P. (2006). Chemical and biological activity of triterpene saponins from *Medicago* species. Nat. Prod. Commun. **1**:1159–1180.

Tava, A., and Pecetti, L. (2012). Chemical investigation of saponin from twelve annual *Medicago* species and their bioassay with the brine shrimp *Artemia salina*. Nat. Prod. Commun. **7**:837–840.

Tava, A., Oleszek, W., Jurzysta, M., Berardo, N., and Odoardi, M. (1993). Alfalfa saponins and sapogenins: isolation and quantification in two different cultivars. Phytochem. Anal. **4**:269–274.

Tava, A., Mella, M., Avato, P., Argentieri, M.P., Bialy, Z., and Jurzysta, M. (2005). Triterpenoid glycosides from leaves of *Medicago arborea* L. J. Agric. Food Chem. **53**:9954–9965.

Tava, A., Mella, M., Avato, P., Biazzi, E., Pecetti, L., Bialy, Z., and Jurzysta, M. (2009). New triterpenic saponins from the aerial parts of *Medicago arabica* (L.) Huds. J. Agric. Food Chem. **57**:2826–2835.

Tava, A., Pecetti, L., Romani, M., Mella, M., and Avato, P. (2011a). Triterpenoid glycosides from the leaves of two cultivars of *Medicago polymorpha* L. J. Agric. Food Chem. **59**:6142–6149.

Tava, A., Scotti, C., and Avato, P. (2011b). Biosynthesis of saponins in the genus *Medicago*. Phytochem. Rev. **10**:459–469.

Watson, B.S., Bedair, M.F., Urbanczyk-Wochniak, E., Huhman, D.V., Yang, D.S., Allen, S.N., Li, W., Tang, Y., and Sumner, L.W. (2015). Integrated metabolomics and transcriptomics reveal enhanced specialized metabolism in *Medicago truncatula* root border cells. Plant Physiol. **167**:1699–1716.