דו"ח לתכנית מחקר מספר **261-0929-14**

מיזם איטלקי-ישראלי לשיפור ערכים תזונתיים של גידולים ממשפחת הסולניים
ITalian-IsraeLi initiative for the NUTRItional improvement of SOLanaceus crops (IT-IL - NUTRI-SOL)

דוח סופי, מוגש לקרן המדען הראשי במשרד החקלאות

עידית גינזברג ועדנה פוגלמן	חקר ירקות, מינהל המחקר החקלאי, מכון וולקני
מיכל אורן-שמיר ורינת עובדיה	פרחים וצמחי נוי, מינהל המחקר החקלאי, מכון וולקני
יוסף הירשברג	בוטניקה וגנטיקה, האוניברסיטה העברית, ירושלים
אסף אהרוני	מדעי הצמח, מכון ויצמן למדע, רחובות

Idit Ginzberg, Dept. of Vegetable Research, ARO, the Volcani Center, PO Box 6, Bet Dagan, Israel 50250. E-mail: iditgin@volcani.agri.gov.il
Michal Oren-Shamir, Dept. of Ornamental Horticulture, ARO, the Volcani Center, PO Box 6, Bet Dagan, Israel 50250. E-mail: vhshamir@volcani.agri.gov.il
Joseph Hirschberg, Dept. of Genetics, The Hebrew University, Jerusalem. E-mail: hirschu@vms.huji.ac.il
Asaph Aharoni, Dept. of Plant Sciences and the Environment, Weizmann Institute of Science, Rehovot. E-mail: asaph.aharoni@weizmann.ac.il

תקציר

הצגת הבעיה - תפו"א הינו אחד מהמזונות הבסיסיים, עשיר בחלבונים, סיבים תזונתיים, ויטמינים ומינרלים. תכניות השבחה של זנים חדשים מכוונות לשיפור הערך הבריאותי של הפקעת באמצעות העלאה של קרוטנואידים ואנתוציאנינים, בעלי פעילות אנטי-חימצונית, והפחתה של גליקואלקלואידים להם פעילות שלילית על הבריאות. רמות מטבוליטים אלו ייחודיות לזן, אולם גורמים התפתחותיים, רקמתיים ותנאי גידול קובעים את הרמה הסופית שלהם ברקמה.
מטרת המחקר - המחקר הנוכחי הינו שיתוף פעולה בין קבוצות מחקר באיטליה ובישראל ומטרתו לנצל גישות גנומיות על מנת לגלות את המנגנונים שמבקרים את רמת מטבוליטים אלו, מה שיאפשר פיתוח זני תפו"א בעלי ערך בריאותי גבוה. דוח זה מסכם את שלוש שנות המחקר של הצוות הישראלי.
שיטות - טיפול עקת חום נעשה על צמחי תפו"א שגודלו בחממה בכלי גידול המאפשרים חימום של אזור הפקעות. השפעת טמפרטורות קרקע על ביטוי גנים במסלול המטבוליזם של אנתוציאנינים, גליקואלקלואידים וקרוטנואידים בקליפה ובפארנכימה של הפקעת נבחנה באמצעות q-PCR ובאנליזת טרנסקריפטום (RNA-seq). במקביל נבדק פרופיל המטבוליטים בפקעות של זנים איטלקיים ובזנים מקומיים. אנליזות מטבולית וטרנסקריפטומית נערכו לעגבניה וחציל ללימוד המטבוליזם של גליקואלקלואידים בגידולים אלה.
תוצאות עיקריות ומסקנות - טיפול חום של כשבועיים הוביל לירידה מובהקת ברמת המטבוליטים של גליקואלקלואידים, במקביל לירידה בביטוי גנים קריטיים במסלול הביוסינטטי. בדומה, נמדדה ירידה בביטוי הגנים הביוסינטטיים לאנתוציאנינים אולם רמת הפיגמנטים לא ירדה, כנראה בשל ייצוב הפיגמנטים הקיימים בוקואולה התאית. בעקבות טיפול החום רמת הקרוטנואידים לא השתנתה בפארנכימת הפקעת אולם הרכבם השתנה. בנוסף התוצאות בעבודה זו היוו חלק חשוב בהבנת מסלול היצירה של הגליקואלקלואידים במשפחת הסולניים שכן מספר גנים מבניים בודדו ואופיינו במהלך הפרוייקט. גנים מבניים נוספים אותרו באמצעות סריקה של רמת הגליקואלקלואידים בעלים של אוכלוסיית אינטרוגרסיה של עגבניה עבור מקטע המשפיע על רמות גליקואלקלואידים, ובודדו שני גני המעורבים בייצור של Esculeosides, הגליקואלקלואידים העיקריים בעגבניה הבשלה. פרט לגנים מבניים אופיין גן בקרה השולט ככל הנראה ברמת ייצור הכולסטרול, אבן הבניין לייצור הגליקואלקלואידים. שיתוף פעולה הדוק עם קבוצה איטלקית הוביל ליצירת קווי חציל מהונדסים המכילים רמה נמוכה משמעותית של גליקואלקלואידים בפירות. בנוסף, נבנה מאגר מידע מטבולי וביטוי גנים בכ- 77 רקמות ואיברים של צמח החציל המשמש לאיתור ואיפיון מסלול הביוסינטזה והבקרה עליו בגידול חשוב זה.

.....

הצהרת החוקר הראשי:

הממצאים בדו"ח זה הינם תוצאות ניסויים.

הניסויים מהווים המלצות לחקלאים: לא

חתימת החוקר _____ תאריך: _1-7-2015_____

Final Scientific Report

Cover page

Italy-Israel Project Number: ID16/ID4 (IL: 261-0929-11)

Project Title: ITalian-IsraeLi initiative for the NUTRItional improvement of SOLanaceus crops (IT - IL - NUTRI - SOL)

Participants (IL):	**Affiliated Institutions:**
Idit Ginzberg	ARO - Volcani Center, Bet Dagan
Michal Oren-Shamir	ARO - Volcani Center, Bet Dagan
Joseph Hirschberg	The Hebrew University, Jerusalem
Asaph Aharoni	Weizmann Institute of Science, Rehovot

Keywords: health-promoting secondary metabolites, growth in the Mediterranean hot climate, steroidal glycoalkaloids, anthocyanins, carotenoids, potato, tomato, eggplant

Israeli Budget (Total NIS): 1,320,000 NIS

Signature

Italy Project leader Italy — Authorizing Official, Principal Institution

Idit Ginzberg

Israel Project leader Israel — Authorizing Official, Principal Institution

Appendix G7-1

- **Abbreviations**

Carotenoid genes

Bch1 beta-carotene hydroxylase 1
Bch2 beta-carotene hydroxylase 2
Ccd1 carotene cleavage dioxygenase 1
Ccd4 carotene cleavage dioxygenase 4
Dxs 1-deoxy-D-xylulose-5-phosphate synthase
GGPS geranylgeranyl diphosphate synthase.
Idi1 IPP isomerase (plastidial)
Lcy-b lycopene beta-cyclase
Lcy-e lycopene epsilon-cyclase
Pds phytoene desaturase
Psy1 phytoene synthase 1
Zep zeaxanthin epoxidase

Anthocyanin genes

CHS chalcone synthase
CHI chalcone isomerase
F3H flavonoid 3-hydroxylase
F3'H flavonoid 3'-hydrolase
F3'5'H flavonoid 3'5'-hydroxylase
DFR dihydroflavonol 4-reductase
ACS anthocyanidin synthase
3GT UDP-glucose:flavonoid 3-*O*-glucosyltransferase

Glycoalkaloid genes

HMG1 3-hydroxy-3-methylglutaryl-coenzyme A reductase 1
GAME glycoalkaloid-metabolism
SQS1 squalene synthase 1
SMT1 sterol methyltransferase 1
CAS1 cycloartenol synthase 1
SGT1 solanidine galactosyltransferase
SGT2 solanidine glucosyltransferase
SGT3 β-chaconine/β-solanine rhamnosyltransferase
SA Steroidal alkaloids
SGA Steroidal glycoalkaloids

Techniques

HPLC High Performance Liquid Chromatography
ELSD Evaporative Light Scattering Detector
GC-MS Gas Chromatography – Mass Spectrometer
LC-MS Liquid Chromatography – Mass Spectrometer
UPLC-QTOF-MS Ultra Performance Liquid Chromatography-Quadrupole Time of Flight Mass Spectrometer
ESI Electro Spray Ionization
qPCR Quantitative Real Time PCR
RNA-Seq RNA sequencing
RAD Restriction-site Associated DNA
PCA Principal Component Analysis
HC Hierarchical Clustering
NCA Network Correlation Analysis

Plant materials

ILs Introgression lines
RIL Recombinant Inbred Lines

- **Publication Summary**

	Joint Israel/Italy authorship	Italy Authors only	Israeli Authors only	Total
Refereed (published, in press, accepted)			1	1
Submitted, in review, in preparation			1 (in review)	1
Invited reviews			1	1
Book chapters			1	1
Books				
Master theses				
Ph.D. theses				
Abstracts				
Not refereed (proceedings, reports, etc.)				

- **Abstract**

The potato is the fourth most important staple food for mankind. Breeding of potatoes for nutritional value is mainly focused on carotenoid, anthocyanin and glycoalkaloid contents. Carotenoids and anthocyanins exhibit antioxidant activity, therefore increasing their level in the tuber is a desirable trait. Glycoalkaloids contribute to the characteristic taste of the potato and have been reported to exhibit anti-cancer activity, but are toxic to humans and their level in the tuber must be regulated carefully. The glycoalkaloids are common constituents of other members of the Solanaceae family, such as tomato and eggplant. The levels of the abovementioned secondary metabolites are genetically determined; however, developmental, tissue specific and environmental cues affect their final content – most relevant for both Israel and Italy are the hot conditions that prevail in the Mediterranean basin.

To this end the following objectives (as listed in the approved proposal) were addressed: (a) The level of carotenoid, anthocyanin and glycoalkaloid metabolites and their biosynthetic gene expression were monitored in a panel of potato genotypes, eggplant introgression and recombinant inbred lines and of tomato mapping population, including carotenoid- and anthocyanin- enriched clones and their crosses, and divergent crosses for glycoalkaloid content; (b) Stability of carotenoid, anthocyanin and glycoalkaloid traits in potato were measured in selected clones under heat stress conditions typical of the Mediterranean region; (c) Potato, eggplant and tomato lines divergent for their glycoalkaloid content were subjected to metabolomic and transcriptomic analyses to identify new glycoalkaloid-related genes and QTL for low glycoalkaloid level.

Potato cultivars from Italy and Israel enriched with carotenoids in tuber flesh or with anthocyanins in their peel or exhibiting whitish peel and tuber flesh, were used to study the effect of high temperature on the level of anthocyanins, carotenoids and steroidal glycoalkaloids (SGAs) metabolites and the respective biosynthetic gene expression in the tuber. Gene expression studies, using RNA-seq and qPCR, showed heat-induced down regulation of specific biosynthetic genes of SGAs, anthocyanin and carotenoids. By mapping the heat transcriptome with KEGG database we showed that the reduced gene expression was not a global heat response but affected specific metabolic pathways. Targeted metabolomics indicated that the concentration of potato SGA metabolites was reduced, in accordance to the transcriptomic data; however the concentration of anthocyanin pigments remained unchanged probably due to their stabilization in the vacuole. Total carotenoids level did no change significantly in potato tuber flesh; however their composition was changed.

Overall, results suggest that growth in hot climate may alter tuber secondary metabolism in a selective manner, altering its nutritive value and the composition of health-promoting components.

In a second major part of the joint project we have been working on several species of the Solanacea family, more specifically tomato, potato and eggplant to elucidate the biosynthesis and regulation of SGAs. Through a combination of gene expression and metabolite data we revealed the biosynthetic pathway leading to SGAs biosynthesis starting from the precursor cholesterol. Metabolic screening of an introgression population resulted in the identification of 2 genes associated with the biosynthesis of Esculeosides, the major SGAs produced by ripening tomato fruit. We have also revealed the first component of the regulatory network controlling the biosynthesis of SGAs in tomato, potato and likely other Solanacea species. A regulatory gene identified in the course of the project appears to control the biosynthesis of the SGA precursor cholesterol. Through intensive collaboration with an Italian partner we have made major strides towards the understanding of SGAs biosynthesis in eggplant. Transgenic plants were generated which are silenced in GAME4, a key gene in SGAs biosynthesis and their fruit were shown to possess significantly reduced levels of SGAs. We also generated a large-scale data base that includes metabolite and gene expression data from 77 eggplant tissues and organs. This data will be of major importance to future work on SGAs and other fruit quality traits in this important and relative neglected scientifically crop plant.

All together the work on SGAs will have a major scientific impact on our understanding of secondary metabolism in general and more specifically regarding its role in chemical defense mechanisms. The results are also important with respect to lowering the levels of anti-nutritional compounds (i.e. SGAs) in crops of the Solanaceae family as well as for our understanding of other fruit quality traits.

- **Achievements**

o **<u>Chapter I</u>** - ***ARO-Volcani/HUJI labs***: The effect of high soil temperatures on anthocyanin, glycoalkaloids and carotenoids biosynthetic pathways and metabolite level

Anthocyanin, glycoalkaloids and carotenoids in potato tuber. The peel of potato tuber is made of a periderm tissue that accumulates protective components. The outer skin cells (phellem) have suberized cell walls, and the inner parenchyma-like phelloderm accumulate secondary metabolites, among which are the anthocyanins (in red skin cultivars) and the steroidal glycoalkaloids (SGAs). Yellow cultivars accumulate carotenoids in their tuber flesh. Hence, defined tuber tissues were used in the present proposal to monitor the target metabolites.

The experimental setting included growth of potato plants in a greenhouse at the Volcani Center. Heat treatment was applied 7-weeks after sprout emergence by circulating hot water (33-35°C) in tubes lining the internal side of the pots. Tubers from heat treatment and control were collected at 8 weeks post sprouting when the skin and the phelloderm layers can be easily separated from the tuber flesh. RNA extractions and expression analyses using qPCR and RNA-seq, as well as determination of tuber color and metabolite profiling of anthocyanins, SGAs and carotenoids were performed as detailed in the original proposal. The main scientific achievements are detailed below.

1. Heat-induce fading of Desirée tubers red pigmentation and down-regulation of anthocyanin biosynthetic pathway

The red tint of Desirée tubers was reduced following the heat treatment (Table 1, a* values). In accordance, RNA-seq by Illumina indicated reduced expression of anthocyanin biosynthetic genes in tuber phelloderm (Fig 1a) which was verified by qPCR (Table 2). Interestingly, HPLC determination of the anthocyanidin and phenolic acids in the phelloderm (Table 3) indicated no significant change in their levels in that tissue. Please note that the level of the anthocyanidins pelargonidin and peonidin in the phelloderm is significantly greater – as expected – than in the skin.

The contradiction between down regulation of the biosynthetic pathway while no change in pigments level may be explained by stabilization of the anthocyanins in the vacuole. This hypothesis is supported by unchanged expression level of the *vATPase-PH4* – whose ortholog from petunia was shown to regulate vacuolar acidity and final color of the petals – upon heat treatment (Table 2 and Fig 1a). The fading of the red tint of the tuber may result from heat induced alteration in skin morphology, as we showed before (Ginzberg et al., 2009). It was shown by Noda et al. (1994) that cell shape affects the proportion of light that enters it which in turn affects light absorption by the pigments and hence color intensity.

2. Heat-induce reduction of the steroidal glycoalkaloids chaconine and solanine

Similar approach was conducted with respect to SGA metabolism. Previous reports in the literature indicated of accumulation of SGA metabolites upon stresses, including heat stress. Our data showed reverse phenomenon in the phelloderm (the main SGA producing cells in the tuber) – SGA biosynthesis (Fig 1b and Table 2) and accumulation (Table 4) were reduced upon the heat conditions applied in our system. These included the *HMG1* gene whose expression we showed previously to correlate SGA levels (Ginzberg et al., 2012) and the newly discovered GAME genes by the group of Aharoni, albeit GAME4 was not reduced significantly (Table 2). Of the major SGAs of the cultivar Desirée, both chaconine and solanine levels were reduced, although the later was not statistically significant (Table 4)

3. Heat-induce alteration in carotenoids composition

Transcriptome analysis of genes involved with carotenoid metabolism in Desirée tuber flesh showed mainly up-regulation or no change in gene expression with only few down regulated genes (Figure 1c). Down regulation of *ZEP* with no change in *VDE* expression was expected to results with accumulation of zeaxanthine, β-carotene and lutein, and reduction in the levels of antheraxanthin and violaxanthin, compared to untreated controls. Up-regulation of *NXS* may result with increase in neoxanthine (Figure 2). This pattern of carotenoid accumulation was obtained in heat-treated Desirée tuber flesh as expected, but also in the tuber flesh of Melrose and tuber peel of Andean Sunrise (Figure 3). Total carotenoid levels remained the same or increased upon the heat exposure (Figure 3), suggesting that heat stress may alter carotenoid composition with no reduction in total level.

4. Heat-induce alteration was limited to specific metabolic branches

To test that the above described down regulated pathways are not a result of heat-induced down regulation of global tuber phelloderm and flesh transcriptomes, our sequence data was aligned against the KEGG database and respective pathway maps were prepared. The phenylpropanoid pathway from which the anthocyanin pathway is branching (Figures 4a and 4b, respectively), the mevalonate/phytosterol pathways from which the SGA biosynthesis is branching (Figure 4c), and the carotenoid biosynthesis (Figure 4d). Most of the pathways that provide precursors for the anthocyanins, SGAs and tuber carotenoids, are not altered by the heat treatment (Figure 4), but the specific biosynthetic steps of the tested metabolites.

- **Chapter II** - ***Weizman lab*:** Elucidation and manipulation of the glycoalkaloid pathway in tubers and fruit of solanaceae family crop species

1. Gene Discovery of SGAs pathway genes in tomato

As part of task no. 6 we carried out transcriptome analysis of tomato tissues (all together 21 different tissues) by RNAseq. The data from tomato was used for co-expression analysis in which the previously identified SGA genes were used as baits. We also generated similar co-expression data in potato (publically available), employing the same gene baits as in tomato, to identify SGA genes in potato. Thus, we ended up with a list of genes that were highly co-expressed with the same gene baits in both tomato and potato. This list pointed us to the discovery that the SGAs pathway genes in the Solanacea family are positioned in metabolic gene clusters. In potato and tomato these are in chromosome 7 and chromosome 12. All together we could identify 10 genes involved in the pathway. In tomato they include 4 cytochrome p450s (GAME4, 6, 7, and 8); 4 glycosyltransferases (GAME1, 2, 17 and 18); a dioxygenase (GAME 11) and a transaminase (GAME12) (see Figure 5). The genes were characterized through VIGS analyses in tomato, enzyme assays of recombinant proteins in E.coli and in stable transgenic lines. In one example, we generated tomatine from tomatidine by combining tomatidine in a single reaction tube with UDP -glucose, -galactose and –galactose and the 4 recombinant glycosyltransferase enzymes (i.e. GAME 1, 2, 17 and 18; see Figure 6).

Some specialized plant metabolites, particularly terpenoids, are the result of activities from clusters of genes. The existence of metabolic gene clusters raises questions regarding the advantages of such genomic organization. Reducing the distance between loci, resulting in coinheritance of advantageous combinations of alleles, may be one benefit of clustering. Clustering glycosyltransferases and core pathway genes, as observed here for SGAs, could maintain allelic combinations that support the metabolic outcome needed by the plant and reduce formation of phytotoxic, aglycone, compounds. We found that the regions of coexpressed genes in both chromosomes (i.e., 7 and 12) were flanked by similarly annotated genes and positioned identically along the genome, although poorly coexpressed with GAME1/SGT1 and GAME4 and likely not related to SGAs metabolism. This suggests a duplication event that facilitated the positioning alongside on chromosome 12 of GAME4 and GAME12, both STSs-SGAs branch point genes. Subsequent evolution of enzyme function of this gene pair likely allowed plants in the Solanaceae family to start producing the nitrogen containing steroidal alkaloids.

Furthermore, using the newly identified genes we have shown that SGA levels can be severely reduced in potato tubers by modifying expression of an enzyme in the biosynthetic pathway. The lack of SGAs in such plants might make them sensitive to biotic stress and the increased production of STSs (as occurred in GAME4-silenced plants), which are non-toxic to warm-blooded species, including humans, might provide a compensatory defense mechanism. The findings open the way for developing new strategies, through genetic engineering or more classical breeding programs, to reduce quantities of the antinutritional SGAs in key crops of the Solanaceae including potato, tomato, and eggplant. At the same time, it pro-vides a platform for studying the SGA and STS biosynthetic pathways, transport and regulatory systems that control the production of thousands of these chemicals in specific plant lineages.

2. Deciphering the biosynthesis of Esculeoside A

Previous work in the lab proposed a new pathway for tomatine catabolism and formation of modified SGAs during fruit development and ripening. While in the green tomato fruit α-tomatine is the most abundant SGA, esculeosides and lycoperosides are predominant in the red ripe stage (Figure 8). Total SGAs content remains constant across fruit ripening and this suggested that the entire pool of α-tomatine is converted during this process to the latter compounds. Elucidation of tomato lycoperosides and esculeosides structures revealed hydroxylation, acetylation and glycosylation on the aglycone.

In recent years several groups suggested possible pathway intermediates through which *α*-tomatine is converted to lycoperosides and esculeosides in the course of fruit maturation in tomato Figure 8B). It has also been reported that the majority of pathway reactions in fruit leading to esculeosides formation are independent from the ripening regulatory system and ethylene. In contrast, in *non-ripening* (*nor*) and *ripening-inhibitor* (*rin*) tomato mutants that display altered ripening, levels of esculeoside A were decreased, yet, they exhibited an increased content of upstream acetylated and glycosylated intermediates. This suggested that the glycosylation step in which acetoxy-hydroxytomatine is converted to esculeoside A and further modifications towards lycopersoides formation are associated with ethylene and ripening regulated events (Figure 8). To date, however, no genes or enzymes have been reported that were associated with this pathway.

3. SGAs profiling of leaf extracts derived from a backcross introgression lines (BILs) population and identification of the GAME31 candidate, a putative 2-oxoglutarate-dependant dioxygenase

To profile a set of 8 different SGAs in the 620 lines backcross introgression lines (BILs; Figure 9) we first developed a rapid (10 minutes long gradient per sample), highly sensitive and selective targeted method using an Ultra Performance-LC triple-quadropole MS (UPLC-qqq-MS) in the multiple reaction monitoring mode. A method for SGAs extraction from leaves was adapted from Schilmiller *et al.* (2012) and used for efficient preparation of BILs samples. Eight SGAs were analyzed using UPLC-qqq-MS across the entire population and 11 regions that contained the most significant changes in one or more SGAs were identified. One of the significant introgressed regions in which we detected increased hydroxy- and acetoxy-hydroxy- tomatine contained a minimal set of 17 genes among them 4 homologs of the 2-oxoglutarate-dependant dioxygenases. Member of this family of oxygenases/hydroxylases could likely perform the hydroxylation reaction predicted on the *α*-Tomatine backbone. Analyzing the publically available transcriptome data of the 77 ILs provided evidence that one of the 4 tandem dioxygenases is underlying the change in hydroxy- and acetoxyhydroxy-tomatine levels (see Figure 8). Expression of the specific dioxygenase (i.e. GAME31) was most dramatically elevated in the IL2.1 encompassing the altered introgression region of interest (other 3 dioxygenase homologs did not change in expression between ILs). We are currently investigating the genome sequence of GAME31 to identify a sequence change associated with its altered expression.

4. Identification of the GAME5 UDP-glycosyltransferase putatively catalyzing Esculeoside A formation in ripening tomato fruit

Apart from *GAME31* we currently have evidence that GAME5, a putative UDP-glycosyltransefrase, is the enzyme responsible for the reaction in which acetoxy-hydroxytomatine is converted to Esculeoside A through glycosylation on the side chain (see Figure 8). As mentioned above, this glycosylation step and further modifications towards esculeosides and lycoperosides formation are associated with ethylene and ripening regulated events. Indeed, experiments in our lab and others observed that *GAME5* or the *ERT1* transcript (the ERT1 sequence corresponds to GAME5; Picton *et al.* 1993), increased during the ripening of wild-type fruit and showed reduced accumulation in the ripening *rin* mutant fruit. Virus Induced Gene Silencing (VIGS) experiments in which a large number of UDP-glycosyltransferases have been screened by the A. Bovy and J. Beekwilder lab (WUR, The Netherlands), detected that silencing of *GAME5* in fruit results in decreased levels of Esculeoside A and at the same time an increase in its putative substrate acetoxy-hydroxytomatine (data not shown). In addition, in collaboration with the Beekwilder lab we showed that the recombinant GAME5 produced in *E.coli* cell could perform the same reaction *in vitro* (Figure 10).

5. The α-Tomatine catabolic pathway is blocked in the 'bitter' fruit accession

In preliminary analysis we have shown that α-Tomatine accumulates in the red, ripe fruit of the tomato accession possessing a bitter flavor (identified by the Zamir lab, HUJI), mainly in the fruit peel/skin tissue. Segregation analysis of a cross between the bitter and sweet accessions demonstrated tight correlation between the accumulation of α-Tomatine and bitterness. It also suggested a single recessive mutation. The pathway starting from α-Tomatine is most likely blocked or modified in the reaction catalyzing the conversion of acetoxy-hydroxytomatine to Esculeoside A (see Figure 11). Principal component analysis of peel tissue derived from bitter and sweet fruit representing mass signals from non-targeted LC-MS analysis demonstrated that the 'bitter' phenotype is associated with mid to late stages of fruit development and ripening (Figure 12).

6. Performing large-scale transcriptome and metabolomics analyses of 77 different tissues of eggplant.

In this project, we have characterized both the transcriptome and the metabolome of the aubergine (*Solanum melonga*) plants. To this end, 81 tissues of *S. melongena* cv. DR2 plants were obtained from the group of Prof. Rotino. The tissue samples arrived to our lab in dry ice and were kept frozen until processed. The tissues were

grinded in liquid nitrogen and two aliquots were prepared. On one hand, exactly 100mg ±2mg from each tissue were kept for metabolites extraction, while a second aliquot (100 to 150mg) was kept for RNA extraction.

Eggplant Metabolomics:

Semipolar and polar metabolites from 76 tissues were extracted using our standard extraction procedure. Metabolites from the five developmental stages of eggplant seeds were not extracted. Briefly, 100mg ±2mg frozen and grinded tissues were extracted with 300µl 80% methanol, supplemented with 0.1% formic acid, followed by 20min. sonication. The samples were then cleared by centrifugation (10min. at maximum speed), and filtered through 0.22 µm membrane syringe filters. The samples were then analyzed using UPLC-qTOF mass spectrometry in positive modes. In a non-targeted approach, 485 mass signals were detected from which 84 were putatively identified and additionally 13 steroidal glycolkaloids and saponins were clearly identified (Figure 13).

Eggplant transcriptomics:

RNA from all 81 tissues was extracted using the standard tri-reagent method. RNA samples were then used to create RNAseq libraries. The obtained libraries were aligned to the recently published aubergine genome. A total of 11 700 transcripts were detected with a significant expression level. Among these, we identified 12 genes that are homologues to the genes characterized in our previous work, which are involved in the metabolism of steroidal glycoalkaloids (Figure 13).

Correlation analysis of the transcripts and metabolites related to steroidal glycol-alkaloids/saponins:

The data obtained from the analysis above was analysed to decipher the metabolic pathway of the aubergine's glycolalkaloids/saponins. To present these data, heat maps showing the abundance of the transcripts and metabolites involved were generated (Figure 8). In these data sets, we observed that the genes involved in glycolalkaloid/saponin metabolism cluster in two groups, and notably, the glycol-alkaloids and the saponins also cluster separately. Interestingly, it appears that the genes of the group #1 are expressed mainly in flower tissues where saponins accumulate, while the genes of the group #2 are expressed mainly in fruit tissues where glycolalkaloids accumulate (Figure 13).

7. GAME9 regulates steroidal alkaloid biosynthesis through activation of the cholesterol precursor pathway

Steroidal glycoalkaloids (SGAs) are cholesterol-derived defense compounds. We discovered that *GLYCOALKALOID METABOLISM 9* (*GAME9*), an APETALA2/Ethylene Response Factor, related to regulators of alkaloid production in tobacco and *Catharanthus roseus*, controls *Solanaceae* SGAs biosynthesis. *GAME9* knockdown and over-expression in tomato and potato altered expression of SGAs pathway genes and the recently reported cholesterol biosynthesis gene *STEROL SIDE CHAIN REDUCTASE 2* (*SSR2*). Levels of SGAs, C24-alkylsterols and the upstream mevalonate and cholesterol pathways intermediates were modified in these plants (Figure 14). A GAME9 direct target was functionally characterized and shown to act as a Δ(7)-STEROL-C5(6)-DESATURASE (C-5 SD) in the hitherto unresolved cholesterol pathway. GAME9 exerts its activity through additional interacting and/or downstream regulators besides directly controlling the *SSR2* and *C-5 SD* cholesterol precursor pathway genes. Our findings provide insight into the regulation of SGAs biosynthesis and means for manipulation of these metabolites in crops.

8. Deciphering SGA biosynthetic pathway in Solanaceae.

The work in the past year provided the bases for deciphering the metabolic pathway through which steroidal glycoalkaloids (SGAs) are being modified in the course of fruit development and ripening. We have likely identified a putative dioxygenase enzyme (GAME31) that catalyses the hydroxylation of tomatine, the major SGA in tomato. GAME31 conducts the first step in the modification of SGAs but the pathway to the ripening accumulating Esculeoside A requires several additional enzymes. GAME5, a glycosyltransferase gene showing a ripening-regulated expression pattern was identified in the course of the study as a likely candidate to perform the conversion of acetoxyhydroxytomatine to Esculeoside A. Apart from tomato, we have also been working with another member of the Solanaceae family that produces SGAs. In collaboration with the lab of Giuseppe Rotino in Italy we have prepared RNAseq libraries out of 77 different eggplant plant tissues and organs. The transcriptome analysis appeared successful and together with the metabolomics data (largely secondary metabolites) on the same samples we currently possess an excellent data set and infrastructure to start elucidating the biosynthesis of solmagrine and solasodine, the typical SGAs produced by eggplant that also provide its fruit with a bitter taste.

o Tables

Table 1. Evaluation of Desirée red tint following heat (H) treated and control (C). The color of the tubers was monitored using a portable spectrophotometer and evaluated in terms of the L*a*b* model (L*, brightness; a*, red tint; b*, yellow tint). Values are an average of around 20 tubers. Data were analyzed for statistical significance among means by Tukey HSD test; values followed by different letters are significantly different ($P < 0.05$).

	Tuber color		
	L*	a*	b*
C	46.63 B	18.5 A	13.83 B
H	48.08 A	**15.87 B**	15.4 A

Table 2. Expression profile of anthocyanin and SGA biosynthetic genes and regulators in the phelloderm of Desirée tubers exposed to heat treatment (H) and controls (C). Gene transcription was monitored by qPCR and the expression levels were determined relative to a reference gene *Nac*. Values are an average of 8-9 plants. Data were analyzed for statistical significance among means by Student's *t* test; values followed by different letters are significantly different ($P < 0.05$).

	Phelloderm	
	C	H
Anthocyanin biosynthetic genes		
CHS	1.50 A	0.43 B
CHI	1.27 A	1.49 A
F3H	1.80 A	0.66 B
F35H	2.07 A	0.65 B
DFR	0.99 A	0.41 B
ANS	1.57 A	0.89 B
3GT	1.87 A	0.98 B
RT	1.07 A	0.98 B
AN11	0.77 B	1.60 A
JAF13	1.01 B	1.78 A
vATPase	0.63 B	1.22 A
Glycoalkaloid biosynthetic genes		
HMG1	2.12 A	0.73 B
SQS1	1.30 A	1.43 A
CAS	1.37 A	0.95 A
GAME4	1.42 A	0.75 A
GAME8a	1.54 A	0.61 B
GAME11	1.56 A	0.33 B
GAME12	1.76 A	0.55 B
SGT1	1.98 A	0.80 B
SGT2	1.87 A	1.65 A
SGT3	1.55 A	0.40 B

Table 3. Level of anthocyanins, flavonols and phenolic acids (peak area/gFW /10^6) in the skin and phelloderm of Desirée tubers exposed to heat treatment (H) and controls (C) during experimental year 2013. Values are an average of 8-9 plants ±SE

	Phelloderm	
	C	H
Anthocyanins		
Pelargonidin	1320±167	1226±129
Peonidin	94±9	81±7
Phenolic acids		
Chlorogenic acid like	469±44	438±30
Caffeic acid	1356±96	1359±97
Coumaric acid	349±36	401±22
Ferulic acid like	457±38	526±26
Hydroxycinnamic acid	2506±204	2427±199
Total phenolic acid [including undefined]	5873±476	5968±394
Flavonol		
Kaempferol	67±5.4	54±4.5

Table 4. Level of SGA metabolites in the peel of Desirée tubers exposed to heat treatment for 11 days (H) and controls (C). Values are an average of 3 plants. Data (in columns) were analyzed for statistical significance among means by Student's *t*-test; values followed by different letters are significantly different ($P < 0.05$).

Treatment	SGA mg/1gFW		
	Solanine	Chaconine	Total SGA
C	52.94 A	118.81 A	171.75 A
H	37.88 A	85.48 B	123.36 B

o Figures

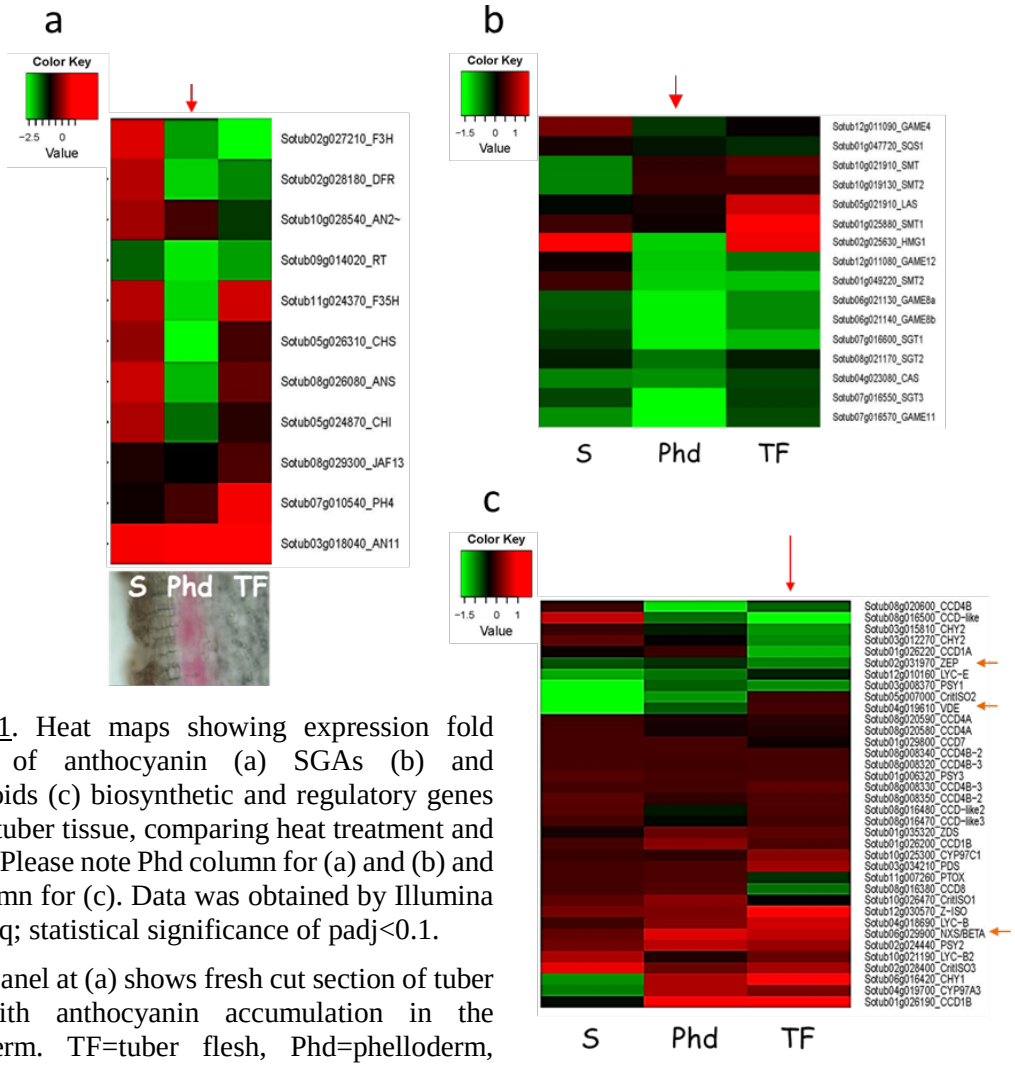


Figure 1. Heat maps showing expression fold change of anthocyanin (a) SGAs (b) and carotenoids (c) biosynthetic and regulatory genes in each tuber tissue, comparing heat treatment and control. Please note Phd column for (a) and (b) and TF column for (c). Data was obtained by Illumina RNA-seq; statistical significance of padj<0.1.

Lower panel at (a) shows fresh cut section of tuber peel with anthocyanin accumulation in the phelloderm. TF=tuber flesh, Phd=phelloderm, S=skin

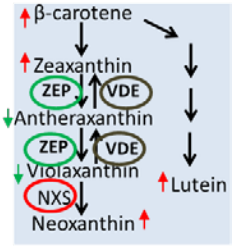


Figure 2. Schematic presentation of partial of the carotenoid biosynthetic pathway (enzymes, upper letters; metabolites, lower letters) in potato tuber flesh and the expected metabolite accumulation (red arrows) or reduction (green arrows) based on transcriptome data in Figure 3).

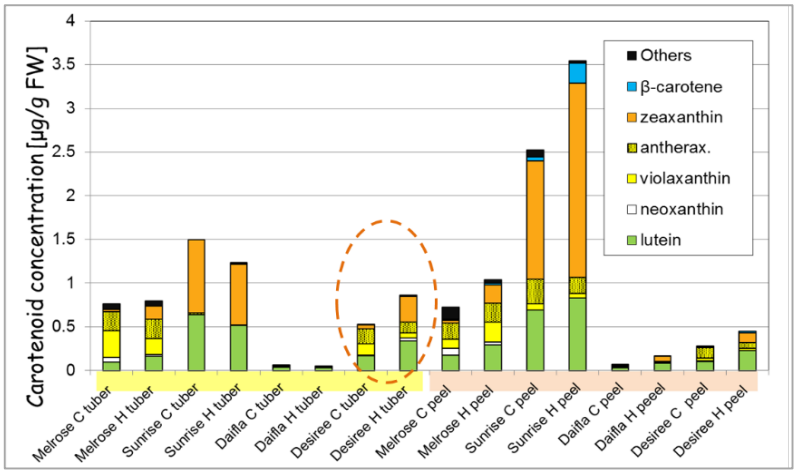


Figure 3. Carotenoid concentration in tuber flesh and peel of Melrose, and Andean Sunrise that are enriched with carotenoids, Desirée with moderate level and Daifla with no carotenoids, following heat treatment (H) and control (C). Bars indicate total carotenoids. Colored sections donate carotenoid types. Values are an average of three independent plants.

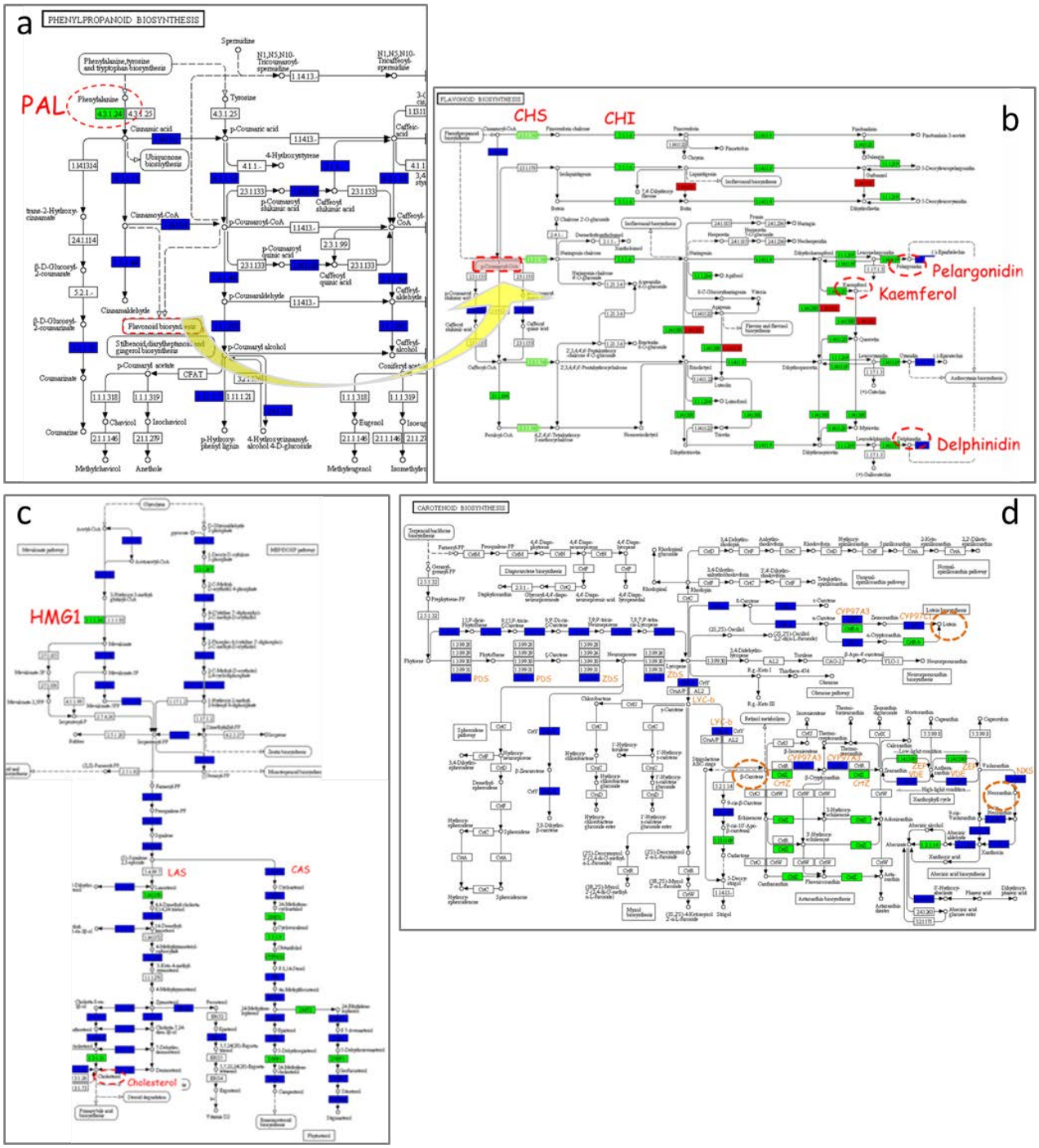


Figure 4. KEGG maps of phenylpropanoid (a), flavonoid/anthocyanin (b), mevalonate/ isoprenoid/phytosterol (c) and carotenoid (d) metabolic pathways. Potato tuber transcriptome data was aligned against KEGG database to produce the maps. Green or red boxes – indicate significant down- or up- regulation of gene expression, respectively, following heat treatment; blue boxes – no significant change; blanc boxes indicate no homology found. Gene or metabolite annotations that are relevant to the present project were written or encircled in red. Data show up-stream pathways are mainly in blue indicating little change in gene expression, while specific branch pathways are mainly green indicating heat-induced down-regulation.

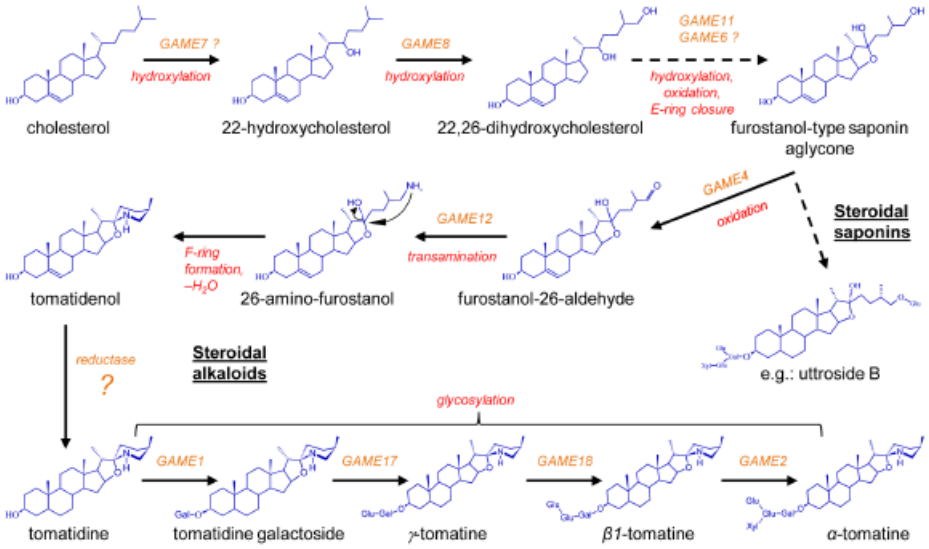


Figure 5. Biosynthesis of glycoalkaloids in Solanaceae. Suggested biosynthetic pathway from cholesterol towards α-tomatine. Dashed and solid arrows represent multiple or single enzymatic reactions in the pathway, respectively.

Figure 6. Invitro generation of α-tomatine from tomatidine with the 4 identified recombinant glycosyltransefrases. Tomatidine and the corresponding UDP-sugars were reacted in the same tube with the 4 recombinant GAME proteins and α-tomatine could be identified using its standard (lower chromatogram). LC-MS was used for the analysis.

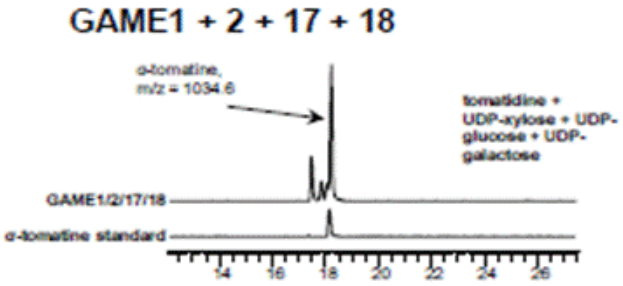
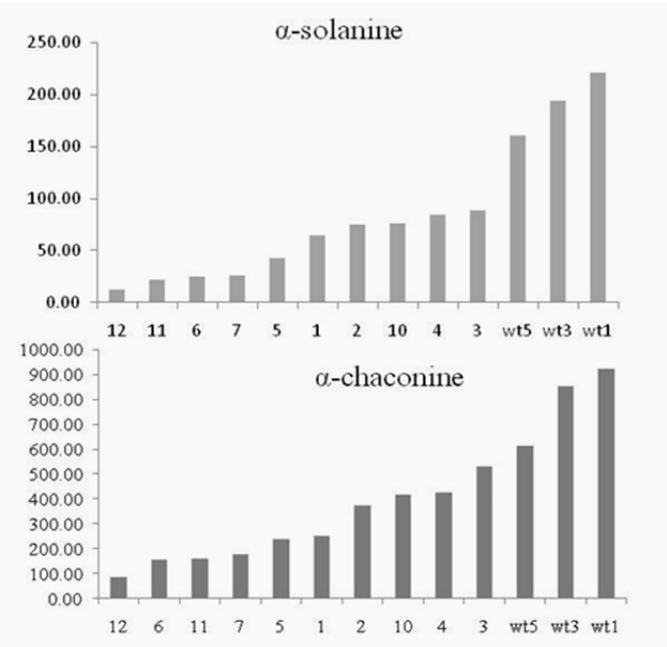


Figure 7. Reducing the levels of toxic glycoalkaloids in potato tubers. Tubers of transgenic lines (marked as numbers) silenced for the *GAME4* gene through RNA interference display dramatically reduced levels as compared to wild type (wt 1-3) tubers. The analysis of chaconine and solanine was performed by LC-MS.



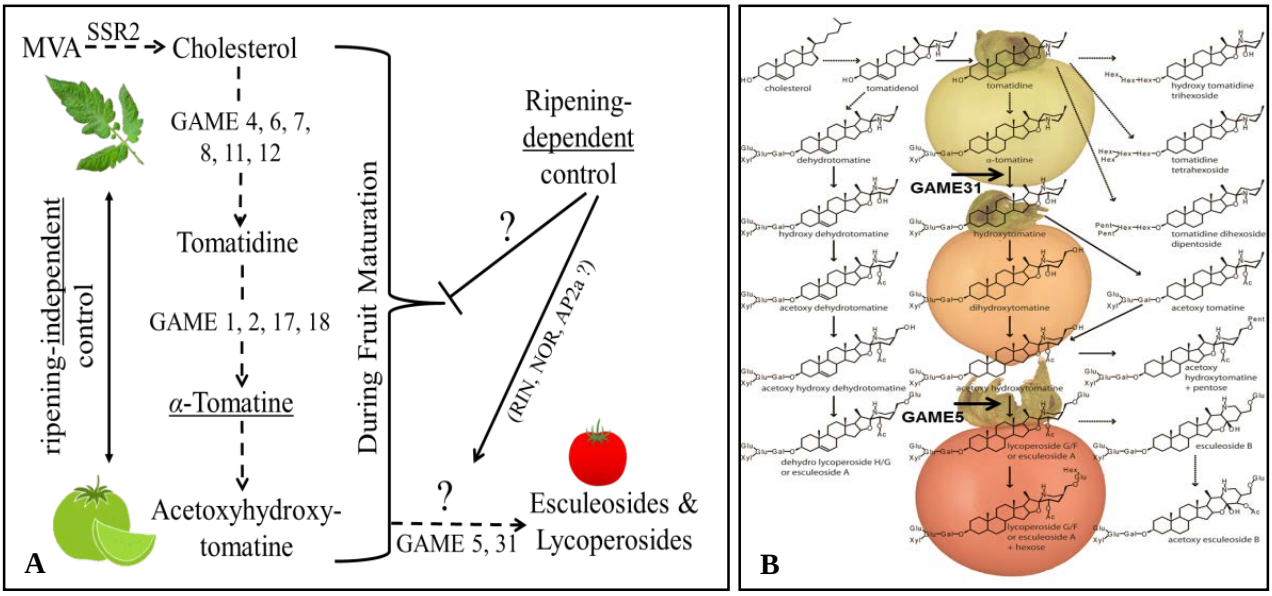


Figure 8. Known and predicted pathways, structural and regulatory genes involved in the biosynthesis of modified steroidal glycoalkaloids (SGAs). In **(A)**, the biosynthetic pathway leading to the formation of modified α-Tomatine metabolites (i.e. esculeosides and lycoperosides) consists of 3 main parts including (i) through the mevalonic acid pathway (MVA) to cholesterol, (ii) from cholesterol to the aglycone tomatidine and its glycosylation at the C_3 position to form α-Tomatine , and (iii) modification of α-Tomatine to form esculeosides and lycoperosides. The pathway up to acetoxy-hydroxytomatine (intermediate detected in leaves and green fruit) appears to be independent of the fruit ripening control machinery while ripening regulators such as RIN, NOR and AP2a might be those directly or indirectly controlling the steps downstream of acetoxytomatine towards the esculeosides and lycoperosides in late fruit maturation. These transcription factors might be controlling GAME5 and GAME31 that are planned to be characterized in the project. Our working hypothesis is that also late in fruit maturation, the repression of genes associated with cholesterol to α-Tomatine up to acetoxy-hydroxytomatine is mediated by the ripening machinery. In **(B)**, modification of α-Tomatine involves multiple yet undescribed enzymes; two primary candidates GAME5 and GAME31, representing a glycosyltransferase and a dioxygenase, were discovered and characterized in the course of the project. The pathway was predicted by our lab according to metabolic profiling data along fruit development. The intermediates are at maximum level in the indicated fruit stages but are still present in other stages albeit in reduced levels.

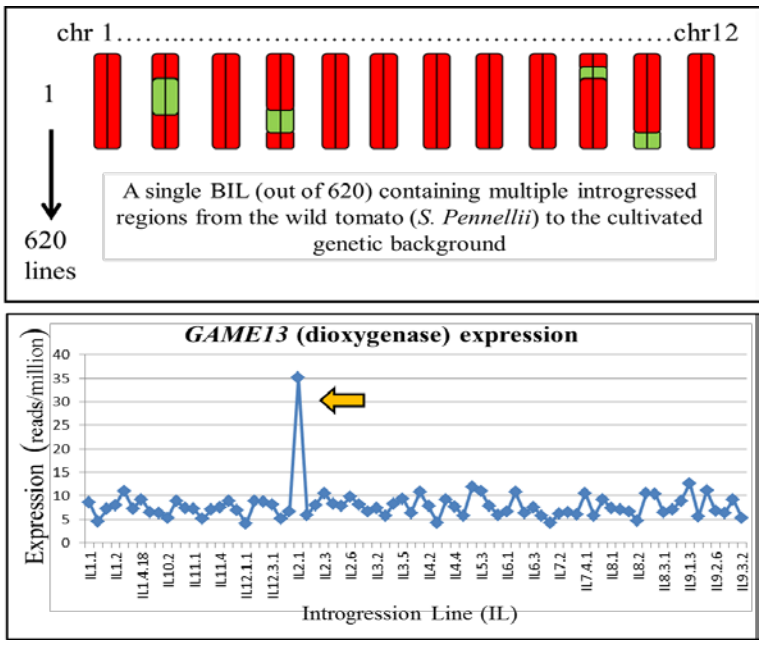


Figure 9. Identification of the *GAME31* (dioxygenase) candidate using a tomato backcross introgression lines (BILs) population to unravel the *α*-Tomatine to esculeoside and lycoperosides pathway**.** In the upper panel, an example of an hypothetical line in the BIL population carrying multiple introgressed segments is presented. Red and green colors represent the cultivated and wild tomato genome segments, respectively. Bottom panel; expression one of the tandem 4 dioxygenases is dramatically elevated in the IL2.1 containing the introgressed region impacting hydroxyl- and acetoxy- tomatine levels in leaves.

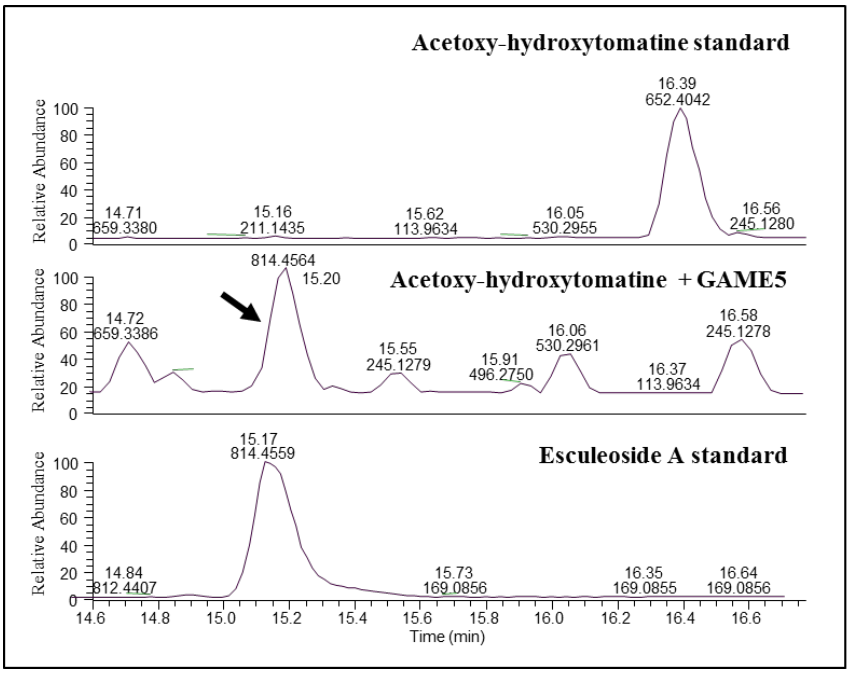


Figure 10**.** The recombinant GAME5 produced in *E.coli* cells glycosylates acetoxy-hydroxytomatine and forms Esculeoside A *in vitro*. Esculeoside A (arrow) is formed in the reaction containing the GAME5 recombinant enzyme as well as the acetoxy-hydroxytomatine substrate. The retention time and spectra of the Esculeoside A in the reaction matches the one of the standard (lower chromatogram).

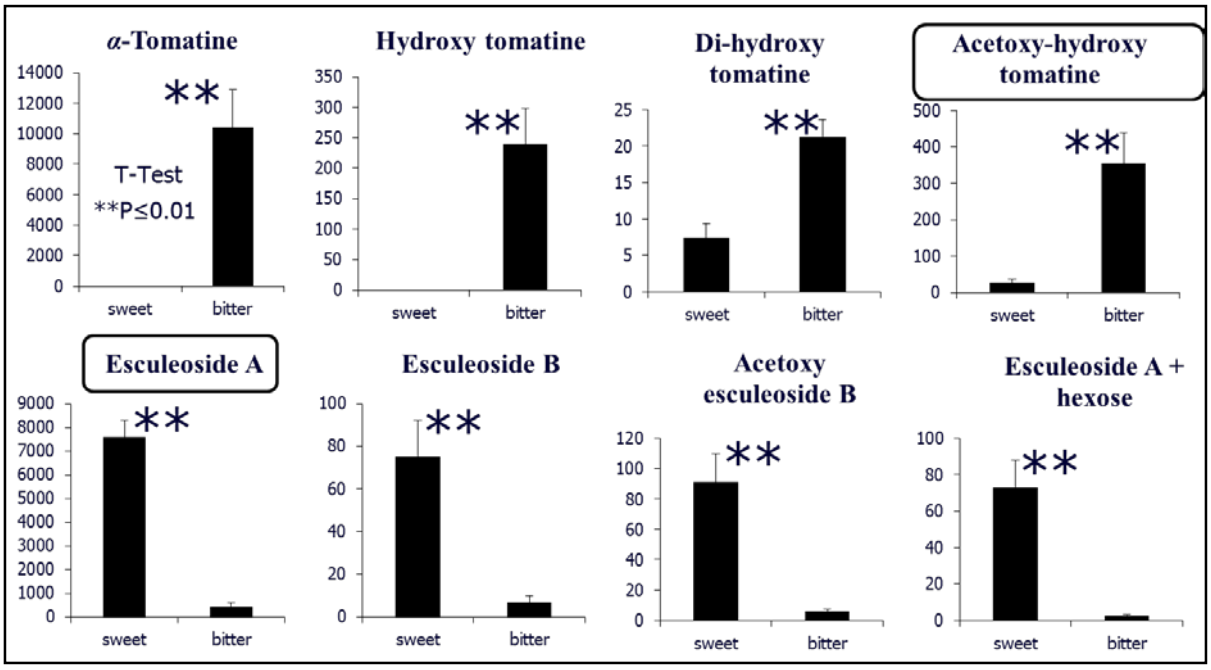


Figure 11**.** Metabolic analysis of the 'bitter' accession red fruit demonstrates the block point in *α*-Tomatine catabolism**.** Fruit of the bitter and a sweet accession at the red stage were analyzed for intermediates in the *α*-Tomatine catabolic pathway. The bitter fruit accumulates *α*-Tomatine and metabolic intermediates up to and including Acetoxy-hydroxytomatine. The reaction in which Acetoxy-hydroxytomatine is converted to Esculeoside A is likely blocked (metabolite names boxed). Esculeoside A and downstream esculeosides (and lycoperosides, not shown) hardly accumulate in the bitter fruit accession. Samples were obtained from a red stage fruit and analyzed by high-resolution LC-MS.

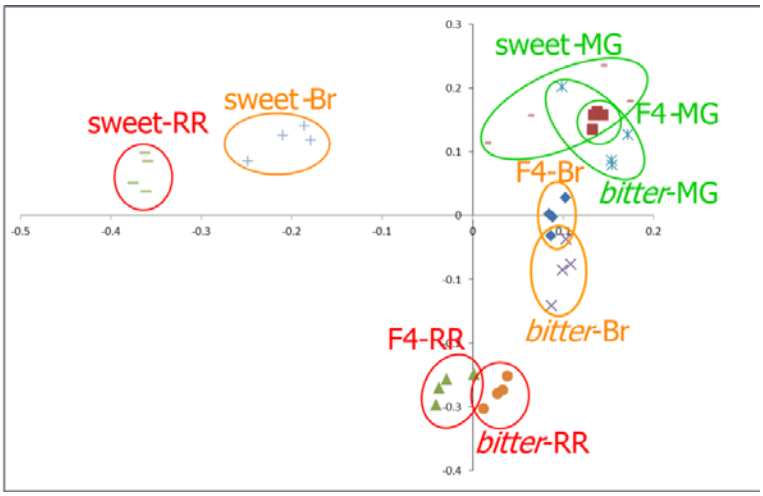


Figure 12. Principal component analysis demonstrated that the 'bitter' phenotype is associated with mid to late stages of fruit development and ripening**.** Mass signals profiles were obtained by non-targeted LC-MS analysis of peel tissue derived from bitter fruit of the original bitter accession and F4 fruit (of the cross bitter x sweet accessions with bitter flavor) and sweet fruit. The analysis demonstrated that the 'bitter' phenotype is associated with mid to late stages of fruit development and ripening. Note that breaker and red stages were well separated between bitter and sweet samples while samples from the mature green stage (MG) were not separated by the PCA. Samples were analyzed by high-resolution LC-MS (n = 4).

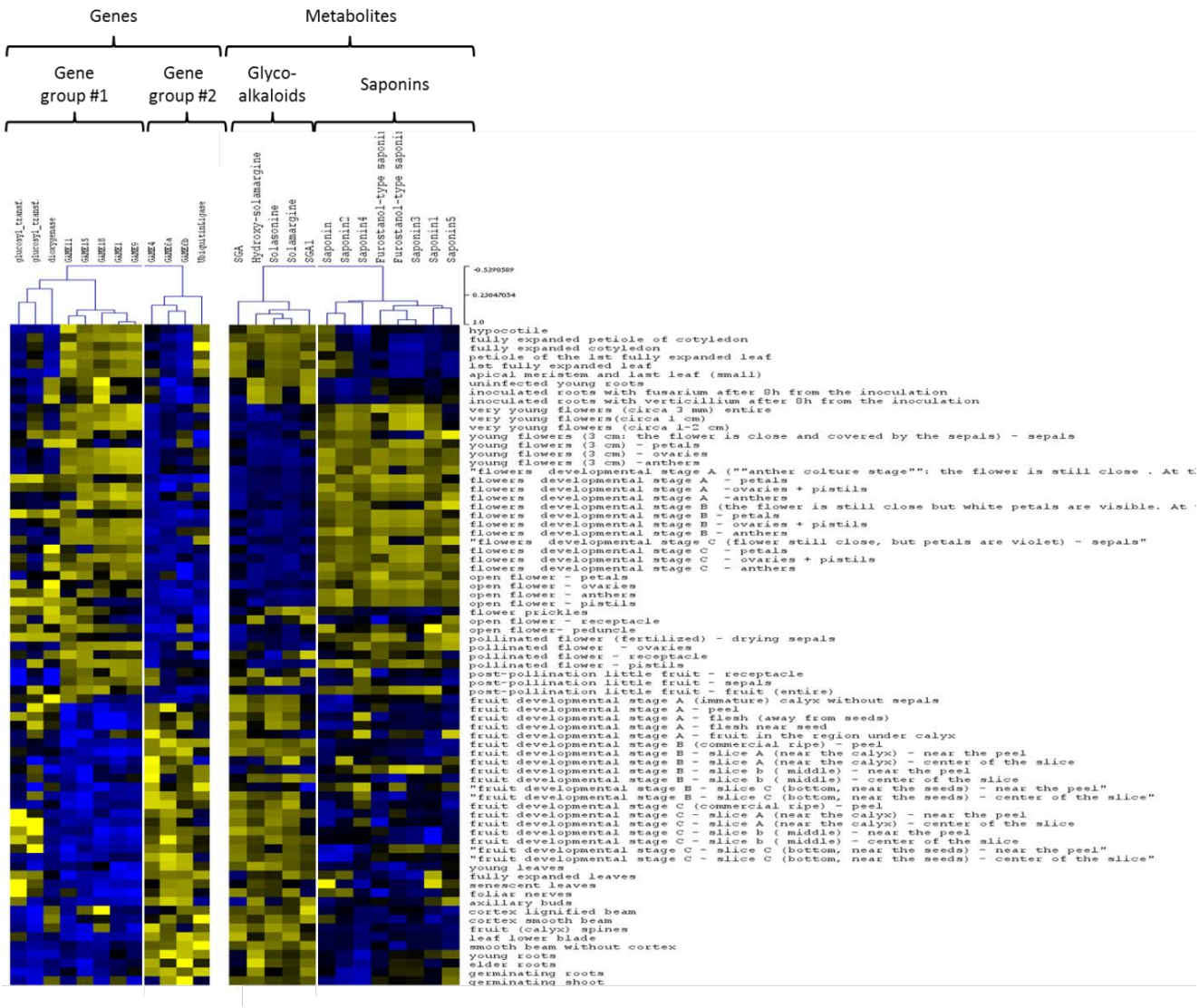


Figure 13: Heat map representing the hierarchical clustering of expression level of the genes previously identified as involved in glycol-alkaloid/saponin metabolism, and the abundance of glycol-alkaloid/saponin metabolites obtained by UPLC-qTOF-MS analysis Yellow, higher gene expression or metabolite abundance, blue, lower gene expression or metabolite abundance

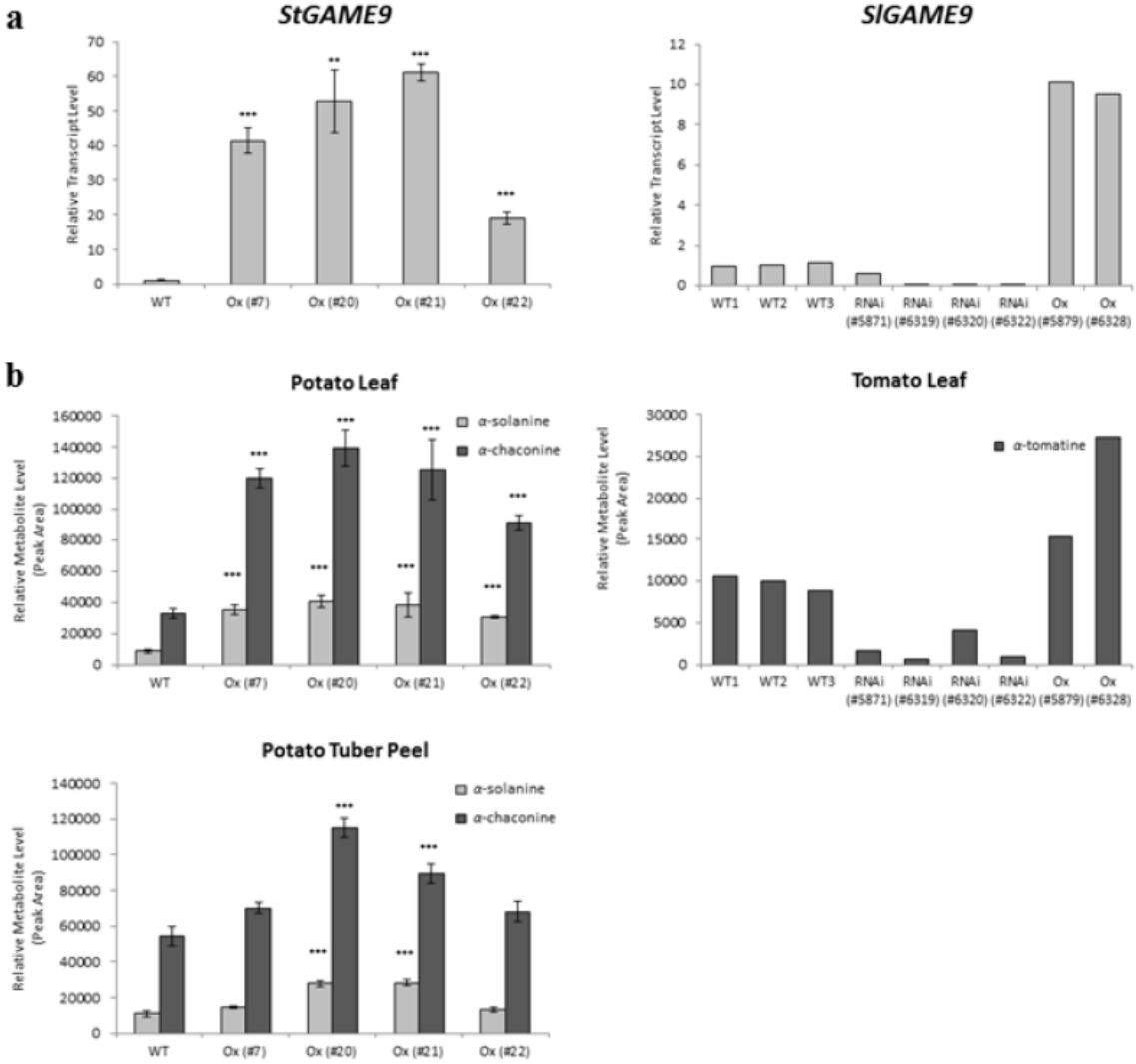


Figure 14. Changes in *GAME9* expression in potato and tomato results in altered levels of the predominant SGAs.

(a) *GAME9* gene expression (qRT-PCR) in *GAME9*-Ox and *GAME9*-RNAi lines in potato and tomato.

(b) Levels of *α*-solanine and *α*-chaconine in leaves and peel of potato tubers of *GAME9*-Ox lines and levels of *α*-tomatine in leaves of *GAME9*-RNAi and *GAME9*-Ox tomato lines.

Values represent means ± standard errors (n=3). Student's *t*-test was used to assess whether the transgenic lines significantly differ from wild-type plants: ***P value < 0.001.

- Details of cooperation

Within the planned collaboration with the Italian group from CRA-CIN, potato cultivars enriched with carotenoids including the Melrose that was developed by Dr. Mandolino, were imported to Israel and were included in the experimental system.

The ARO-Volcani/HUJI groups worked cooperatively to fulfill the objectives of the project. Dr. Ginzberg grew the potato cultivars under controlled conditions in a greenhouse, sampled the specific tuber tissues and delivered them to the Israeli collaborators for the required analyses. Expression analysis of anthocyanin and glycoalkaloids biosynthetic genes was performed by the group of Dr. Ginzberg, including new discovered SGA genes by the group of Aharoni within the frame of the present project. Carotenoid analysis was performed by Prof. Hirschberg and anthocyanin pigment profiling was done by the group of Dr. Oren-Shamir.

We have received seeds of transgenic eggplant plants that possess RNAI of GAME 4. These plants are now growing and will be examined for changes in SGAs. In addition, the Rotino lab sent us samples of 76 tissues of eggplant from which we have isolated RNA and will perform transcriptome analysis with. The Rotino lab obtains information from the Aharoni lab regarding new SGAs genes identified in tomato and potato so they can be investigated in eggplant.

In addition, the Israeli and the Italian collaborators met each year to summarize the yearly achievements and to coordinate the next experimental year. Meetings in 2012 and 2014 took place in Israel (Hebrew University of Jerusalem and ARO, the Volcani Center, respectively), and in 2013 and 2015 in Italy (CRA-Rome and CRA-Milano, respectively).

- Publications

- M. Itkin, U. Heinig, O. Tzfadia, A. Bhide, Y. Chikate, A. Bovy, S. Malitsky, P. Singh, I. Rogachev, J. Beekwilder, A. P. Giri and A. Aharoni (2013). Biosynthesis of Anti-nutritional Glycoalkaloids in Solanaceous Crops is Mediated by Clustered Pathway Genes in Duplicated Genomic Regions. *Science*, 341, 175 - 199.

- Uwe Heinig and Asaph Aharoni (2013). Analysis of Steroidal Alkaloids and Saponins in Solanaceae Plant Extracts using UPLC-qTOF Mass Spectrometry. In: Plant Isoprenoids, in the series Methods in Molecular Biology, published by Springer USA.

- Cárdenas PD, Sonawane PD, Bocobza SE, Aharoni A (2014). The Bitter Side of the Nightshades: Steroidal Alkaloids Metabolism in Solanaceae. Phytochemistry, *Phytochemistry*, 113, 24-32.

- Pablo D. Cárdenas, Prashant D. Sonawane, Jacob Pollier, Robin Vanden Bossche, Efrat Weithorn, Lior Tal, Sagit Meir, Ilana Rogachev, Sergey Malitsky, Ashok P. Giri, Alain Goossens, Saul Burdman, Asaph Aharoni (2015). GAME9 Regulates Steroidal Alkaloid Biosynthesis through Activation of the Cholesterol Precursor Pathway. *In revision.*

- **סיכום עם שאלות מנחות**

נא להתייחס לכל השאלות בקצרה ולעניין, ב-3 עד 4 שורות לכל שאלה (לא תובא בחשבון חריגה מגבולות המסגרת המודפסת).

שיתוף הפעולה שלך יסייע לתהליך ההערכה של תוצאות המחקר.

הערה: נא לציין הפנייה לדו"ח אם נכללו בו נקודות נוספות לאלה שבסיכום.

מטרות המחקר תוך התייחסות לתוכנית העבודה.
תפו"א הינו אחד מהמזונות הבסיסיים, עשיר בחלבונים, סיבים תזונתיים, ויטמינים ומינרלים. תכניות השבחה של זנים חדשים מכוונות לשיפור הערך הבריאותי של הפקעת באמצעות העלאה של קרוטנואידים ואנתוציאנינים, בעלי פעילות אנטי-חימצונית, והפחתה של גליקואלקלואידים להם פעילות שלילית על הבריאות. רמות מטבוליטים אלו ייחודיות לזן, אולם גורמים התפתחותיים, רקמתיים ותנאי גידול קובעים את הרמה הסופית שלהם ברקמה. המחקר הנוכחי הינו שיתוף פעולה בין קבוצות מחקר באיטליה ובישראל ומטרתו לנצל גישות גנומיות על מנת לגלות את המנגנונים שמבקרים את רמת החומרים האלה מה שיאפשר לפתח זני תפו"א בעלי ערך בריאותי גבוה.
עיקרי התוצאות.
השפעת טמפרטורות קרקע על ביטוי גנים במסלול המטבוליזם של אנתוציאנינים, גליקואלקלואידים וקרוטנואידים בפקעות תפו"א נבחנה באמצעות q-PCR ובאנליזת טרנסקריפטום (Illumina). תוצאות הראו כי טיפול החום הוביל לירידה מובהקת ברמת ביטוי הגנים בשני המסלולים הראשונים, בעוד שעבור מסלול הקרוטנואידים התגובה מורכבת יותר. אנליזת מטבוליטים הראתה ירידה תואמת ברמת הגליקואלקלואידים ותגובה תלוית זן ורקמה עבור קרוטנואידים. בנוסף, נבחנה רמת הגליקואלקלואידים בעלים של אוכלוסיית אינטרוגרסיה של עגבניה עבור מקטע המשפיע על רמות גליקואלקלואידים, ובודדו שני גני המעורבים במודיפיקציה של Esculeoside A , הגליקואלקלואידים העיקרי בעגבניה. וכן נערכה אנליזה מטבולית וטרנסקריפטומית ל- 77 רקמות של חציל.
מסקנות מדעיות וההשלכות לגבי יישום המחקר והמשכו. האם הושגו מטרות המחקר לתקופת הדו"ח?
הושגו מטרות המחקר, אולם יש להשלים אנליזה של אנתוציאנינים.
בעיות שנותרו לפתרון ו/או שינויים (טכנולוגיים, שיווקיים ואחרים) שחלו במהלך העבודה; התייחסות המשך המחקר לגביהן, האם יושגו מטרות המחקר בתקופה שנותרה לביצוע תוכנית המחקר?
לא חלו שינויים בתכנית המחקר
הפצת הידע שנוצר בתקופת הדו"ח: פרסומים בכתב - ציטט ביבליוגרפי כמקובל בפרסום מאמר מדעי; פטנטים - יש לציין שם ומס' פטנט; הרצאות וימי עיון - יש לפרט מקום, תאריך, ציטוט ביבליוגרפי של התקציר כמקובל בפרסום מאמר מדעי.
M. Itkin, U. Heinig, O. Tzfadia, A. Bhide, Y. Chikate, A. Bovy, S. Malitsky, P. Singh, I. Rogachev, J. Beekwilder, A. P. Giri and A. Aharoni (2013). Biosynthesis of Anti-nutritional Glycoalkaloids in Solanaceous Crops is Mediated by Clustered Pathway Genes in Duplicated Genomic Regions. Science, 341, 175 - 199. Uwe Heinig and Asaph Aharoni (2013). Analysis of Steroidal Alkaloids and Saponins in Solanaceae Plant Extracts using UPLC-qTOF Mass Spectrometry. In: Plant Isoprenoids, in the series Methods in Molecular Biology, published by Springer USA. Cárdenas PD, Sonawane PD, Bocobza SE, Aharoni A (2014). The Bitter Side of the Nightshades: Steroidal Alkaloids Metabolism in Solanaceae. Phytochemistry, 113, 24-32. Pablo D. Cárdenas, Prashant D. Sonawane, Jacob Pollier, Robin Vanden Bossche, Efrat Weithorn, Lior Tal, Sagit Meir, Ilana Rogachev, Sergey Malitsky, Ashok P. Giri, Alain Goossens, Saul Burdman, Asaph Aharoni (2015). GAME9 Regulates Steroidal Alkaloid Biosynthesis through Activation of the Cholesterol Precursor Pathway. In revision.
פרסום הדו"ח: אני ממליץ לפרסם את הדו"ח: (סמן אחת מהאופציות)
➢ללא הגבלה (בספריות ובאינטרנט)
➢חסוי – לא לפרסום: יש לצרף אישור ומידע ממוסד המחקר X
האם בכוונתך להגיש תוכנית המשך בתום תקופת המחקר הנוכחי? *
*יש לענות על שאלה זו רק בדו"ח שנה ראשונה במחקר שאושר לשנתיים, או בדו"ח שנה שניה במחקר שאושר לשלוש שנים

Sciencexpress Reports

Biosynthesis of Antinutritional Alkaloids in Solanaceous Crops Is Mediated by Clustered Genes

M. Itkin,[1]* U. Heinig,[1] O. Tzfadia,[1] A. J. Bhide,[1,2] B. Shinde,[1,2] P. Cardenas,[1] S. E. Bocobza,[1] T. Unger,[4] S. Malitsky,[1] R. Finkers,[3] Y. Tikunov,[3] A. Bovy,[3] Y. Chikate,[1,2] P. Singh,[1,2] I. Rogachev,[1] J. Beekwilder,[3] A. P. Giri,[1,2] A. Aharoni[1]†

[1]Department of Plant Sciences, Weizmann Institute of Science, Rehovot 76100, Israel. [2]Plant Molecular Biology Unit, Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Pune 411008, MS, India. [3]Plant Research International, Wageningen University and Research Centre, POB 16, 6700 AA, Wageningen, Netherlands. [4]Israel Structural Proteomics Center, Weizmann Institute of Science, Rehovot 76100, Israel.

*Present address: Agricultural Research Organization, The Volcani Center, Bet Dagan 50250, Israel.

†Corresponding author. E-mail: asaph.aharoni@weizmann.ac.il

Steroidal glycoalkaloids (SGAs) such as α-solanine found in Solanaceous food plants—as, for example, potato—are antinutritional factors for humans. Comparative coexpression analysis between tomato and potato coupled to chemical profiling revealed an array of 10 genes that partake in SGA biosynthesis. We discovered that six of them exist as a cluster on chromosome 7 while an additional two are adjacent in a duplicated genomic region on chromosome 12. Following systematic functional analysis, we suggest a revised SGA biosynthetic pathway starting from cholesterol up to the tetrasaccharide moiety linked to the tomato SGA aglycone. Silencing *GLYCOALKALOID METABOLISM 4* prevented accumulation of SGAs in potato tubers and tomato fruit. This may provide a means for removal of unsafe, antinutritional substances present in these widely used food crops.

Our demand for more and better food continues to increase. Improved nutritional qualities as well as removal of antinutritional traits are both needed. Various approaches have been used to add nutritional qualities to food crops. We focus here on reducing the level of endogenous, antinutritional factors in existing crops (*1*). Antinutritional substances range from lethal toxins to compounds that disrupt digestion and nutrient absorption (*2*). In the course of crop domestication, levels of antinutrients were reduced by selection and/or breeding, although some of such substances remain in the general food source. In addition, wild germplasm, which can be useful as a source of novel traits such as pathogen resistance, may also be complicated by co-occurrence of antinutritional compounds that need to be removed. Current technologies include extensive backcrossing, which can be a slow and imperfect process (*3*).

Steroidal glycoalkaloids (SGAs), found in staple vegetable crops such as potato (*Solanum tuberosum*) and tomato (*S. lycopersicum*), are a class of antinutritional substances that remain in our food chain and daily diet (*4*). The glycoalkaloids α-solanine (*5*) and α-chaconine are the principle toxic substances in potato. These SGAs cause gastrointestinal and neurological disorders and, at high concentrations, may be lethal to humans. Mechanisms of toxicity include disruption of membranes and inhibition of acetylcholine esterase activity (*6*). For this reason, total SGA levels exceeding 200 mg per kilogram fresh weight of edible tuber are deemed unsafe for human consumption (*7*). SGA biosynthesis requires genes encoding UDP-glycosyltransferases (UGTs) that decorate the steroidal alkaloid (SA) skeleton with various sugar moieties (*8*, *9*). The tomato GLYCOALKALOID METABOLISM 1 (GAME1) glycosyltransferase, a homolog of the potato SGT1 (*8*), catalyzes galactosylation of the alkamine tomatidine (*9*). Cholesterol is the proposed common precursor for biosynthesis of both steroidal alkaloids (SAs) and non-nitrogenous steroidal saponins (STSs) (Fig. 1 and fig. S1) (*10*). Conversion of cholesterol to the alkamine SA should require several hydroxylation, oxidation and transamination reactions (*10*). Here, we identify genes encoding enzymes performing the conversion of cholesterol to SGAs and use them to engineer *Solanaceae* plants with reduced SGA content.

To discover genes associated with SGA biosynthesis, we carried out coexpression analysis using transcriptome data from tomato and potato plants (*11*). Sixteen genes from each species were coexpressed with *GAME1*/*SGT1* (Fig. 2). One of these genes, which we named *GLYCOALKALOID METABOLISM 4* (*GAME4*), encodes a member of the 88D subfamily of cytochrome P450 proteins (fig. S2). *GAME4* and *GAME1/SGT1* display a very similar expression profile in tomato and potato (fig. S3, B and C, and fig. S4). We then discovered that the *GAME1/SGT1* and *GAME4* genes in tomato and potato are positioned in chromosomes 7 and 12, respectively, such that they are physically next to several of their coexpressed genes (Fig. 3).

A cluster of *GAME1*/*SGT1* coexpressed genes spans a ~200 Kbp genomic region on chromosome seven. Together with *GAME1*, the tomato cluster is composed of 7 coexpressed genes. These include, 3 UDP-glycosyltransferases [*GAME2* (termed *SGT3 in potato*); *GAME17* and *GAME18*], a cytochrome P450 of the 72A subfamily (*GAME6*), a 2-oxoglutarate–dependent dioxygenase (*GAME11*), and a cellulose synthase-like protein. It appears that in potato this cluster contains 5 coexpressed genes as it lacks homologs of the tomato GAME17 and GAME18 UDP-glycosyltransferases. We performed enzyme activity assays with the four recombinant clustered tomato UDP-glycosyltransferases. GAME17 and GAME18 exhibited UDP-glucosyltransferase activity when incubated with tomatidine galactoside (T-Gal) and γ-tomatine (T-Gal-Glu) as a substrate, respectively, whereas GAME2 was shown to have an UDP-xylosyltransferase activity when incubated with *β1*-tomatine (T-Gal-Glu-Glu) as a substrate (Fig. 4, F to H, and fig. S5). GAME1 was previously shown to act as a tomatidine UDP-galactosyltransferase in tomato (*9*). When incubating the 4 recombinant UGT enzymes in a single test tube, with tomatidine, and all glycoside donors (UDP -galactose, -glucose and -xylose), we observed the accumulation of the final SGA product α-tomatine (Fig. 4I and fig. S5). The role of *GAME18* in creating the tetrasaccharide moiety of α-tomatine was supported by Virus Induced Gene Silencing (VIGS) assays as *GAME18*-silenced fruit accumulated γ-tomatine which was not present in the control sample (fig. S6, A to E). Analysis of the VIGS-silenced leaves of GAME11, a putative dioxygenase in the cluster, revealed a significant reduction in α-tomatine levels and accumulation of several cholestanol-type steroidal saponins, confirming its function in the SGA pathway (Fig. 4B and fig. S6, F to I). Additionally, GAME6,

encoded by another cluster gene, was previously suggested to be associated with SGA metabolism (*12*).

GAME4 and a putative transaminase (*GAME12*) that was highly coexpressed were positioned alongside on chromosome 12 of both species (Fig. 3). Silencing *GAME4* in potato by RNAi (*GAME4*i plants), showed up to 74 times reduction in the levels of α-solanine/chaconine and other SGAs in both leaves and tubers (fig. S7, A to E). In the dark, normal quantities of α-solanine and α-chaconine are 200 and 370 mg/kg, respectively (fig. S7C). After light exposure, levels of α-solanine and α-chaconine increase in tuber skin and quantities are 510 and 870 mg/kg, respectively. With the *GAME4* gene silenced, the concentrations of both α-solanine and α-chaconine remained below 5 mg/kg and did not change with light exposure (fig. S7, C to E).

In the domesticated tomato, the dominant SGA in leaves and mature green fruit is α-tomatine that was reduced ~100-fold in *GAME4i* plants (figs. S7F and S8, and table S14). During the transition from green to red fruit, α-tomatine is converted to lycoperosides and esculeosides. These two classes of compounds represent hydroxylated, glycosylated, and often acetylated α-tomatine derivatives (*13*). Hence, reduced α-tomatine accumulation in the green fruit stage resulted in reduced accumulation of lycoperosides and esculeosides in the red-ripe fruit stage (fig. S7G). Complementary results were obtained in *GAME4*-overexpressing tomato plant leaves (*GAME4*oe) as they accumulated 2.5 times more α-tomatine than the controls (fig. S8B). Furthermore, *GAME4*oe red-ripe fruit exhibited 2.9 times more esculeoside A (fig. S8C), demonstrating once more the central role of GAME4 in SGAs biosynthesis. It appeared that SGA precursors [i.e., cholesterol, cycloartenol and (S)-2,3-oxidosqualene] and the phytosterols campesterol and β-sitosterol, accumulated in leaves of *GAME4*-silenced tomato plants (fig. S9). Despite altered phytosterol levels, *GAME4*-silenced plants were not affected in their morphology under the conditions examined in this study (*14*).
Tomato and potato *GAME4*i plants with decreased levels of SGAs accumulated nitrogen-lacking compounds identified as steroidal saponins (STSs) (fig. S7, H and I, and fig. S10). Greater reduction in SGAs correlated with greater accumulation of STSs (fig. S7, D, E, H and I). Levels of STSs were significantly induced by light in several wild-type and *GAME4*i lines examined (fig. S7, H and I). These results indicate that SGAs and STSs originate from the same precursor, and that GAME4 is positioned in a branch point prior to the incorporation of nitrogen for SGA generation in the diverging biosynthetic pathways that produce these two classes of steroidal compounds (Fig. 1 and fig. S1).

GAME12 (transaminase) silenced tomato leaves were found enriched with a furostanol-type saponin (Fig. 4D and fig. S6, J to M), suggesting additional hydroxylation of its accumulated substrate. We also functionally examined genes that were tightly coexpressed and positioned elsewhere in the genome that belong to the CYP72 subfamily of cytochrome P450s (i.e., GAME7 and GAME8). GAME7 was coexpressed in both species while StGAME8a and StGAME8b were strongly coexpressed with *StSGT1* and *StGAME4* in potato. At present, we could not demonstrate SGA-related activity for GAME7 although as for GAME6 it was suggested to be involved in SGA metabolism (*12*). Yet, *GAME8*-silenced tomato leaves accumulated 22-(*R*)-hydroxycholesterol (fig. S6N to S6Q), a proposed intermediate in the SGA biosynthetic pathway (Fig. 1).

The above findings allowed us to propose a pathway from cholesterol to α-tomatine. Cholesterol is hydroxylated at C22 by GAME7 (*12*) followed by GAME8 hydroxylation at the C26 position (Fig. 1). The 22,26-dihydroxycholesterol is than hydroxylated at C16 and oxidized at C22 followed by closure of the *E*-ring by GAME11 and GAME6 to form the furostanol-type aglycone. This order of reactions is supported by the accumulation of cholestanol-type saponins, lacking hydroxylation at C16 and the hemi-acetal *E*-ring when silencing *GAME11* (Fig. S6, F to I). The furostanol-intermediate is oxidized by GAME4 to its 26-aldehyde which is the substrate for transamination catalyzed by GAME12. Nucleophilic attack of the amino-nitrogen at C22 leads to the formation of tomatidenol which is dehydrogenated to tomatidine. Tomatidine is subsequently converted by GAME1 to T-Gal (*9*). T-Gal in its turn is glucosylated by GAME17 into γ-tomatine, which is further glucosylated by GAME18 to *β1*-tomatine that is finally converted to α-tomatine by GAME2 (Fig. 1).

Some specialized plant metabolites, particularly terpenoids, are the result of activities from clusters of genes (*15*, *16*). The existence of metabolic gene clusters raises questions regarding the advantages of such genomic organization (*17*). Reducing the distance between loci, resulting in coinheritance of advantageous combinations of alleles, may be one benefit of clustering (*17*). Clustering glycosyltransferases and core pathway genes, as observed here for SGAs, could maintain allelic combinations that support the metabolic outcome needed by the plant and reduce formation of phytotoxic, aglycone, compounds (*9*, *18*). We found that the regions of coexpressed genes in both chromosomes (i.e., 7 and 12) were flanked by similarly annotated genes and positioned identically along the genome, although poorly coexpressed with *GAME1*/*SGT1* and *GAME4* and likely not related to SGAs metabolism (fig. S11 and table S13). This suggests a duplication event that facilitated the positioning alongside on chromosome 12 of *GAME4* and *GAME12*, both STSs-SGAs branch point genes. Subsequent evolution of enzyme function of this gene pair likely allowed plants in the *Solanaceae* family to start producing the nitrogen containing steroidal alkaloids.

We have shown that SGA levels can be severely reduced in potato tubers by modifying expression of an enzyme in the biosynthetic pathway. The lack of SGAs in such plants might make them sensitive to biotic stress and the increased production of STSs (as occurred in *GAME4*-silenced plants), which are non-toxic to warm-blooded species, including humans (*19*), might provide a compensatory defense mechanism (*20*). The findings open the way for developing new strategies, through genetic engineering or more classical breeding programs, to reduce quantities of the antinutritional SGAs in key crops of the *Solanaceae* including potato, tomato, and eggplant. At the same time, it provides a platform for studying the SGA and STS biosynthetic pathways, transport and regulatory systems that control the production of thousands of these chemicals in specific plant lineages.

References and Notes

1. N. N. Narayanan, U. Ihemere, C. Ellery, R. T. Sayre, Overexpression of hydroxynitrile lyase in cassava roots elevates protein and free amino acids while reducing residual cyanogen levels. *PLoS ONE* **6**, e21996 (2011). doi:10.1371/journal.pone.0021996 Medline
2. L. C. Dolan, R. A. Matulka, G. A. Burdock, Naturally occurring food toxins. *Toxins* **2**, 2289–2332 (2010). doi:10.3390/toxins2092289 Medline
3. L. L. Sanford, S. P. Kowalski, C. M. Ronning, K. L. Deahl, Leptines and other glycoalkaloids in tetraploid *Solanum tuberosum* x *Solanum chacoense* F2 hybrid and backcross families. *Am. J. Potato Res.* **75**, 167–172 (1998). doi:10.1007/BF02853568
4. M. Friedman, Potato glycoalkaloids and metabolites: Roles in the plant and in the diet. *J. Agric. Food Chem.* **54**, 8655–8681 (2006). doi:10.1021/jf061471t Medline
5. M. Desfosses, Extrait d'une lettre à M. Robiquet. *J. Pharmacie* **6**, 374–376 (1820).
6. J. G. Roddick, The acetylcholinesterase-inhibitory activity of steroidal glycoalkaloids and their aglycones. *Phytochemistry* **28**, 2631–2634 (1989). doi:10.1016/S0031-9422(00)98055-5
7. FDA Poisonous Plant Database, http://www.accessdata.fda.gov/scripts/Plantox/
8. K. F. McCue, L. V. T. Shepherd, P. V. Allen, M. M. Maccree, D. R. Rockhold, D. L. Corsini, H. V. Davies, W. R. Belknap, Metabolic compensation of steroidal glycoalkaloid biosynthesis in transgenic potato tubers: Using reverse genetics to confirm the in vivo enzyme

function of a steroidal alkaloid galactosyltransferase. *Plant Sci.* **168**, 267–273 (2005). doi:10.1016/j.plantsci.2004.08.006
9. M. Itkin, I. Rogachev, N. Alkan, T. Rosenberg, S. Malitsky, L. Masini, S. Meir, Y. Iijima, K. Aoki, R. de Vos, D. Prusky, S. Burdman, J. Beekwilder, A. Aharoni, GLYCOALKALOID METABOLISM1 is required for steroidal alkaloid glycosylation and prevention of phytotoxicity in tomato. *Plant Cell* **23**, 4507–4525 (2011). doi:10.1105/tpc.111.088732 Medline
10. E. Eich, Solanaceae and Convolvulaceae: Secondary Metabolites: Biosynthesis, Chemotaxonomy, Biological and Economic Significance (Springer, Berlin, 2008).
11. The Potato Genome Sequencing Consortium, Genome sequence and analysis of the tuber crop potato. *Nature* **475**, 189–195 (2011). doi:10.1038/nature10158 Medline
12. N. Umemoto, S. Katsunori, U.S. Patent application 20120159676 A1 (2012).
13. T. Yamanaka, J. P. Vincken, H. Zuilhof, A. Legger, N. Takada, H. Gruppen, C22 isomerization in alpha-tomatine-to-esculeoside A conversion during tomato ripening is driven by C27 hydroxylation of triterpenoidal skeleton. *J. Agric. Food Chem.* **57**, 3786–3791 (2009). doi:10.1021/jf900017n Medline
14. Materials and methods are available as supporting materials on *Science* Online.
15. D. J. Kliebenstein, A. Osbourn, Making new molecules: Evolution of pathways for novel metabolites in plants. *Curr. Opin. Plant Biol.* **15**, 415–423 (2012). doi:10.1016/j.pbi.2012.05.005 Medline
16. T. Winzer, V. Gazda, Z. He, F. Kaminski, M. Kern, T. R. Larson, Y. Li, F. Meade, R. Teodor, F. E. Vaistij, C. Walker, T. A. Bowser, I. A. Graham, A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science* **336**, 1704–1708 (2012). doi:10.1126/science.1220757 Medline
17. A. E. Osbourn, Gene clusters for secondary metabolic pathways: An emerging theme in plant biology. *Plant Physiol.* **154**, 531–535 (2010). doi:10.1104/pp.110.161315 Medline
18. B. Field, A. E. Osbourn, Metabolic diversification: Independent assembly of operon-like gene clusters in different plants. *Science* **320**, 543–547 (2008). doi:10.1126/science.1154990 Medline
19. P. F. Dowd, M. A. Berhow, E. T. Johnson, Differential activity of multiple saponins against omnivorous insects with varying feeding preferences. *J. Chem. Ecol.* **37**, 443–449 (2011). doi:10.1007/s10886-011-9950-3 Medline
20. S. G. Sparg, M. E. Light, J. van Staden, Biological activities and distribution of plant saponins. *J. Ethnopharmacol.* **94**, 219–243 (2004). doi:10.1016/j.jep.2004.05.016 Medline
21. C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T. L. Madden, BLAST+: Architecture and applications. *BMC Bioinform.* **10**, 421 (2009). doi:10.1186/1471-2105-10-421
22. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003). doi:10.1101/gr.1239303 Medline
23. M. Gouy, S. Guindon, O. Gascuel, SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224 (2010). doi:10.1093/molbev/msp259 Medline
24. M. Expósito-Rodríguez, A. A. Borges, A. Borges-Pérez, J. A. Pérez, Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. *BMC Plant Biol.* **8**, 131 (2008). doi:10.1186/1471-2229-8-131 Medline
25. C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, L. Pachter, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012). doi:10.1038/nprot.2012.016 Medline
26. J. T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E. S. Lander, G. Getz, J. P. Mesirov, Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011). doi:10.1038/nbt.1754 Medline
27. D. Orzaez, A. Medina, S. Torre, J. P. Fernández-Moreno, J. L. Rambla, A. Fernández-Del-Carmen, E. Butelli, C. Martin, A. Granell, A visual reporter system for virus-induced gene silencing in tomato fruit based on anthocyanin accumulation. *Plant Physiol.* **150**, 1122–1134 (2009). doi:10.1104/pp.109.139006 Medline
28. M. Itkin, H. Seybold, D. Breitel, I. Rogachev, S. Meir, A. Aharoni, TOMATO AGAMOUS-LIKE 1 is a component of the fruit ripening regulatory network. *Plant J.* **60**, 1081–1095 (2009). doi:10.1111/j.1365-313X.2009.04064.x Medline
29. R. Li, Y. Zhou, Z. Wu, L. Ding, ESI-QqTOF-MS/MS and APCI-IT-MS/MS analysis of steroid saponins from the rhizomes of *Dioscorea panthaica*. *J. Mass Spec.* **41**, 1–22 (2006). doi:10.1002/jms.988
30. S. C. Sharma, R. Chand, O. P. Sati, A. K. Sharma, Oligofurostanosides from *Solanum nigrum*. *Phytochemistry* **22**, 1241–1244 (1983). doi:10.1016/0031-9422(83)80231-3
31. H. H. Shwe, M. Aye, M. M. Sein, K. T. Htay, P. Kreitmeier, J. Gertsch, O. Reiser, J. Heilmann, Cytotoxic steroidal saponins from the rhizomes of *Tacca integrifolia*. *Chem. Biodivers.* **7**, 610–622 (2010). doi:10.1002/cbdv.200900042 Medline

Acknowledgments: We thank A. Tishbee and R. Kramer for LC-MS analysis and the European Research Council (SAMIT-FP7 program) for supporting the work in the A.A. laboratory. A.A. is the incumbent of the Peter J. Cohn Professorial Chair. J.B. was supported by the EU 7th Frame ATHENA Project (FP7-KBBE-2009-3-245121-ATHENA). U.H. was partially supported by fellowship AZ.: I/82 754, Volkswagen Foundation, Hannover, Germany. We thank the Council of Scientific and Industrial Research (India) for support to A.P.G. (Raman Research Fellowship); A.J.B., Y.C., and P.S. (Research Fellowship) and University Grants Commission (India) for supporting B.S. We also thank D. R. Nelson for assistance with CYP450 gene classification and R. Last for critically reading of the manuscript. A.A. and M.I. are inventors on publication number WO2012095843 A1 submitted by Yeda Research and Development Co. Ltd that covers the use of the *GAME4* gene for generating low-alkaloid fruit and tubers.

Supplementary Materials
www.sciencemag.org/cgi/content/full/science.1240230/DC1
Materials and Methods
Figs. S1 to S15
Tables S1 to S16
References (*21–31*)

8 May 2013; accepted 6 June 2013
Published online 20 June 2013
10.1126/science.1240230

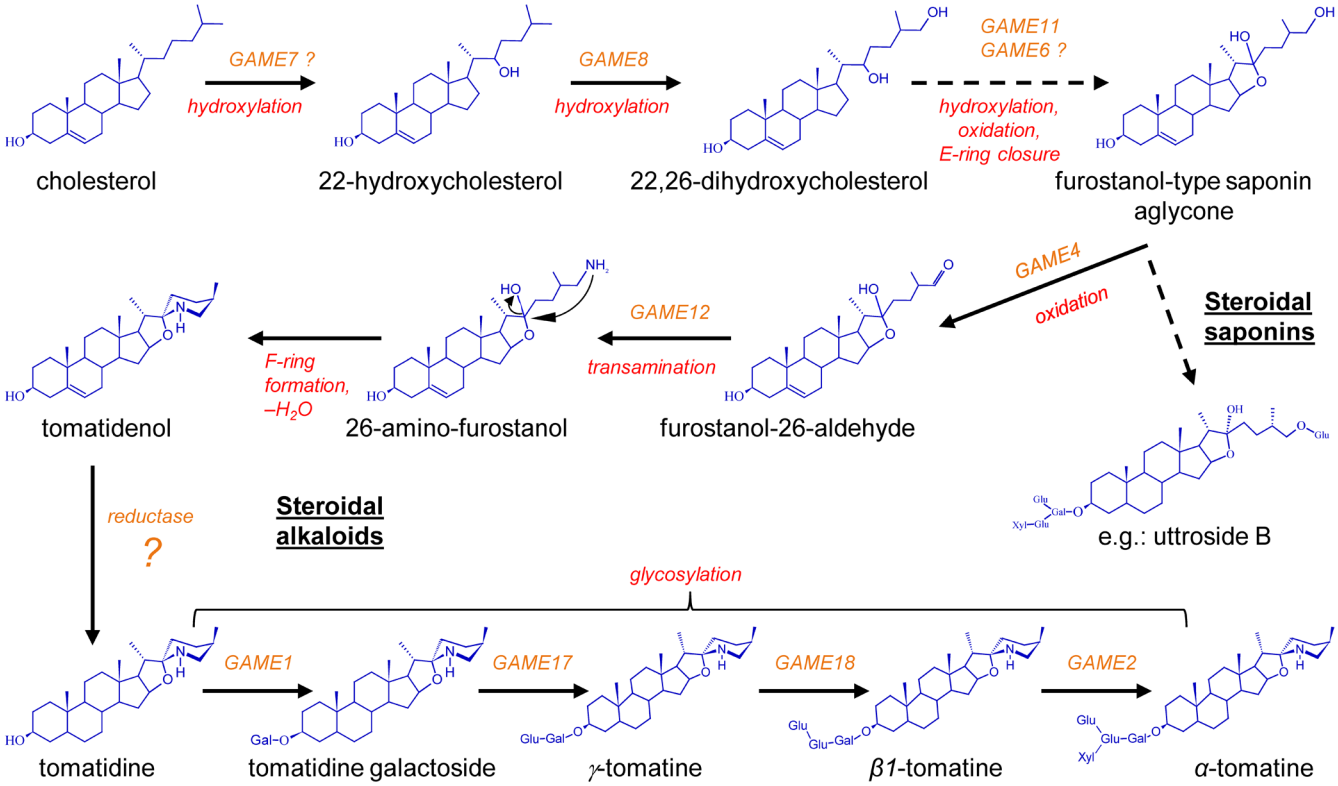


Fig. 1. Biosynthesis of steroidal alkaloids and saponins in the triterpenoid biosynthetic pathway in *Solanaceae* plants. Suggested biosynthetic pathway from cholesterol toward α-tomatine. Dashed and solid arrows represent multiple or single enzymatic reactions in the pathway, respectively. The proposed activity of *GAME1*, *GAME4* and *GAME8* was supported by investigating transgenic plants, of *GAME11*, *GAME12* and *GAME18* by Virus Induced Gene Silencing (VIGS) assays and of *GAME1*, *GAME17*, *GAME18* and *GAME2* by activity assays of the recombinant enzymes.

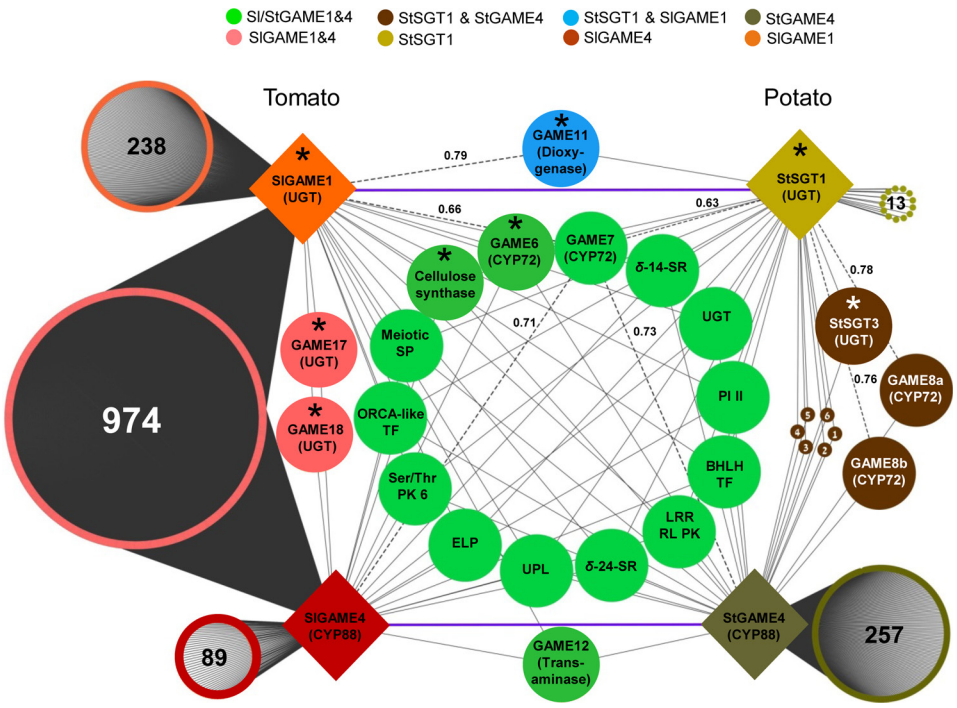


Fig. 2. Steroidal alkaloids gene discovery through coexpression network analysis in *Solanaceae* plants. Shared homologs of coexpressed genes for 'baits' from tomato (*SlGAME1 and SlGAME4*) and potato (*StSGT1 and StGAME4*). Continuous (*r*-value > 0.8) and dashed (*r*-value > 0.63) lines connect coexpressed genes. *, located in the tomato or potato chromosome 7 cluster. St, *Solanum tuberosum*; Sl, *S. lycopersicum*. Color background of gene names corresponds to bait they were found to be coexpressed with (legend above). For more details see tables S1 to S10. SP, serine proteinase; PI, proteinase inhibitor; UPL, ubiquitin protein ligase; ELP, extensin-like protein; PK, protein kinase; SR, sterol reductase; RL, receptor-like.

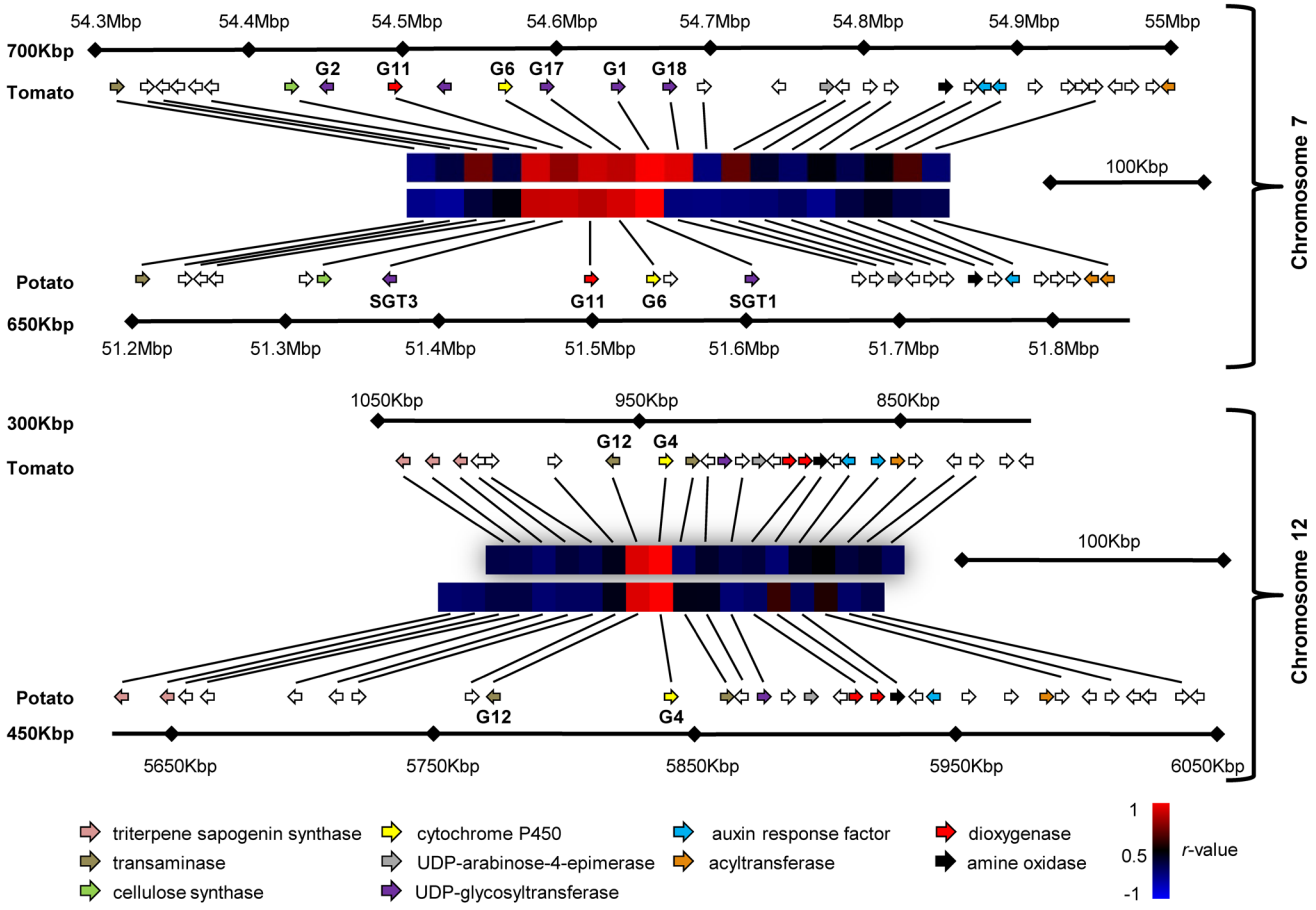


Fig. 3. Schematic map of genes identified in the duplicated genomic regions in tomato and potato and their coexpression. Coexpression with *GAME1/SGT1* (chr. 7) and *GAME4* (chr. 12) as baits in either potato or tomato are presented in a form of a heatmap (table S12). Specific gene families are indicated by colored arrows while members of other gene families are in white arrows. Note the homology in genes flanking the high coexpression regions and positioned in a matching sequence along the genome, suggesting common origin of the regions on both chromosomes (see fig. S11).

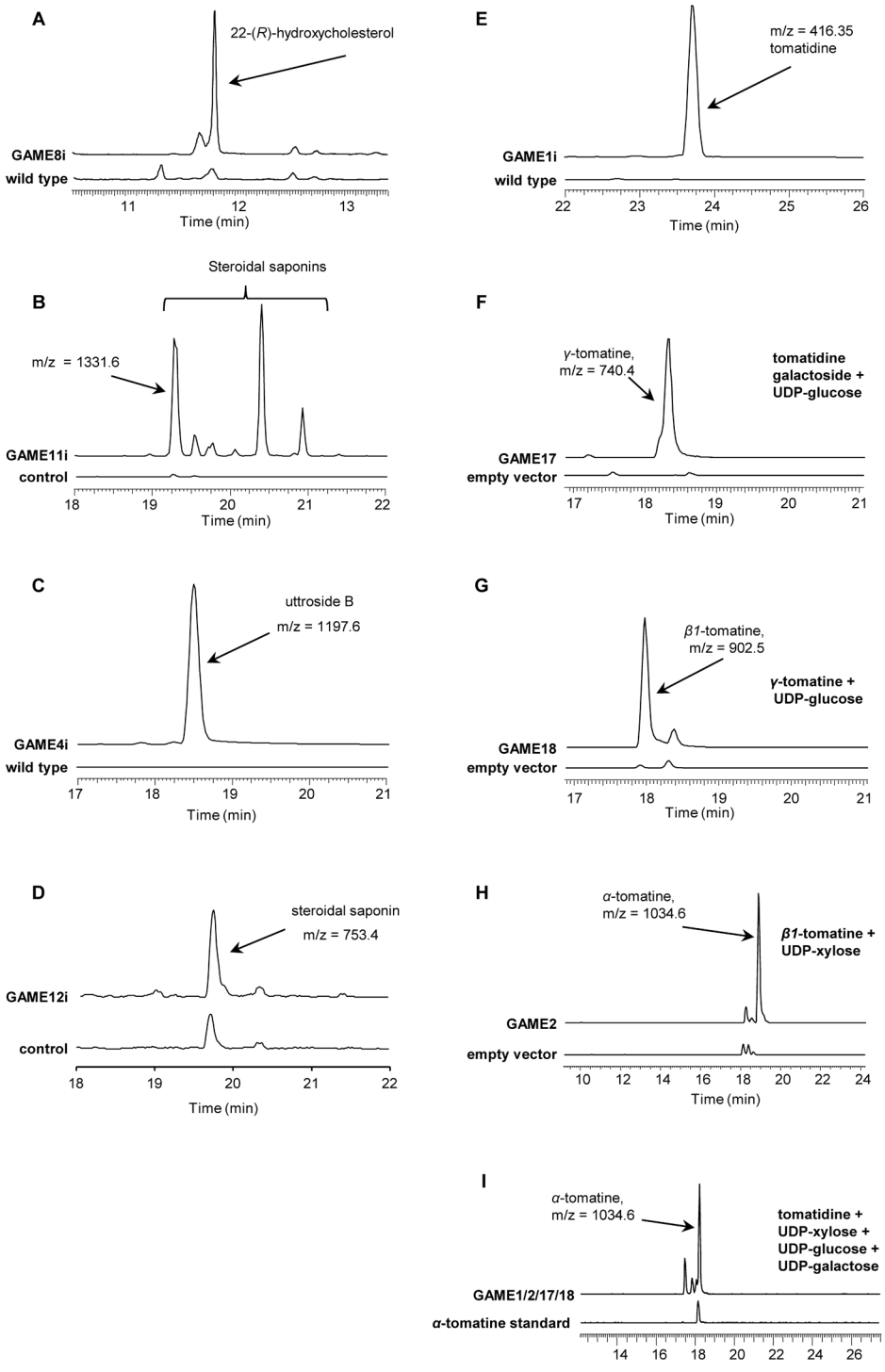


Fig. 4. Functional analysis of tomato GAME genes. (**A**) *GAME8*-silenced transgenic (RNAi) leaves accumulated 22-(*R*)-hydroxycholesterol compared to wild type. (**B**) An array of cholestanol-type steroidal saponins (STSs) accumulates in *GAME11* VIGS-silenced leaves. (**C**) An STS annotated as Uttroside B accumulates in *GAME4*-silenced transgenic leaves. (**D**) An STS (m/z = 753.4) accumulates in *GAME12* VIGS- leaves. (**E**) Tomatidine, the steroidal alkaloid aglycone, accumulates in *GAME1*-silenced transgenic leaves. (**F** to **I**) Enzyme activity assays of the 4 recombinant tomato GAME glycosyltransferases (*14*). Reactions containing GAME17 (F) and GAME18 (G) recombinant proteins with UDP-glucose as donor-substrate, and tomatidine galactoside (T-Gal) or γ-tomatine (T-Gal-Glu) as an acceptor-substrate, respectively, produced products with m/z = 740.4 and m/z = 902.5, respectively. Reaction products were identified as γ-tomatine for GAME17 (F) and *β1*-tomatine (T-Gal-Glu-Glu) for GAME18 (G). Reaction containing *β1*-tomatine, the GAME2 recombinant protein and UDP-xylose, produced α-tomatine (H). Reaction, containing tomatidine as substrate, UDP -galactose, -glucose and -xylose as sugar donors and the GAME1, GAME2, GAME17 and GAME18 recombinant proteins resulted in accumulation of α-tomatine (I). See also figs. S5, S6, and S10.

Phytochemistry 113 (2015) 24–32

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Review

The bitter side of the nightshades: Genomics drives discovery in *Solanaceae* steroidal alkaloid metabolism

P.D. Cárdenas [a,b], P.D. Sonawane [a], U. Heinig [a], S.E. Bocobza [a], S. Burdman [b], A. Aharoni [a,*]

[a] Department of Plant Sciences, Weizmann Institute of Science, Rehovot 76100, Israel
[b] Department of Plant Pathology and Microbiology, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot 76100, Israel

ARTICLE INFO

Article history:
Available online 31 December 2014

This contribution is in honor of Professor Vincenzo de Luca's 60th birthday

Keywords:
Solanaceae
Steroidal alkaloids
Metabolic gene clusters

ABSTRACT

Steroidal alkaloids (SAs) and their glycosylated forms (SGAs) are toxic compounds largely produced by members of the *Solanaceae* and *Liliaceae* plant families. This class of specialized metabolites serves as a chemical barrier against a broad range of pest and pathogens. In humans and animals, SAs are considered anti-nutritional factors because they affect the digestion and absorption of nutrients from food and might even cause poisoning. In spite of the first report on SAs nearly 200 years ago, much of the molecular basis of their biosynthesis and regulation remains unknown. Aspects concerning chemical structures and biological activities of SAs have been reviewed extensively elsewhere; therefore, in this review the latest insights to the elucidation of the SAs biosynthetic pathway are highlighted. Recently, co-expression analysis combined with metabolic profiling revealed metabolic gene clusters in tomato and potato that contain core genes required for production of the prominent SGAs in these two species. Elaborating the knowledge regarding the SAs biosynthetic pathway, the subcellular transport of these molecules, as well as the identification of regulatory and signaling factors associated with SA metabolism will likely advance understanding of chemical defense mechanisms in *Solanaceae* and *Liliaceae* plants. It will also provide the means to develop, through classical breeding or genetic engineering, crops with modified levels of anti-nutritional SAs.

Contents

1. Introduction

Solanaceae is a family of flowering plants comprising about 100 genera and 2500 species (Olmstead et al., 2008). Many of them are among the world's most important agricultural species, including tomato (*Solanum lycopersicum*), potato (*Solanum. tuberosum*), eggplant (*Solanum. melongena*), tobacco (*Nicotiana tabacum*), pepper (*Capsicum annuum*) and petunia (*Petunia* spp.). Indeed, *Solanaceae* is one of the most valuable and variable family among vegetable crops in terms of number of species that have been domesticated and its wide variety of uses (Mueller et al., 2005). Apart from their

* Corresponding author.
E-mail address: asaph.aharoni@weizmann.ac.il (A. Aharoni).

http://dx.doi.org/10.1016/j.phytochem.2014.12.010
0031-9422/

economic value, some *Solanaceae* plant species have become important model systems in plant biology research as they allow effective studies of plant development, fruit ripening and defense responses.

Originating from the Andes region of South America, both tomato and potato have been subjected to intensive breeding for resistance to biotic and abiotic stresses, yield, shelf-life, taste and nutritional quality (Bai and Lindhout, 2007). Along with selection for improved agricultural traits, co-occurrence of compounds that exert a negative impact on the nutritional quality (i.e. anti-nutritional factors) is a common issue in the breeding of *Solanaceae* species. A major group of anti-nutritionals found in *Solanaceae*, but also present in *Liliaceae* species, are steroidal alkaloids (SAs) and their glycosylated forms known as steroidal glycoalkaloids (SGAs) (Heftmann, 1983). Additionally, some related steroidal alkaloids have also been found in the *Apocynaceae* and the *Buxaceae* plant families (Atta-ur-Rahman and Choudhary, 1999; Schakirov and Yunusov, 1990). SAs are nitrogen-containing compounds derived from the cytosolic terpenoid (mevalonate) biosynthetic pathway (Fig. 1) and have been proposed to contribute to plant resistance to pathogens and predators, showing a concentration dependent toxicity to different organisms, including bacteria, fungi, viruses, insects and humans (Milner et al., 2011). High SGA concentration in food is associated with bitterness and unpleasant sensations when eaten, including convulsions, dizziness and headache (Eich, 2008; Roddick, 1996). Accordingly, the US Food and Drug Administration (FDA) limits the amount of SGAs in potato tubers to a maximum of 200 mg/kg fresh weight (Dolan et al., 2010).

2. Historical overview: Desfosses's dog and the discovery of SA toxicity

The history of SGAs research started almost 200 years ago when Desfosses, a French pharmacist, isolated an alkaline base that he named "solanée" from the berries of *Solanum nigrum*. A small amount of this extract mixed with meat, provoked narcotic symptoms like somnolence and vomiting when fed to a dog (Desfosses, 1820). Years later, Baup isolated from potato what seemed to be the same compound found by Desfosses and renamed it as solanine (**29,** see Fig. 1) (Baup, 1826). After four decades, Zwenger and Kind (1861) reported that solanine (**29**) was glycosylated, referring to the aglycone as solanidine (**22**). More than 130 years after Desfosses's findings, Kuhn and Low (1954) showed that SGA composition in potato tubers was mainly provided by the α, β and γ forms of solanine (**27**–**29**) and chaconine (**30**–**32**), having the same aglycone (i.e. solanidine (**22**)), but a different glycosylation pattern (Fig. 1B). Around the same years, Irving and coworkers reported "tomatine" (**26**) as a new antibiotic agent isolated from the tomato plant (Irving et al., 1945, 1946). Subsequently, further purification and characterization of tomatine (**26**) identified it as a glycoside alkaloid composed of the aglycone tomatidine (**17**) and a tetrasaccharide moiety composed of xylose, galactose and two glucose units (Fig. 1B; Fontaine et al., 1948; Ma and Fontaine, 1950). Four decades later, it was demonstrated that the compound known as tomatine (**26**), was in fact a mixture of two SGAs, α-tomatine (**26**) and dehydrotomatine (**28**) (Bushway et al., 1994; Friedman et al., 1994).

Subsequent to these seminal contributions, hundreds of different SAs/SGAs have been reported in *Solanaceae* species. In potato, over 50 SAs have been identified, with α-chaconine (**32**) and α-solanine (**29**) comprising >90% of the total SGA content in the tubers (Itkin et al., 2013; Shakya and Navarre, 2008). In tomato, about 100 SAs were reported (Iijima et al., 2013; Itkin et al., 2011; Moco et al., 2006), and of those α-tomatine (**26**) and dehydrotomatine (**28**) are predominant in green tissues and esculeosides (**39**–**43**) in the red ripe fruit (Fujiwara et al., 2004; Itkin et al., 2011; Mintz-Oron et al., 2008). Investigated to a much less extent, α-solasonine and α-solamargine are the two major SGAs found in eggplant fruit (Mennella et al., 2012; Sánchez-Mata et al., 2010).

3. Biosynthesis of steroidal alkaloids in the *Solanaceae*

3.1. From acetyl-CoA to cholesterol

Steroidal glycoalkaloids consist of two structural components; the aglycone, a cholesterol derived hydrophobic 27-carbon skeleton with a nitrogen atom in the side-chain, and a carbohydrate side-chain at the 3-OH position on the A ring, which typically consists of different combinations of D-glucose, D-galactose, D-xylose and L-rhamnose (Milner et al., 2011). The cellular localization of the steroidal alkaloids is not completely understood. It was suggested that in tomato and potato they are localized to the soluble phase of the cytoplasm, most probably in the vacuoles (Roddick, 1976, 1977). However, recently it has been shown that they are also present in the apoplast (Ökmen et al., 2013). The common precursor for SGA biosynthesis is cholesterol (**10**), which undergoes several steps of hydroxylation, oxidation, transamination and glycosylation generating the observed diversity of metabolites of this natural product class (Eich, 2008). Cholesterol (**10**) is produced from acetyl-CoA through the isoprenoid mevalonate (MEV) pathway (Fig. 1). Recent feeding experiments have shown that radioactive acetate, mevalonate (**2**), lanosterol (**7**), cycloartenol (**8**) and deuterium labeled cholesterol (**10**) constitute precursors of SGAs in potato and tomato (Heftmann, 1983; Petersson et al., 2013).

Squalene (**6**), an intermediate in the triterpene synthesis pathway, is oxidized by squalene epoxidase to squalene 2,3-epoxide, that is cyclized, either by cycloartenol synthase or by lanosterol synthase, to form cycloartenol (**8**) or lanosterol (**7**) (C_{30}), respectively. Both of these compounds are further metabolized by demethylation, desaturation, isomerization and reduction reactions to generate cholesterol (**10**) (C_{27}) (Nes, 2011; Ohyama et al., 2009; Suzuki and Muranaka, 2007). Until recently, little was known regarding the association between SGA production and the corresponding biosynthesis of isoprenoid precursors. Two key steps in the pathway for the synthesis of cholesterol (**10**) are the conversion of 3-hydroxy-3-methylglutaryl coenzyme A to mevalonic acid [catalyzed by a family of 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR) enzymes] and the synthesis of the squalene (**6**) precursor by squalene synthase (SQS) (Nes, 2011) (Fig. 1A). A survey of various potato genotypes, known to exhibit different levels of SGAs, showed a positive correlation between SGA content and the endogenous expression levels of *HMGR* and *SQS* genes (Krits et al., 2007). However, only recently, transgenic potato lines overexpressing these two genes, confirmed the association between HMGR and SQS with SGA content (Ginzberg et al., 2012). Additionally, conversion of cycloartenol (**8**) to plant sterols, other than cholesterol (**10**), has been studied in relation to SGA synthesis. A key step in this process is the methylation of the sterol side-chain of cycloartenol (**8**) by *S*-adenosyl-L-methionine:sterol C24-methyltransferases type 1 (SMT1) enzyme (Fig. 1A). Overexpression of *SMT1* in potato plants resulted in an increased content of total alkylated sterols (**9**) at the expense of cholesterol (**10**) (non-alkylated sterol) (Arnqvist et al., 2003). As a consequence of *SMT1* overexpression, the level of free cholesterol (**10**) was decreased together with a reduction in total SGA content. Recently, a sterol side-chain reductase 2 (SSR2) was identified in both potato and tomato. Enzymatic *in vitro* assays and yeast complementation on the background of various ergosterol (*erg*) mutant strains showed SSR2 as a key step in the branching point between

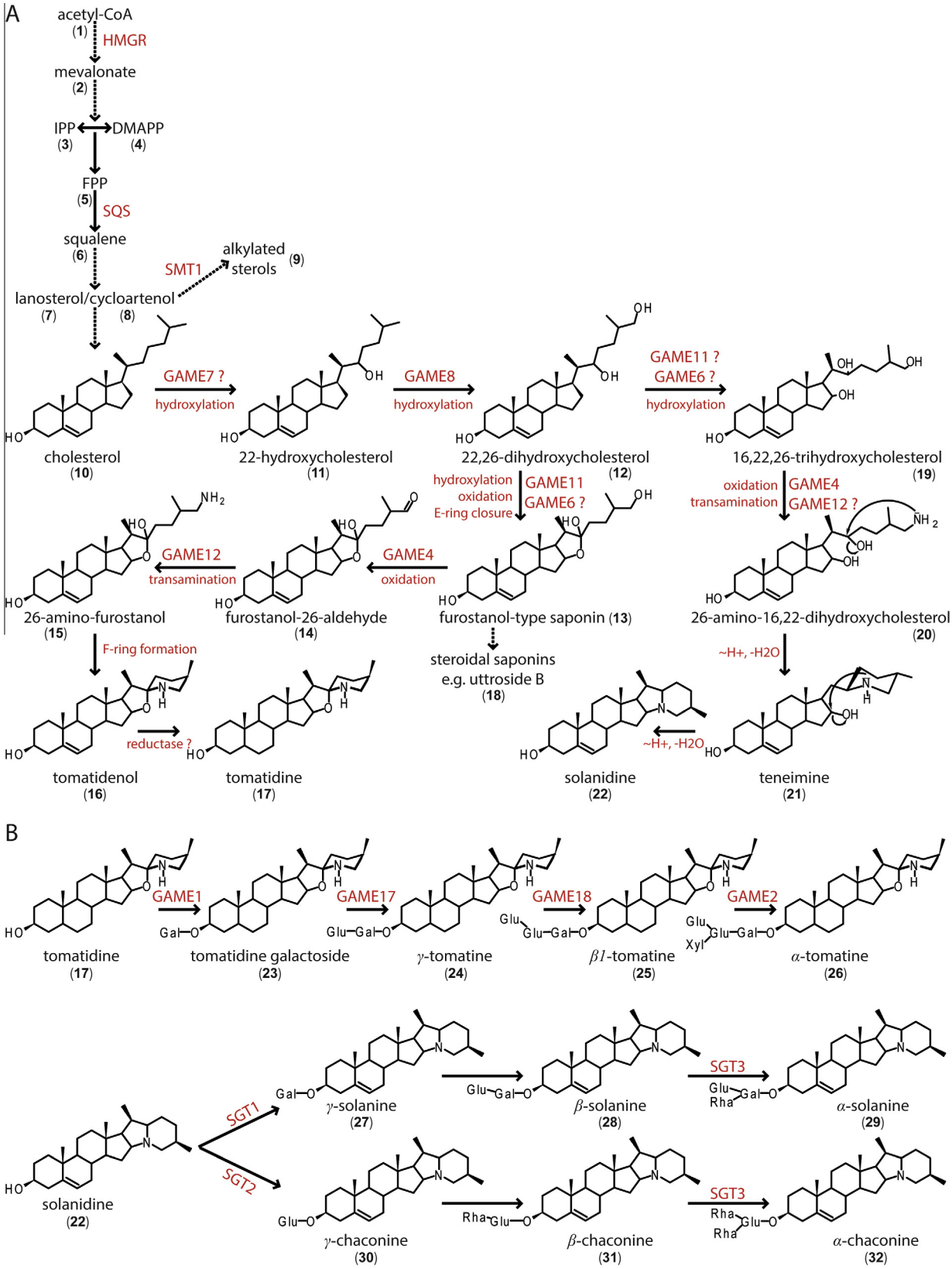


Fig. 1. Biosynthesis of steroidal glycoalkaloids in *Solanaceae* species. Steroidal glycoalkaloids (SGAs) are derived from the cytosolic mevalonic acid pathway. (A) Starting from acetyl-CoA (**1**), cholesterol (**10**) is generated and further modified by glycoalkaloid metabolism (GAME) enzymes through hydroxylation, oxidation and transamination to generate the aglycone tomatidine (**17**) in tomato (Itkin et al., 2013) and solanidine (**22**) in potato. (B) In tomato, glycosylation of the aglycone tomatidine (**17**) generates α-tomatine (**26**). In potato, the aglycone solanidine (**22**) is glycosylated by solanidine glycosyltransferases (SGTs) enzymes to generate α-solanine (**29**) and α-chaconine (**32**). Some enzymes involved in the synthesis of precursors upstream of cholesterol (**10**) in the mevalonate pathway have been linked to SGA content in *Solanaceae* species, including 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR), squalene synthase (SQS) and S-adenosyl-L-methionine:sterol C24-methyl transferases type 1 (SMT1). Dashed arrows represent multiple biosynthetic reactions whereas solid arrows represent a single step.

C_{24} alkylsterols and cholesterol (**10**) biosynthesis. Silencing of SSR2 resulted in a significant reduction of cholesterol (**10**) and SGAs in tomato and potato (Sawai et al., 2014).

3.2. From cholesterol to the steroidal alkaloid aglycone

Up to recently, the predicted biosynthetic route from cholesterol (**10**) to the SA aglycone (**17**, **22**) in plants of the *Solanaceae* family was based on the detection of pathway intermediates discovered in the liliaceous solanidine (**22**) producing plant *Veratrum grandiflorum* (Kaneko et al., 1977). In *Veratrum*, solanidine (**22**) was considered a precursor for jerveratrum alkaloids (Kaneko et al., 1972) and its synthesis was proposed to involve multiple steps of hydroxylation, oxidation and amination of the cholesterol (**10**) skeleton, with arginine suggested as the main source of nitrogen incorporated into the side-chain (Kaneko et al., 1976). However, recently, Itkin et al. (2013) identified a set of *GLYCOALKALOID METABOLISM* (*GAME*) genes that participate in the SGA biosynthesis pathway in potato and tomato (Fig. 1). Through comparative co-expression analysis, it was found that in tomato, six SGA genes (*GAME1*, *GAME17*, *GAME18*, *GAME2*, *GAME11* and *GAME6*) are arranged in a cluster on chromosome 7, whereas two other genes (*GAME4* and *GAME12*) are located next to each other on chromosome 12 (Fig. 2). Furthermore, two additional genes, not located in these clusters, were also associated with SGA biosynthesis (i.e. *GAME7* and *GAME8*). Out of these, six genes were shown to be specifically involved in the SA aglycone synthesis and four in the generation of the tetra-saccharide moiety. Moreover, in potato, four SGA-related genes are located in chromosome 7 (i.e. *SGT3*, *GAME11*, *GAME6* and *SGT1*) and two in chromosome 12 (i.e. *GAME12* and *GAME4*; Fig. 2).

Functional characterization supported by transgenic plants, virus-induced gene silencing (VIGS), enzyme activity assays, and data reported by Umemoto and Sasaki (2012) (patent application for *GAME6* and *GAME7*), allowed Itkin et al. (2013) to propose a revised SGA biosynthetic pathway in the *Solanaceae* species (Fig. 1). In the newly predicted pathway, cholesterol (**10**) is putatively hydroxylated at the C_{22} position by GAME7 (cytochrome P450 72A186), and further hydroxylated at the C_{26} position by GAME8 (cytochrome P450 72A208). The role of GAME8 was supported by analysis of *GAME8*-silenced tomato lines that accumulated the intermediate 22-(R)-hydroxycholesterol (**11**) (Itkin et al., 2013). The 22, 26-dihydroxycholesterol (**12**) is then hydroxylated at C_{16} and oxidized at C_{22}, followed by closure of the E-ring by GAME11 (2-oxoglutarate-dependent dioxygenase) and GAME6 (cytochrome P450 72A188) to form the furostanol-type aglycone (**13**). Silencing of *GAME11* in tomato resulted in decrease of α-tomatine (**26**) levels in leaves and accumulation of cholestanol-type steroidal saponins (**18**). The furostanol-intermediate (**13**) is subsequently oxidized by GAME4 (a cytochrome P450 88D) to its 26-aldehyde (**14**), which is the substrate for transamination catalyzed by GAME12 (aminotransferase-like protein). Silencing of both *GAME4* and *GAME12* resulted in a reduction in SGA content and accumulation of non-nitrogenous steroidal saponins, revealing their importance in the amination step (Itkin et al., 2013). The 26-aldehyde intermediate (**14**) has been recently proposed in an independent report based on feeding of isotope-labeled cholesterol (**10**) (Ohyama et al., 2013). Nucleophilic attack of the amino nitrogen at C_{22} leads to the formation of tomatidenol (**16**), which is dehydrogenated to tomatidine (**17**) (Itkin et al., 2013).

The detailed functional analysis of SGAs enzymes in potato has yet to be carried out. Yet, the similarity of potato genes to the tomato ones and the chemical structure of the potato SGAs allows the prediction of a mechanism/reaction sequence towards solanidine (Fig. 1A). We propose the divergence of biosynthetic pathways between tomato and potato at 22,26-dihydroxycholesterol (**12**). In contrast to tomato, GAME11 or GAME6 orthologs might catalyze the hydroxylation at C_{16} followed by oxidation and transamination (catalyzed by the GAME4 and the GAME12 orthologs) at C_{26} without prior closure of the E-ring leading to furostanol-type intermediates. Two subsequent nucleophilic substitutions, via a teneimine (**21**) intermediate, similarly to the initially proposed SGA pathway

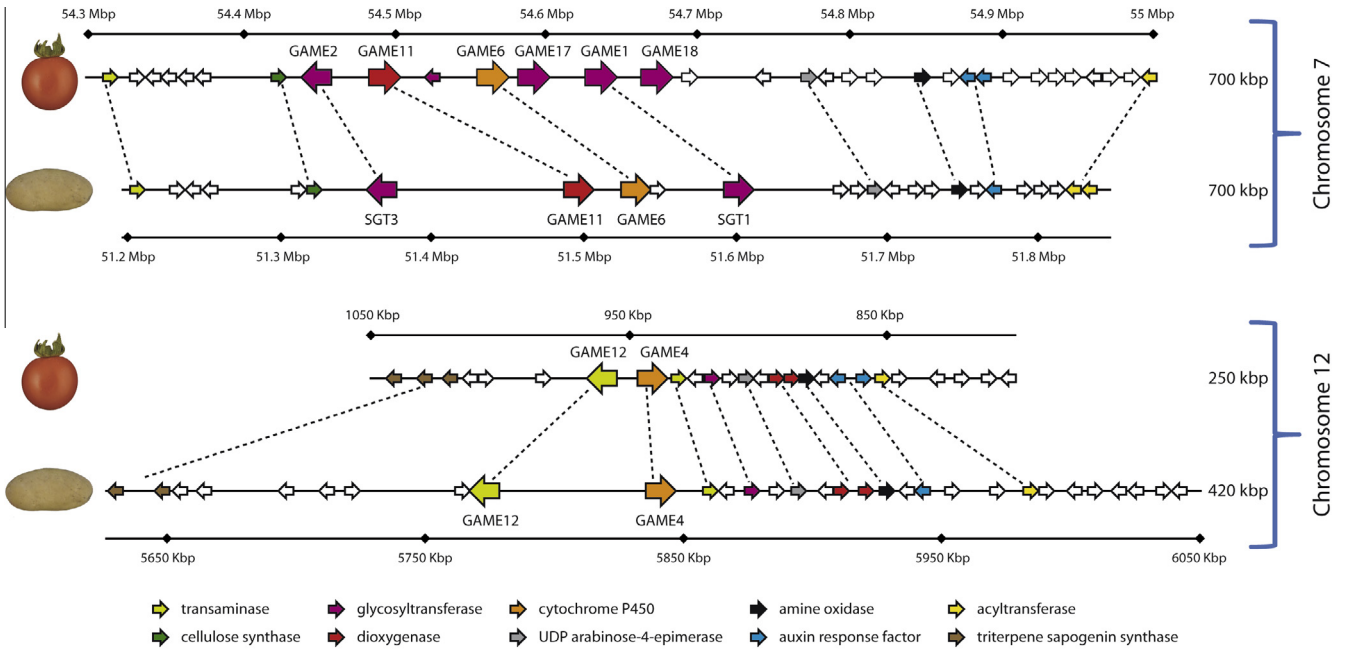


Fig. 2. Steroidal alkaloids genes are present in metabolic gene clusters in tomato and potato. In tomato, six genes (i.e. *GAME 2*, *11*, *6*, *17*, *1*, *18*) participating in SGAs biosynthesis are located in chromosome 7 and two (i.e. *GAME 12* and *4*) in chromosome 12. In potato, four SGA-related genes are located in chromosome 7 (i.e. *SGT3*, *GAME11*, *GAME6* and *SGT1*) and two in chromosome 12 (i.e. *GAME12* and *4*). Conserved genes (connected by dashed lines) in both species revealed a putative common origin of the regions on both chromosomes.

in the *Liliaceae* SGAs (Friedman, 2002; Kaneko et al., 1977) could explain the formation of the solanidine aglycone (**22**).

3.3. Glycosylation of the steroidal alkaloid aglycones

The major tomato SGAs, α-tomatine (**26**) and dehydrotomatine (**28**), differ in the aglycone (tomatidine (**17**) and tomatidenol (**16**), respectively) but both contain the same tetra-saccharide moiety (i.e. lycotetraose), composed of a single D-xylose and D-galactose and two D-glucose units. In potato, α-chaconine (**32**) and α-solanine (**29**) have the same aglycone, solanidine (**22**), but are glycosylated with chacotriose (i.e. a single D-galactose and two L-rhamnose units) and solatriose (i.e. D-galactose, D-glucose and L-rhamnose), respectively. In eggplant, α-solamargine and α-solasonine are also glycosylated by chacotriose and solatriose, respectively, but the aglycone moiety is solasodine (Milner et al., 2011).

Early molecular studies of the SA aglycone glycosylation were performed in potato. Three different solanidine glycosyltransferase (SGT) enzymes were associated with the conversion of the aglycone solanidine (**22**) into α-chaconine (**32**) and α-solanine (**29**) (Fig. 1B). Silencing of *SGT1*, a UDP-galactose:solanidine galactosyltransferase, resulted in inhibition of α-solanine (**29**) accumulation that was compensated with high levels of α-chaconine (**32**) (McCue et al., 2005; Moehs et al., 1997). On the other hand, silencing of a UDP-glucose:solanidine glucosyltransferase, *SGT2*, resulted in plants having reduced levels of α-chaconine (**32**) and accumulation of α-solanine (**29**) (McCue et al., 2006). Functional characterization of SGT1 and SGT2, showed that these enzymes are responsible for the specific synthesis of either α-solanine (**29**) or α-chaconine (**32**). Eventually, SGT3, a UDP-rhamnose:β-solanine/β-chaconine rhamnosyltransferase, was shown to catalyze the terminal step in formation of the potato SGA sugar moiety. In *SGT3*-silenced potato lines there was a dramatic decrease in the accumulation of α-chaconine (**32**) and α-solanine (**29**), while β-solanine (**28**) and β-chaconine (**31**) levels increased (McCue et al., 2007).

In tomato, the first gene reported to be involved in the synthesis of the SGA saccharide moiety was *GAME1*, encoding a UDP-galactose:tomatidine galactosyltransferase homologous to the potato *SGT1*. Phylogenetic analysis showed that in the large family of plant glycosyltransferases, the three potato SGTs and GAME1 form a separate protein clade involved in the transfer of sugar moieties to SAs and steroidal saponins (Itkin et al., 2011). Most glycosyltransferases involved in the biosynthesis of the saccharide moiety of the main tomato and potato glycoalkaloids were located in the metabolic gene clusters. Interestingly, the tomato chromosome 7 cluster included all the required glycosyltransferases (i.e. *GAME1*, *2*, *17* and *18*) while in potato, only *SGT1* and *SGT3* are present in the cluster while *SGT2* is located elsewhere in the genome (Itkin et al., 2013; Fig. 2).

In the proposed pathway for SGA biosynthesis in tomato, tomatidine (**17**) is converted by GAME1 to tomatidine galactoside (**23**). This compound is glucosylated by GAME17 (UDP-glucosyltransferase) into γ-tomatine (**24**), which is further glucosylated by GAME18 (UDP-glucosyltransferase) to β_1-tomatine (**25**), which is finally converted to α-tomatine (**26**) by GAME2 (UDP-xylosyltransferase) (Fig. 1B) (Itkin et al., 2013). Virus-induced gene silencing assays provided a rapid evaluation method of the metabolic role of candidate genes. For instance, *GAME18*-silenced plants accumulated γ-tomatine (**24**) thus identifying the role of GAME18 as a glucosyltransferase (Fig. 3). Enzyme activity assays containing these four recombinant glycosyltransferases in a single reaction with tomatidine (**17**) as a substrate and UDP-galactose, -glucose and -xylose as sugar donors resulted in the accumulation of α-tomatine (**26**) (Itkin et al., 2013).

4. Modification of steroidal glycoalkaloids during tomato fruit development and ripening

In tomato fruit, α-tomatine (**26**) is extensively modified during development and ripening (Iijima et al., 2008; Mintz-Oron et al., 2008; Moco et al., 2006). While in the green tomato fruit α-tomatine (**26**) is the most abundant SGA, lycoperosides and esculeosides (**39**–**43**) are predominant in the red ripe stage (Fig. 4). Total SGA content remains constant across fruit ripening and this suggested that the entire pool of α-tomatine (**26**) is converted during this process to the latter compounds (Iijima et al., 2008; Nohara et al., 2006; Yamanaka et al., 2009). However, no genes or enzymes have been characterized in this pathway to date.

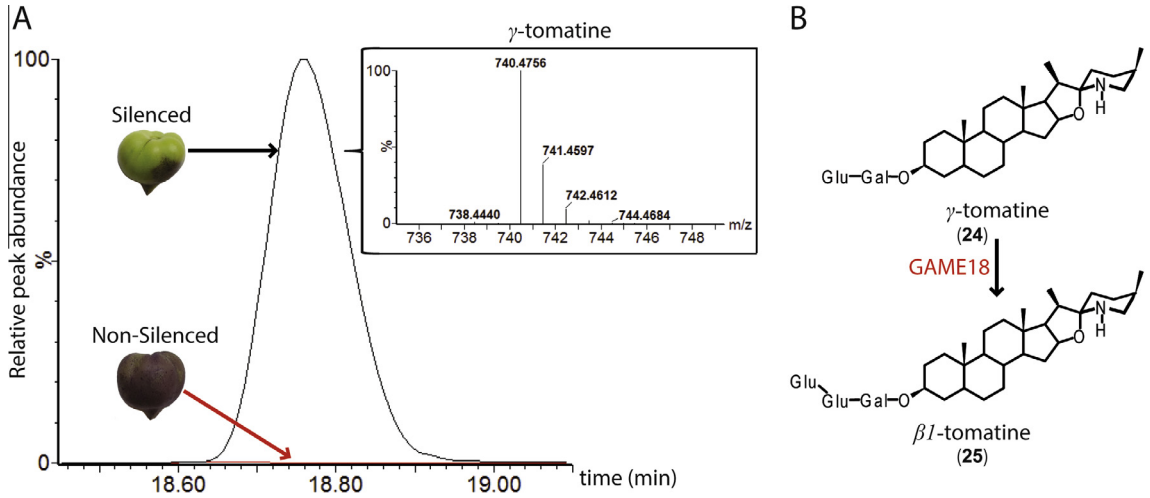


Fig. 3. Functional characterization of candidate steroidal alkaloid-biosynthesis genes by means of virus induced gene silencing assays in tomato fruit. (A) Rapid evaluation of the metabolic role of candidate steroidal alkaloid (SA)-related genes was performed by virus induced gene silencing (VIGS) assays coupled to metabolic profiling with high resolution liquid chromatography–mass spectrometry analysis. Comparison of silenced to non-silenced tomato tissues established γ-tomatine (**24**) accumulation when *GAME18* was suppressed, demonstrating its role as a glucosyltransferase (Itkin et al., 2013). (B) Glucosylation of γ-tomatine (**24**) to *β1*-tomatine (**25**) by GAME18. Silencing was performed in the background of a transgenic tomato line expressing the *Antirrhinum majus DELILA* and *ROSEA1* (*DEL*/*ROS*) transcription factors, that convey a purple anthocyanin-rich phenotype to the fruit (Orzaez et al., 2009). The VIGS vector includes the candidate gene as well as the DEL/ROS sequences in a way that allows locating green fruit patches in which the candidate gene was likely silenced. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

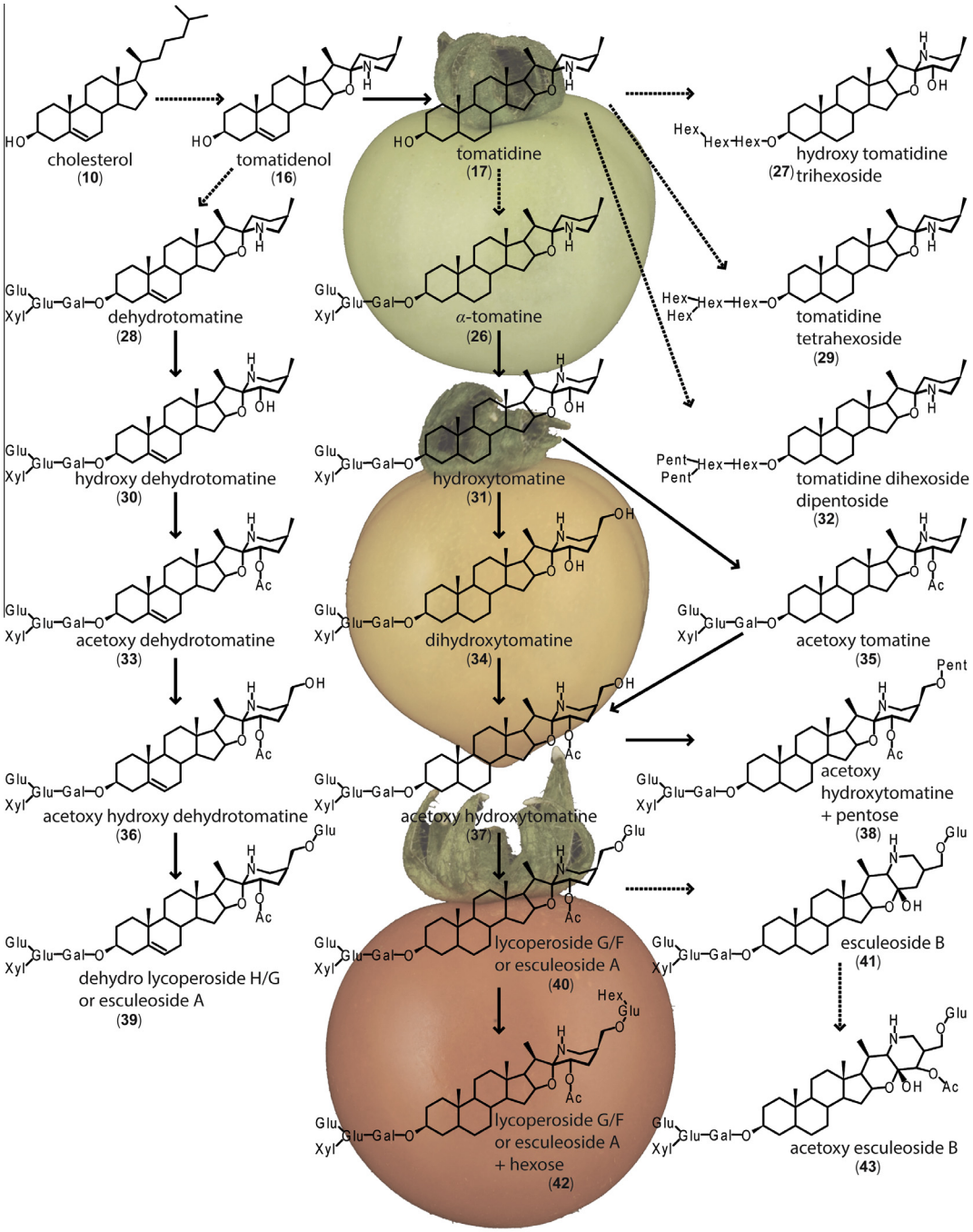


Fig. 4. Modification of steroidal glycoalkaloids structures during tomato fruit development and ripening. The major steroidal glycoalkaloids (SGAs) in green tomato fruit are α-tomatine (**26**) and dehydrotomatine (**28**). Subsequently, hydroxy- and acetoxy-derivatives (**30**–**38**) accumulate at the breaker stage. In the red ripe tomato fruit, the most abundant SGAs correspond to the lycoperosides and esculeosides (**39**–**43**). Dashed arrows represent multiple biosynthetic reactions whereas solid arrows represent a single step. The pathway was adapted from Mintz-Oron et al. (2008). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Lycoperosides A–D and F–H were first isolated from leaves and fruit of *S. lycopersicum* (Yahara et al., 1996, 2004). Elucidation of tomato lycoperosides structure by NMR spectroscopy established hydroxylation, acetylation and glycosylation on the aglycone (**17**). With similar chemical modifications, esculeosides A (**40**) and B (**41**) are the major glycosylated SGAs isolated from red ripe tomato fruit (Fujiwara et al., 2003, 2004). Recently, it has been suggested that the final step of esculeoside A (**40**) biosynthesis, starting from α-tomatine (**26**), is controlled by the ripening hormone ethylene (Iijima et al., 2008, 2009). In *non-ripening* (*nor*) and *ripening-inhibitor* (*rin*) tomato mutants that display altered ripening, levels of esculeoside A (**40**) were decreased; however; they exhibited an increased content of upstream acetylated and glycosylated intermediates (**35**, **37**) (Iijima et al., 2008). This suggested that only the final glycosylation step of esculeoside A (**40**) biosynthesis is associated with ethylene-ripening regulated events (Fig. 4).

Steroidal glycoalkaloids are subject to further metabolism including modifications of the saccharide moiety. In potato, it has been reported that endogenous enzymes are able to remove the

rhamnose from α-chaconine (**32**) and α-solanine (**29**) (Bushway et al., 1988, 1990; Swain et al., 1978). In *Solanum torvum*, a β-glucosidase was characterized which specifically cleaves glucose units attached to the saponins torvosides A and H (Arthan et al., 2006). Enzymes with similar activities, known as tomatinases, are produced in *Solanum* pathogens that actively detoxify SGAs. These enzymes remove single or multiple sugar molecules from the saccharide moiety of α-tomatine (**26**), creating less cytotoxic compounds for the pathogen (Ford et al., 1977; Ökmen et al., 2013) and have been characterized in fungi such as *Fusarium oxysporum* f. sp. *lycopersici* and *Cladosporium fulvum*. A tomatinase gene was recently reported in the tomato bacterial pathogen *Clavibacter michiganensis* subsp. *michiganensis* (*Cmm*). Activity was demonstrated *in vitro* and *Cmm* mutants impaired in the tomatinase showed growth inhibition but were not affected in virulence (Kaup et al., 2005). As the glycosylated forms of SAs could increase the potency against pathogens (Hoagland, 2009), the presence of glycosidases in the plant can be associated with modulation of SGA toxicity in defense response.

5. Metabolic gene clusters and steroidal alkaloid biosynthesis

Organization of metabolic genes in clusters is typically associated with operons present in bacteria or fungal genomes (Osbourn, 2010a). Nevertheless, nearly a dozen metabolic gene clusters have been reported in the past years in multiple plant species, all without exception associated with the biosynthesis of a certain class of specialized metabolites (Boycheva et al., 2014). Similarly to bacterial operons, plant gene clusters are composed of non-homologous genes that contribute to the same biosynthetic pathway. Yet, unlike their bacterial counterparts, plant genes are transcribed as separate mRNAs rather than as a single polycistronic mRNA (Osbourn, 2010b). Interestingly, to date such clusters were mostly associated with terpenoid biosynthesis. Metabolic gene clusters have been described for the synthesis of cyclic hydroxamic acids in maize (Frey et al., 1997); thalianol, marneral and avenacin triterpenoids in Arabidopsis and oat, respectively (Field and Osbourn, 2008; Field et al., 2011; Qi et al., 2004); phytocassane and momilactone diterpenes in rice (Swaminathan et al., 2009; Wilderman et al., 2004); cyanogenic glucosides and a novel triterpene dihydrolupeol precursor of 20-hydroxybetulinic acid in *Lotus japonicus* (Krokida et al., 2013; Takos et al., 2011); noscapine alkaloids in Papaver (Winzer et al., 2012) as well as the recently described gene cluster for steroidal alkaloids in *Solanum* species described above (Fig. 2; Itkin et al., 2013).

An important question remains with respect to the advantage of having the SGAs and other metabolic gene clusters in plant genomes. Most of the metabolites produced through the clusters described up to now were implicated in the synthesis of defense compounds, suggesting that one of the advantages might be the co-inheritance of an important fitness trait as discussed by Osbourn (2010b). Cluster disruption, can lead to partial activity of pathways generating toxic intermediates for the plant. As an example, the silencing of *GAME1* (encoding a glycosyltransferase) in tomato plants led to a reduction in α-tomatine (**26**) content, but an increase in its aglycone tomatidine (**17**). *GAME1*-silenced plants showed severe developmental defects including retarded growth, deformed leaves and partial abortion of flower buds (Itkin et al., 2011). The *GAME1* gene is part of the gene cluster located on the tomato and potato chromosome 7 (Fig. 2) and this may possibly secure its presence in the next generations, to avoid phytotoxicity by the SGA pathway intermediates.

Another major benefit provided by gene clustering over randomly dispersed genes is in facilitating the co-regulation of clustered genes. Interestingly, Itkin et al. (2013) showed that both in potato and tomato, the regions of clustered genes in chromosomes 7 and 12 are flanked by similar genes positioned identically along the genome indicating co-inheritance from a common *Solanum* ancestor. While the set of genes in the clusters flanking regions were not co-expressed, the clustered genes were found to be tightly co-expressed, suggesting a possible coordinated regulatory process for their control (Itkin et al., 2013). The mechanism underlying this co-regulation remains to be elucidated, but specific chromatin decondensation might be involved as shown for the avenacin-producing cluster in oat (Wegel et al., 2009).

6. Concluding remarks: future prospects in the study of steroidal alkaloids

The recent findings with respect to the biosynthesis of SAs/SGAs in *Solanaceae* species extend the contribution of non-homologous clustered genes in plant specialized metabolism. Nevertheless, extensive research is required in several areas including the discovery of new enzymes generating the large SA diversity, identification of regulatory mechanisms and links to signaling pathways of plant defense response as well as transport and localization of SAs at the cell type and subcellular level. The availability of sequenced *Solanaceae* genomes, transcriptomics data and metabolic profiling of multiple plant species will allow us to select new putative gene candidates involved in SA metabolism. Of outstanding interest will be to determine whether the SA/SGA genes are conserved in other *Solanaceae* and in the *Liliaceae* plant species and if they participate in the metabolism of the related steroidal saponins or exclusively in SAs synthesis.

Additionally, recent advances in the discovery of metabolic gene clusters provide a platform for the production of highly valuable SAs or steroidal saponins and/or its derivatives for medical and industry purposes. For instance, new strategies can be developed for the production of steroidal drugs derived from SA/saponin precursors (e.g. diosgenin derivatives; Patel et al., 2012), either through heterologous expression systems or modified plants. Conventional breeding using naturally occurring variants and single gene genetic engineering had proven to be efficient ways to modulate specialized metabolites content. Nevertheless, with the discovery of gene clusters in plants, new strategies can be proposed for their generation in naturally non-producing plants. The use of whole clusters in heterologous systems has the potential to advance the current approaches for specialized metabolites production. The dual properties of SAs providing a protective role for plants against pathogens and predators and their anti-nutritional effects in animals and humans, raise the opportunity to generate functional crops with modified levels of specific SGAs to improve either nutritional value or crop protection of *Solanaceae* species.

References

Arnqvist, L., Dutta, P.C., Jonsson, L., Sitbon, F., 2003. Reduction of cholesterol and glycoalkaloid levels in transgenic potato plants by overexpression of a type 1 sterol methyltransferase cDNA. Plant Physiol. 131, 1792–1799.

Arthan, D., Kittakoop, P., Esen, A., Svasti, J., 2006. Furostanol glycoside 26-O-β-glucosidase from the leaves of *Solanum torvum*. Phytochemistry 67, 27–33.

Atta-ur-Rahman, Choudhary, M.I., 1999. Diterpenoid and steroidal alkaloids. Nat. Prod. Rep. 16, 619–635.

Bai, Y., Lindhout, P., 2007. Domestication and breeding of tomatoes: what have we gained and what can we gain in the future? Ann. Bot. 100, 1085–1094.

Baup, M., 1826. Extrait d'une lettre. Ann. Chim. Phys. 31, 108–109.

Boycheva, S., Daviet, L., Wolfender, J.L., Fitzpatrick, T.B., 2014. The rise of operon-like gene clusters in plants. Trends Plant Sci.. http://dx.doi.org/10.1016/j.tplants.2014.01.013.

Bushway, A.A., Bushway, R.J., Kim, C.H., 1988. Isolation, partial purification and characterization of a potato peel glycoalkaloid glycosidase. Am. Potato J. 65, 621–631.

Bushway, A.A., Bushway, R.J., Kim, C.H., 1990. Isolation, partial purification and characterization of a potato peel α-solanine cleaving glycosidase. Am. Potato J. 67, 233–238.

Bushway, R.J., Perkins, L.B., Paradis, L.R., Vanderpant, S., 1994. High-performance liquid chromatographic determination of the tomato glycoalkaloid, tomatine, in green and red tomatoes. J. Agric. Food Chem. 42, 2824–2829.
Desfosses, M., 1820. Extrait d'une lettre. J. Pharm. 6, 374–376.
Dolan, L.C., Matulka, R.A., Burdock, G.A., 2010. Naturally occurring food toxins. Toxins 2, 2289–2332.
Eich, E., 2008. *Solanaceae* and *Convolvulaceae*: Secondary Metabolites. Springer, Berlin.
Field, B., Fiston-Lavier, A.S., Kemen, A., Geisler, K., Quesneville, H., Osbourn, A.E., 2011. Formation of plant metabolic gene clusters within dynamic chromosomal regions. Proc. Natl. Acad. Sci. U.S.A. 108, 16116–16121.
Field, B., Osbourn, A.E., 2008. Metabolic diversification-independent assembly of operon-like gene clusters in different plants. Science 320, 543–547.
Fontaine, T.D., Irving, G.W., Ma, R., Poole, J.B., Doolittle, S.P., 1948. Isolation and partial characterization of crystalline tomatine, an antibiotic agent from the tomato plant. Arch. Biochem. 18, 467–475.
Ford, J.E., McCance, D.J., Drysdale, R.B., 1977. The detoxification of α-tomatine by *Fusarium oxysporum* f. sp. *lycopersici*. Phytochemistry 16, 545–546.
Frey, M., Chomet, P., Glawischnig, E., Stettner, C., Grun, S., Winklmair, A., Eisenreich, W., Bacher, A., Meeley, R.B., Briggs, S.P., Simcox, K., Gierl, A., 1997. Analysis of a chemical plant defense mechanism in grasses. Science 277, 696–699.
Friedman, M., 2002. Tomato glycoalkaloids: role in the plant and in the diet. J. Agric. Food Chem. 50, 5751–5780.
Friedman, M., Levin, C.E., McDonald, G.M., 1994. α-Tomatine determination in tomatoes by HPLC using pulsed amperometric detection. J. Agric. Food Chem. 42, 1959–1964.
Fujiwara, Y., Takaki, A., Uehara, Y., Ikeda, T., Okawa, M., Yamauchi, K., Ono, M., Yoshimitsu, H., Nohara, T., 2004. Tomato steroidal alkaloid glycosides, esculeosides A and B, from ripe fruits. Tetrahedron 60, 4915–4920.
Fujiwara, Y., Yahara, S., Ikeda, T., Ono, M., Nohara, T., 2003. Cytotoxic major saponin from tomato fruits. Chem. Pharm. Bull. 51, 234–235.
Ginzberg, I., Thippeswamy, M., Fogelman, E., Demirel, U., Mweetwa, A.M., Tokuhisa, J., Veilleux, R.E., 2012. Induction of potato steroidal glycoalkaloid biosynthetic pathway by overexpression of cDNA encoding primary metabolism HMG-CoA reductase and squalene synthase. Planta 235, 1341–1353.
Heftmann, E., 1983. Review of steroids in *Solanaceae*. Phytochemistry 22, 1843–1860.
Hoagland, R.E., 2009. Toxicity of tomatine and tomatidine on weeds, crops and phytopathogenetic fungi. Allelopathy J. 23, 425–436.
Iijima, Y., Fujiwara, Y., Tokita, T., Ikeda, T., Nohara, T., Aoki, K., Shibata, D., 2009. Involvement of ethylene in the accumulation of esculeoside A during fruit ripening of tomato (*Solanum lycopersicum*). J. Agric. Food Chem. 57, 3247–3252.
Iijima, Y., Nakamura, Y., Ogata, Y., Tanaka, K., Sakurai, N., Suda, K., Suzuki, T., Suzuki, H., Okazaki, K., Kitayama, M., Kanaya, S., Aoki, K., Shibata, D., 2008. Metabolite annotations based on the integration of mass spectral information. Plant J. 54, 949–962.
Iijima, Y., Watanabe, B., Sasaki, R., Takenaka, M., Ono, H., Sakurai, N., Umemoto, N., Suzuki, H., Shibata, D., Aoki, K., 2013. Steroidal glycoalkaloid profiling and structures of glycoalkaloids in wild tomato fruit. Phytochemistry 95, 145–157.
Irving, G.W., Fontaine, T.D., Doolittle, S.P., 1945. Lycopersicin, a fungistatic agent from the tomato plant. Science 102, 9–11.
Irving, G.W., Fontaine, T.D., Doolittle, S.P., 1946. Partial antibiotic spectrum of tomatin, an antibiotic agent from tomato plant. J. Bacteriol. 52, 601–607.
Itkin, M., Heinig, U., Tzfadia, O., Bhide, A.J., Shinde, B., Cardenas, P.D., Bocobza, S.E., Unger, T., Malitsky, S., Finkers, R., Tikunov, Y., Bovy, A., Chikate, Y., Singh, P., Rogachev, I., Beekwilder, J., Giri, A.P., Aharoni, A., 2013. Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. Science 341, 175–179.
Itkin, M., Rogachev, I., Alkan, N., Rosenberg, T., Malitsky, S., Masini, L., Meir, S., Iijima, Y., Aoki, K., de Vos, R., Prusky, D., Burdman, S., Beekwilder, J., Aharoni, A., 2011. GLYCOALKALOID METABOLISM1 is required for steroidal alkaloid glycosylation and prevention of phytotoxicity in tomato. Plant Cell 23, 4507–4525.
Kaneko, K., Tanaka, M.W., Mitsuhashi, H., 1976. Origin of the nitrogen in the biosynthesis of solanidine by *Veratrum grandiflorum*. Phytochemistry 15, 1391–1393.
Kaneko, K., Tanaka, M.W., Mitsuhashi, H., 1977. Dormantinol, a possible precursor in solanidine biosynthesis, from budding *Veratrum grandiflorum*. Phytochemistry 16, 1247–1251.
Kaneko, K., Watanabe, M., Taira, S., Mitsuhashi, H., 1972. Conversion of solanidine to jerveratrum alkaloids in *Veratrum grandiflorum*. Phytochemistry 11, 3199–3202.
Kaup, O., Gräfen, I., Zellermann, E.M., Eichenlaub, R., Gartemann, K.H., 2005. Identification of a tomatinase in the tomato-pathogenic actinomycete *Clavibacter michiganensis* subsp. *michiganensis* NCPPB382. Mol. Plant–Microbe Interact. 18, 1090–1098.
Krits, P., Fogelman, E., Ginzberg, I., 2007. Potato steroidal glycoalkaloid levels and the expression of key isoprenoid metabolic genes. Planta 227, 143–150.
Krokida, A., Delis, C., Geisler, K., Garagounis, C., Tsikou, D., Peña-Rodríguez, L.M., Katsarou, D., Field, B., Osbourn, A.E., Papadopoulou, K.K., 2013. A metabolic gene cluster in *Lotus japonicus* discloses novel enzyme functions and products in triterpene biosynthesis. New Phytol. 200, 675–690.
Kuhn, R., Low, I., 1954. Die Konstitution des Solanins. Angew. Chem. 66, 639–640.
Ma, R., Fontaine, T.D., 1950. Identification of the sugars in crystalline tomatin. Arch. Biochem. 27, 461–462.
McCue, K.F., Allen, P.V., Shepherd, L.V.T., Blake, A., Maccree, M.M., Rockhold, D.R., Novy, R.G., Stewart, D., Davies, H.V., Belknap, W.R., 2007. Potato glycosterol rhamnosyltransferase, the terminal step in triose side-chain biosynthesis. Phytochemistry 68, 327–334.
McCue, K.F., Allen, P.V., Shepherd, L.V.T., Blake, A., Whitworth, J., Maccree, M.M., Rockhold, D.R., Stewart, D., Davies, H.V., Belknap, W.R., 2006. The primary in vivo steroidal alkaloid glucosyltransferase from potato. Phytochemistry 67, 1590–1597.
McCue, K.F., Shepherd, L.V.T., Allen, P.V., Maccree, M.M., Rockhold, D.R., Corsini, D.L., Davies, H.V., Belknap, W.R., 2005. Metabolic compensation of steroidal glycoalkaloid biosynthesis in transgenic potato tubers: using reverse genetics to confirm the in vivo enzyme function of a steroidal alkaloid galactosyltransferase. Plant Sci. 168, 267–273.
Mennella, G., Lo Scalzo, R., Fibiani, M., D'Alessandro, A., Francese, G., Toppino, L., Acciarri, N., de Almeida, A.E., Rotino, G.L., 2012. Chemical and bioactive quality traits during fruit ripening in eggplant (*S. melongena* L.) and allied species. J. Agric. Food Chem. 60, 11821–11831.
Milner, S.E., Brunton, N.P., Jones, P.W., O'Brien, N.M., Collins, S.G., Maguire, A.R., 2011. Bioactivities of glycoalkaloids and their aglycones from *Solanum* species. J. Agric. Food Chem. 59, 3454–3484.
Mintz-Oron, S., Mandel, T., Rogachev, I., Feldberg, L., Lotan, O., Yativ, M., Wang, Z., Jetter, R., Venger, I., Adato, A., Aharoni, A., 2008. Gene expression and metabolism in tomato fruit surface tissues. Plant Physiol. 147, 823–851.
Moco, S., Bino, R.J., Vorst, O., Verhoeven, H.A., de Groot, J., van Beek, T.A., Vervoort, J., de Vos, C.H.R., 2006. A liquid chromatography–mass spectrometry-based metabolome database for tomato. Plant Physiol. 141, 1205–1218.
Moehs, C.P., Allen, P.V., Friedman, M., Belknap, W.R., 1997. Cloning and expression of solanidine UDP-glucose glucosyltransferase from potato. Plant J. 11, 227–236.
Mueller, L.A., Solow, T.H., Taylor, N., Skwarecki, B., Buels, R., Binns, J., Lin, C., Wright, M.H., Ahrens, R., Wang, Y., Herbst, E.V., Keyder, E.R., Menda, N., Zamir, D., Tanksley, S.D., 2005. The SOL Genomics Network. A comparative resource for *Solanaceae* biology and beyond. Plant Physiol. 138, 1310–1317.
Nes, W.D., 2011. Biosynthesis of cholesterol and other sterols. Chem. Rev. 111, 6423–6451.
Nohara, T., Ikeda, T., Fujiwara, Y., Matsushita, S., Noguchi, E., Yoshimitsu, H., Ono, M., 2006. Physiological functions of solanaceous and tomato steroidal glycosides. J. Nat. Med. 61, 1–13.
Ohyama, K., Okawa, A., Moriuchi, Y., Fujimoto, Y., 2013. Biosynthesis of steroidal alkaloids in *Solanaceae* plants: involvement of an aldehyde intermediate during C-26 amination. Phytochemistry 89, 26–31.
Ohyama, K., Suzuki, M., Kikuchi, J., Saito, K., Muranaka, T., 2009. Dual biosynthetic pathways to phytosterol via cycloartenol and lanosterol in Arabidopsis. Proc. Natl. Acad. Sci. U.S.A. 106, 725–730.
Ökmen, B., Etalo, D.W., Joosten, M.H., Bouwmeester, H.J., de Vos, R.C.H., Collemare, J., de Wit, P.J.G.M., 2013. Detoxification of α-tomatine by *Cladosporium fulvum* is required for full virulence on tomato. New Phytol. 198, 1203–1214.
Olmstead, R.G., Bohs, L., Migid, H.A., Santiago-Valentin, E., Garcia, V.F., Collier, S.M., 2008. A molecular phylogeny of the *Solanaceae*. Taxon 57, 1159–1181.
Orzaez, D., Medina, A., Torre, S., Fernández-Moreno, J.P., Rambla, J.L., Fernández-Del-Carmen, A., Butelli, E., Martin, C., Granell, A., 2009. A visual reporter system for virus-induced gene silencing in tomato fruit based on anthocyanin accumulation. Plant Physiol. 150, 1122–1134.
Osbourn, A., 2010a. Secondary metabolic gene clusters: evolutionary toolkits for chemical innovation. Trends Genet. 26, 449–457.
Osbourn, A., 2010b. Gene clusters for secondary metabolic pathways: an emerging theme in plant biology. Plant Physiol. 154, 531–535.
Patel, K., Gadewar, M., Tahilyani, V., Patel, D.K., 2012. A review on pharmacological and analytical aspects of diosgenin: a concise report. Nat. Prod. Bioprospect. 2, 46–52.
Petersson, E.V., Nahar, N., Dahlin, P., Broberg, A., Tröger, R., Dutta, P.C., Jonsson, L., Sitbon, F., 2013. Conversion of exogenous cholesterol into glycoalkaloids in potato shoots, using two methods for sterol solubilisation. PLoS One 8, e82955.
Qi, X., Bakht, S., Leggett, M., Maxwell, C., Melton, R., Osbourn, A., 2004. A gene cluster for secondary metabolism in oat: implications for the evolution of metabolic diversity in plants. Proc. Natl. Acad. Sci. U.S.A. 101, 8233–8238.
Roddick, J.G., 1976. Intracellular distribution of the steroidal glycoalkaloid α-tomatine in *Lycopersicon esculentum* fruit. Phytochemistry 16, 475–477.
Roddick, J.G., 1977. Subcellular localization of steroidal glycoalkaloids in vegetative organs of *Lycopersicon esculentum* and *Solanum tuberosum*. Phytochemistry 16, 805–807.
Roddick, J.G., 1996. Steroidal glycoalkaloids: nature and consequences of bioactivity. Adv. Exp. Med. Biol. 404, 277–295.
Sánchez-Mata, M.C., Yokoyama, W.E., Hong, Y.J., Prohens, J., 2010. α-Solasonine and α-solamargine contents of gboma (*Solanum macrocarpon* L.) and scarlet (*Solanum aethiopicum* L.) eggplants. J. Agric. Food Chem. 58, 5502–5508.
Sawai, S., Ohyama, K., Yasumoto, S., Seki, H., Sakuma, T., Yamamoto, T., Takebayashi, Y., Kojima, M., Sakakibara, H., Aoki, T., Muranaka, T., Saito, K., Umemoto, N., 2014. Sterol side chain reductase 2 is a key enzyme in the biosynthesis of cholesterol, the common precursor of toxic steroidal glycoalkaloids in potato. Plant Cell 26, 3763–3774.
Schakirov, R., Yunusov, M.S., 1990. Steroidal alkaloids. Nat. Prod. Rep. 7, 557–564.
Shakya, R., Navarre, D.A., 2008. LC–MS analysis of solanidane glycoalkaloid diversity among tubers of four wild potato species and three cultivars (*Solanum tuberosum*). J. Agric. Food Chem. 56, 6949–6958.
Suzuki, M., Muranaka, T., 2007. Molecular genetics of plant sterol backbone synthesis. Lipids 42, 47–54.

Swain, A.P., Fitzpatrick, T.J., Talley, E.A., Herb, S.F., Osman, S.F., 1978. Enzymatic hydrolysis of α-chaconine and α-solanine. Phytochemistry 17, 8–9.
Swaminathan, S., Morrone, D., Wang, Q., Fulton, D.B., Peters, R.J., 2009. CYP76M7 is an ent-cassadiene C11α-hydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice. Plant Cell 21, 3315–3325.
Takos, A.M., Knudsen, C., Lai, D., Kannangara, R., Mikkelsen, L., Motawia, M.S., Olsen, C.E., Sato, S., Tabata, S., Jørgensen, K., Møller, B.L., Rook, F., 2011. Genomic clustering of cyanogenic glucoside biosynthetic genes aids their identification in *Lotus japonicus* and suggests the repeated evolution of this chemical defence pathway. Plant J. 68, 273–286.
Umemoto, N., Sasaki, K., 2012. Protein having glycoalkaloid biosynthetic enzyme activity and gene encoding the same. U.S. patent application 2012/0159676 A1.
Wegel, E., Koumproglou, R., Shaw, P., Osbourn, A., 2009. Cell type-specific chromatin decondensation of a metabolic gene cluster in oats. Plant Cell 21, 3926–3936.
Wilderman, P.R., Xu, M., Jin, Y., Coates, R.M., Peters, R.J., 2004. Identification of syn-pimara-7,15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis. Plant Physiol. 135, 2098–2105.
Winzer, T., Gazda, V., He, Z., Kaminski, F., Kern, M., Larson, T.R., Li, Y., Meade, F., Teodor, R., Vaistij, F.E., Walker, C., Bowser, T.A., Graham, I.A., 2012. A *Papaver somniferum* 10-gene cluster for synthesis of the anticancer alkaloid noscapine. Science 336, 1704–1708.
Yahara, S., Uda, N., Nohara, T., 1996. Lycoperosides A–C, three steroisomeric 23-acetoxyspirosolan-3β-ol β-lycotetraosides from *Lycopersicon esculetum*. Phytochemistry 42, 169–172.
Yahara, S., Uda, N., Yoshio, E., Yae, E., 2004. Steroidal alkaloid glycosides from tomato (*Lycopersicon esculentum*). J. Nat. Prod. 67, 500–502.
Yamanaka, T., Vincken, J.P., Zuilhof, H., Legger, A., Takada, N., Gruppen, H., 2009. C22 isomerization in α-tomatine-to-esculeoside A conversion during tomato ripening is driven by C27 hydroxylation of triterpenoidal skeleton. J. Agric. Food Chem. 57, 3786–3791.
Zwenger, C., Kind, A., 1861. Über das Solanin und dessen Spaltungsprodukte. Ann. Chem. Pharm. 118, 129–151.

Pablo D. Cárdenas obtained his degree in Bioengineering and Master in Biochemistry and Bioinformatics from the University of Concepción, Chile. He is currently a PhD student from the Hebrew University of Jerusalem in a collaboration project with the Weizmann Institute of Science. He investigates the transcriptional regulation and metabolism of steroidal alkaloids in Solanaceae and their potential as sources for pathogen-resistance in plants.

Prashant D. Sonawane obtained his Ph.D. in Biotechnology from CSIR-National Chemical Laboratory, India (affiliated to University of Pune, India) in 2014. In his Ph.D. Prashant worked on the cinnamoyl CoA reductase (CCR) enzyme from Leucaena leucocephala. Prashant is currently working as postdoctoral fellow at Weizmann Institute of Science, Rehovot, Israel under the guidance of Prof. Asaph Aharoni. His research is focused on genetic regulation of steroidal glycoalkaloids metabolism in Solanaceae crops.

Uwe Heinig was born in 1979 and received his Diploma in chemistry from the Technical University Darmstadt (Germany) in 2006. After this he did his PhD in the group of Dr. Stefan Jennewein at the Fraunhofer Institute for Molecular Biology and Applied Ecology in Aachen (Germany) studying evolution of Taxol biosynthesis in plant associated endophytic fungi. Currently he is working as a researcher in Prof. Asaph Aharoni's laboratory in the Department of Plant Sciences at the Weizmann Institute of Science in Rehovot, Israel. His major research interests are the isolation and characterization of natural products, mass spectrometry based metabolomics and the examination of secondary metabolite biosynthetic pathways.

Samuel Bocobza obtained his Ph.D. in Life Sciences from the Ben Gurion University of the Negev, Israel, in 2010. His research focus was on how riboswitches regulate gene expression and primary metabolism in plants. Since then and until now, he is conducting postdoctoral research under the guidance of Prof. Asaph Aharoni at the Weizmann Institute of Science, Rehovot, Israel.

Saul Burdman obtained his Ph.D. in Agricultural Microbiology from the Hebrew University of Jerusalem in 2001. In 2003, after postdoctoral research under the guidance of Prof. Pamela Ronald at University of California, Davis, he returned to the Hebrew University where he got a researcher position at the Department of Plant Pathology and Microbiology. Currently, S. Burdman is associate professor in Phytopathology. His research focuses on basic and applied aspects of interactions between plants and phytopathogenic bacteria. He teaches courses on Plant Pathology, Plant-Microbe Interactions and Phytobacteriology. S. Burdman is member of the Israeli Phytopathological Society, the Israeli Society for Microbiology, the American Phytopathological Society, the American Society for Microbiology and the International Society for Molecular Plant-Microbe Interactions.

Asaph Aharoni conducted his B.Sc. and M.Sc. studies at the Faculty of Agriculture of the Hebrew University of Jerusalem, Israel. In 1996, he began his Ph.D. studies in the CPRO-DLO Institute in Wageningen, The Netherlands. His doctorate work combined functional genomics, molecular biology and biochemistry for studying fruit development and ripening. The main focus of the research was fruit flavor biosynthesis. His post-doctoral position was conducted in Plant Research International (Wageningen, The Netherlands), studying the genetic regulation of metabolic pathways in plants. The work involved the use of transposon tagging and metabolic profiling for discovery of genes controlling the biosynthesis of plant surface components. Also during his postdoc Asaph initiated and set-up metabolomics technologies as a tool for the comprehensive analyses of small molecules in plants. His work in this field had an important contribution to the implementation of these technologies for the study of plant biology and metabolism, particularly in the case of secondary metabolites. In August 2004, he started his own research group in the department of Plant Sciences at the Weizmann Institute in Israel. Also today, his lab combines expertise in molecular biology, analytical chemistry and computational biology. The topic of his current research activity is the genetic regulation of metabolic pathways and its co-ordination with developmental and stress response programs in plant biology. The lab investigates several aspects of metabolic regulation including transport, transcriptional and post transcriptional control.

Chapter 12

Analysis of Steroidal Alkaloids and Saponins in *Solanaceae* Plant Extracts Using UPLC-qTOF Mass Spectrometry

Uwe Heinig and Asaph Aharoni

Abstract

Plants of the *Solanaceae* family are renowned for the production of cholesterol-derived steroidal glycosides, including the nitrogen containing glycoalkaloids and steroidal saponins. In this chapter we describe the use of UPLC (Ultra Performance Liquid Chromatography) coupled with qTOF (Quadrupole Time-of-Flight) mass spectrometry for profiling of these two large classes of semipolar metabolites. The presented method includes an optimized sample preparation protocol, a procedure for high resolution chromatographic separation and metabolite detection using the TOF mass spectrometer which provides high resolution and mass accuracy. A detailed description for non-targeted data analysis and a strategy for putative identification of steroidal glycosides from complex extracts based on interpretation of mass fragmentation patterns is also provided. The described methodology allows profiling and putative identification of multiple steroidal glycoside compounds from the assortment of *Solanaceae* species producing these molecules.

Key words UPLC/qTOF mass spectrometry, *Solanaceae*, Steroidal alkaloids, Steroidal saponins, Metabolomics

1 Introduction

Plants of the *Solanaceae* family are well known for the production of numerous natural products, of which sapogenins are one major compound class. Sapogenins are plant terpenoid secondary metabolites, derived from the cytosolic mevalonic acid pathway. Two molecules of farnesyldiphosphate (FPP), formed from the universal terpene precursors isopentenyldiphosphate (IPP) and dimethylallyldiphosphate (DMAPP), are condensed to squalene followed by oxidation toward the precursor for all phytosteroids, 2,3-oxidosqualene. Oxidosqualene is then either converted to triterpenes (C30), such as amyrins or to lanosterol the precursor of cholesterol (C27) [1].

Manuel Rodríguez-Concepción (ed.), *Plant Isoprenoids: Methods and Protocols*, Methods in Molecular Biology, vol. 1153, DOI 10.1007/978-1-4939-0606-2_12,

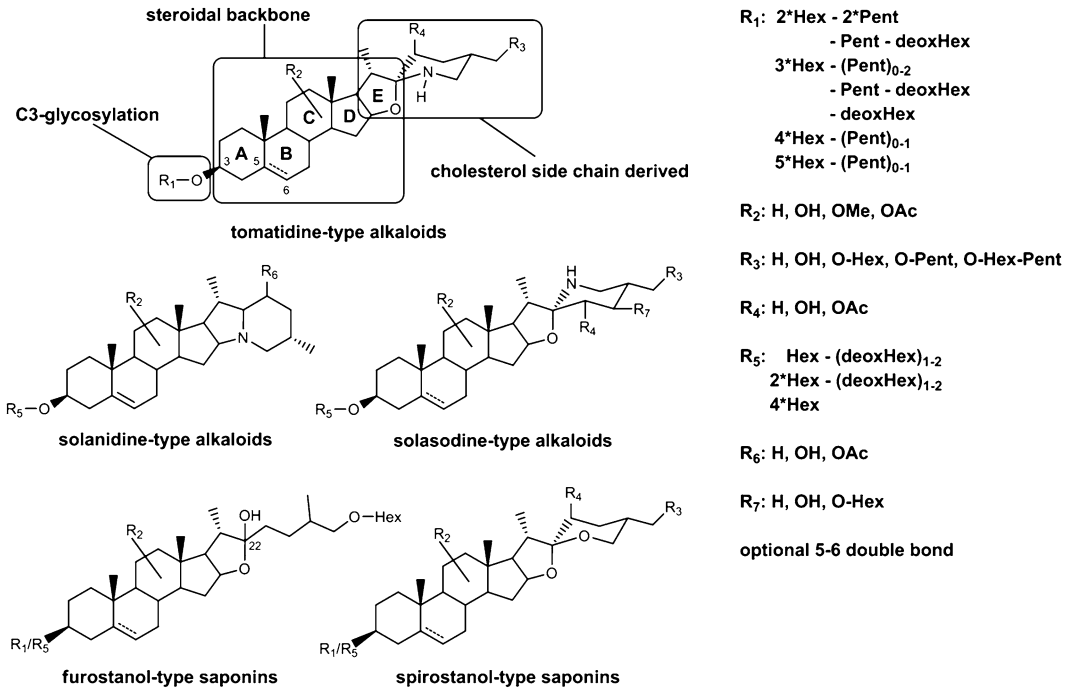


Fig. 1 Diversity of penta-cyclic steroidal saponins and alkaloids. All molecules of these two classes generally consist of a penta-cyclic steroidal backbone, a modified side chain (cyclic, open-chain, nitrogen-containing in case of alkaloids) originated from the side chain of cholesterol and are highly glycosylated at the hydroxyl-group at C3 ($R_{1,5}$). Further substitutions like hydroxylations, acetylations, and O-glycosylations can occur on the backbone or more likely on the side-chain ($R_{2,3,4,6,7}$). *Hex* hexose, *Pent* pentose, *deoxHex* deoxyhexose

In *Solanaceae* species, including tomato (*Solanum lycopersicum*), potato (*Solanum tuberosum*), eggplant (*Solanum melongena*), or black nightshade (*Solanum nigrum*), the predominant semipolar compounds detected are cholesterol-derived steroidal glycoalkaloids and steroidal saponins [2, 3]. Their biosynthesis starting from cholesterol involves hydroxylations, oxidations and in case of the alkaloids a transamination step, followed by extensive glycosylation, resulting in the observed diversity of detected end products. In Fig. 1, the five major aglycone structures observed in the *Solanaceae* are presented: tomatidine, solasodine and solanine (alkaloids) and furostanol and spirostanol (saponins). These carbon backbones can be substituted in different positions (R_{1-7} in Fig. 1) [2, 4–6].

By and large *Solanaceae* steroidal glycosides consist of three major structural parts: (a) a penta-cyclic steroidal backbone (rings marked with A–E in Fig. 1) that can contain a double bond at C5, (b) a modified side-chain derived from C20–C27 cholesterol side-chain, and (c) a glycosyl-moiety at the C3 hydroxyl-group (upper structure in Fig. 1). The occurrence of specific compounds thereby depends strongly on the plant species, the analyzed tissue and the developmental stage (e.g., fruit maturation in tomato [4, 6]).

In this chapter we describe a UPLC/qTOF-MS method for simultaneous chromatographic separation and mass spectrometric detection of the major steroidal alkaloids and saponins found in the *Solanaceae*. Besides sample preparation and chromatography, we focused here on the putative identification of compounds using a non-targeted approach based on accurate mass measurements and interpretation of mass fragmentation patterns.

2 Materials

2.1 Material and Reagents for Sample Preparation

1. Water, double deionized, from a Milli-Q purification system (Millipore), resistivity 18.2 MΩ cm, filtered through a 0.22-μm membrane filter.
2. Methanol, gradient grade for liquid chromatography (e.g., Merck KGaA).
3. Acetonitrile, gradient grade for liquid chromatography (e.g., Merck KGaA).
4. Liquid nitrogen for grinding and freezing plant samples.
5. Standards for QC (quality control) samples: L–Tryptophan (Sigma), L–Phenylalanine (Sigma), Chlorogenic acid (Fluka), Caffeic acid (Sigma), *p*-Coumaric acid (Sigma), Ferulic acid (Aldrich), Sinapic acid (Sigma), Rutin hydrate (Sigma), Quercetin dihydrate (Sigma), Tomatine (Apin), Naringenin (Fluka), Kaempferol (Fluka) (*see* **Note 1**).
6. Mortar and pestle or a ball mill (e.g., Retsch MM301).
7. Balance.
8. Screw-cap Polypropylene (PP) tubes (15 mL, e.g., Greiner) or 2 mL PP safe-lock eppendorf tubes for storage of tissue and extraction, suitable for centrifugation at 3,000 × *g* and 15,000 × *g*, respectively.
9. Ultrasonic bath.
10. Vortex.
11. Single-use sterile latex-free 1 mL syringes (e.g., 0.5 mm × 16 mm syringe, BD Plastipak™).
12. Single-use, 0.22 μm membrane syringe filters (e.g., 4 mm diameter PVDF (Polyvinylidenefluoride, Millex-GV) or 12 mm diameter PTFE (Polytetrafluoroethylene, PALL).
13. Amber-glass 2 mL autosampler vials and caps with a PTFE/Silicone septum. Use suitable 250 μl glass inserts for small extract volumes injection (*see* **Note 2**).

2.2 UPLC-qTOF and Data Analysis

1. UPLC/qTOF system: e.g., a UPLC Waters Acquity instrument connected in-line to a Synapt HDMS detector (tandem quadrupole/time-of-flight mass spectrometer).

The MS detector is equipped with an electrospray ion source (ESI). The Synapt HDMS system is operated in the standard qTOF mode, without using the ion mobility capabilities (*see* **Note 3**).

2. UPLC BEH C18 column (Waters Acquity), 100 × 2.1 mm i.d., 1.7 μm, with a column pre-filter.
3. Solvents for chromatography: mobile phase A, 5 % acetonitrile (ACN) in ddH_2O + 0.1 % formic acid (FA); mobile phase B, ACN + 0.1 % FA.
4. Washing solutions: strong needle wash, 80 % methanol (to remove organic components); weak needle wash, 5 % acetonitrile (re-equilibration to chromatographic starting conditions); seal wash, 10 % methanol.
5. MassLynx 4.1 instrument software (Waters).
6. DataBridge program (Waters).
7. XCMS package for R for peak picking and retention time correction between replicate samples [7].
8. Software for mass peak clustering and further statistical analysis, e.g., Principal component analysis (PCA).

3 Methods

3.1 Sample Preparation

The most widely used method for extraction of semipolar compounds for LC-MS analysis, such as glycoalkaloids and saponins, is sample preparation using acidic methanol [8, 9]. A final methanol content of 75 and 0.1 % formic acid (FA) was found to be a most efficient extraction solution when applied in various plant species and different tissues (*see* **Note 4**). A detailed description of the sample preparation procedure for tomato can also be found in [10] and is carried out as follows:

1. Homogenize frozen tissue using mortar and pestle pre-cooled in liquid nitrogen, or when starting with limited amount of material add a metal ball into the 2 mL eppendorf tube which contains the tissue, place the tubes into the pre-cooled ball mill tray and disrupt the material by shaking with 20 Hz for 2 min. After shaking, immediately place the tray into liquid nitrogen. Make sure that samples do not begin to thaw during the whole procedure.
2. Prepare appropriate number of clean tubes (15 or 2 mL), pre-cool in liquid nitrogen and weigh them on a balance.
3. Transfer the powder from **step 1** to 15 mL PP screw-cap tube (>350 mg of material) or in case of using less than 350 mg into a 2 mL eppendorf tube (*see* **Note 5**).
4. Weigh the biomaterial containing tubes and calculate the weight of the samples. In case of using the ball mill, determine the weight before adding the metal ball and grinding.

5. Add the required amount of acidified MeOH or MeOH/water (ratio *see* **Note 4**) to the frozen samples, to a final MeOH concentration of 75 % (*see* **Note 6**).
6. Vortex until all frozen powder is completely re-suspended in the solvent.
7. Sonicate for 20 min.
8. Vortex for several seconds.
9. Remove debris by centrifugation for 10 min with 3,000 × *g* when using 15 mL tubes or 15,000 × *g* in case of using 2 mL eppendorf tubes.
10. Carefully transfer the supernatant into a fresh tube.
11. Filter extract through a 0.22 μm PTFE (or PVDF) disposable filter into either another 2 mL eppendorf tube or directly into an autosampler vial using a 1 mL syringe. When volume is limited (less than 500 μL), filter an aliquot of the supernatant into an autosampler vial insert (approximately 150 μL). Use not less than 50 μL and make sure that no air bubbles remain at the bottom of the insert. Close vials tightly with PTFE/Silicone septum-caps to avoid evaporation of solvent. In case of storage of the extracts for later injection at −20 °C, check if precipitates are formed and eventually repeat filtering directly before injection.
12. As a blank sample use 75 % MeOH/ddH_2O supplemented with 0.1 % FA.
13. Place vials for injection into the UPLC autosampler, cooled to 12 °C.
14. The remaining extract (e.g., for MS/MS analysis) should be stored at −20 °C immediately.

3.2 UPLC/qTOF-MS Analysis

Chromatographic conditions are chosen according to the amphiphilic nature of the analytes (glycosylated cholesterol-derived molecules). The polarity during chromatographic separation is thereby mainly determined by the sugar moieties, hence, gradient elution starting with 5 % ACN in ddH_2O supplemented with 0.1 % FA to 31.6 % ACN (= 28 % B) over 22 min was found to give good separation and peak shape [10]. Steroidal alkaloids and saponins elute typically between 10 min and the end of the chromatogram. Thereby they are well separated from more polar compounds, like for example organic acids which elute earlier.

Column preparation, system stability testing and chromatographic conditions for analysis are listed below.

1. Equilibrate new columns in 50 % A/50 % B for at least 60 min, followed by equilibration to chromatographic starting conditions (5–10 column volumes A). Columns already used and stored in 50 % A/50 % B are equilibrated with at least 5 column volumes of A.

2. Performance of the system (stability of retention times, signal intensities, mass accuracy) can be tested by several injections of standard QC-mix (*see* **Notes 1** and 7).
3. UPLC conditions: column oven, 35 °C; flow rate, 0.3 mL/min; autosampler temperature, 12 °C; gradient, initial conditions 100 % A, 22 min 72 % A, 22.5 min 60 % A, 23 min 0 % A, 26.5 min 0 % A, 27 min 100 % A, 28 min 100 % A.
4. Inject 4–10 μL of biological sample (*see* **Note 8**).
5. Wash needle between injections with 200 μL strong needle wash solution and 600 μL weak needle wash (*see* **item 4** of Subheading 2.2) solution.
6. For longer data-sets, the QC-mix should be injected at least every ten samples to control the stability of the system (*see* **step 2** of Subheading 3.2).

For mass spectrometric detection, we use the Synapt HDMS detector equipped with an ESI source. "Soft" ESI ionization allows detection of intact metabolites (molecular ions), necessary for calculation of elemental composition and putative identification in a non-targeted manner. The mass detector part (Time of Flight, TOF) is operated in V mode with a mass resolution of 9,000. Spectra are recorded from *m/z* 50 to 1,500 with scan duration of 0.25 s and an interscan delay of 0.02 s in centroid mode. Mass spectrometer parameters are set to: capillary voltage = 3.4 kV (ESI-: 3 kV), cone voltage = 24 eV (ESI-: 28 eV), source temperature = 125 °C, desolvation temperature = 275 °C, cone gas flow = 25 L/h, desolvation gas flow = 650 L/h and collision energy = 4 eV. Argon is used as a collision gas and the mass spectrometer is calibrated with leucine enkaphalin ($[M+H]^+$: *m/z* 556.2771, $[M-H]^-$: *m/z* 554.2620). In a second channel the collision energy is ramped from 10 to 30 eV in the positive mode and from 15 to 35 eV in the negative mode. For MS/MS experiments, product ion spectra of selected masses are recorded with various collision energies (15–50 eV) with scan duration of 0.4 s and an interscan delay of 0.02 s in the same mass range.

3.3 Data Analysis Workflow

Data analysis is performed in two different ways. Compounds like for example α-tomatine that are commercially available can be identified directly by comparison with an authentic standard. However most of the metabolites formed are not known and have to be putatively assigned in a non-targeted manner. Therefore, information obtained from the known compounds, like polarity hence retention on the column, ionization behavior, and especially compound class specific mass fragmentation are necessary for

putative identification. Non-targeted data analysis involves the following steps:

1. Control system stability manually by comparison of QC-control injections at the beginning, during, and at the end of the acquisition sequence (*see* **Note 9**).
2. Convert raw data files to NetCDF format using the MassLynx Databridge program and organize files into sample groups (replicate groups, e.g., Species 1, Species 2). Treat data acquired in positive and negative ionization mode separately. For further analysis use only NetCDF files created from the first channel of raw data files (first channel: mass data collision energy = 4 eV; second channel: mass data collision energy ramp; third channel: lock mass calibration; fourth channel: PDA detector, absorption spectra recorded from 210 to 500 nm).
3. Run mass peak detection and retention time correction with XCMS package for R [7]. Parameters thereby depend on the performance of chromatography as well as mass spectrometer specifications (*see* **Note 10**).
4. Perform quality control for XCMS using the MetaboQC program [11] or manually. XCMS creates different types of outputs, in our case a peak table including all detected mass traces, integrated intensity values and a number of extracted ion chromatograms (comparison between two replicate groups) that also show the borders for integration (*see* **Note 11**).
5. Perform mass peak clustering using software such as "CAMERA" for XCMS [12], MZ-mine [13], or Metalign [14].
6. Filter and sort the output data. Set for example intensity thresholds or sort according to retention time, mass or intensity.
7. Identify masses of peak groups/clusters (= chromatographic peaks) in raw data file.
8. Identify molecular ions of compounds; compare therefore positive and negative ionization datasets in order to avoid selection of the wrong ion, due to for example neutral losses during ionization or formation of adducts.
9. Calculate elemental composition of compounds using the elemental composition calculator of MassLynx 4.1. Molecular formulas are determined using accurate mass and isotopic pattern.
10. Search databases, e.g., the Natural product database or Scifinder for possible structures.
11. Verify putative structures by analysis of mass fragmentation patterns obtained by "collision energy ramp" or MS/MS experiments.

3.4 Identification of Steroidal Alkaloids and Saponins

Whereas the general procedure described above is especially useful to identify differences between samples according to all masses detected in the analysis, putative identification of steroidal alkaloids and saponins has to be done according to their mass spectral characteristics (Fig. 2). Thereby, the procedure can be divided into several steps, including (a) identification of the molecular ions of present compounds, (b) analysis of the glycosylation pattern, (c) analysis of substitution of the aglycones by determination of specific neutral losses and the fragmentation of the aglycones in order to distinguish between different types of saponins and alkaloids as presented in Fig. 3. We suggest doing the analysis in the following manner:

1. Identify molecular ions of compounds by comparison between the mass peak clustering analysis and raw data. It is highly recommended to compare data-sets obtained by positive and negative ionization. Steroidal glycoalkaloids tend to form formic acid adducts in the negative ionization mode ($[M-H+46]^-$), due the presence of formic acid in the solvent. In positive mode $[M+H]^+$ is the most commonly detected ion, although in case of substitutions on the aglycone $[M+H\text{-substituent}]^+$, e.g., minus H_2O or acetate, are observed, too. Steroidal saponins of the furostanol-type loose easily their 22-OH group resulting in exclusive detection of $[M+H-H_2O]^+$ in the positive mode, whereas $[M-H]^-$, and in minor amounts the FA adducts are observed, in the negative mode (Fig. 2). In low energy, doubly charged ions can occur ($M/2+H^+$).
2. Divide compounds into putative alkaloids (even mass, compounds contain one nitrogen atom) and saponins (odd mass, no nitrogen atom in the molecule) (*see* **Note 12**).
3. Filter detected compound ions for masses higher than the mass of possible aglycones, e.g., *m/z* 413 and higher for steroidal saponins (*see* **Note 13**).
4. Perform MS/MS analysis for all detected putative steroidal glycosides (alkaloids and saponins) using different collision energies in order to fragment only the glycosyl moieties (15–30 eV) or the entire molecule including the aglycone (40–50 eV). Especially for the analysis of neutral losses from the aglycone and glycosylation pattern, data obtained by applying the collision energy ramp (*see* Subheading 3.2) can be used (*see* **Note 14**).
5. Analyze glycosylation pattern of compounds. Typically steroidal glycosides contain up to seven glycosyl residues, which are mainly hexoses (galactosyl-, glucosyl-, loss of 162 Da), pentoses (xylosyl-, loss of 132 Da) or deoxyhexoses (rhamnosyl-, loss of 146 Da; Figs. 1 and 2d). In rare cases, other substituents occur (*see* **Note 15**). In the example shown in Fig. 2d, the glycosylation pattern of α-tomatine from *Solanum lycopersicum*

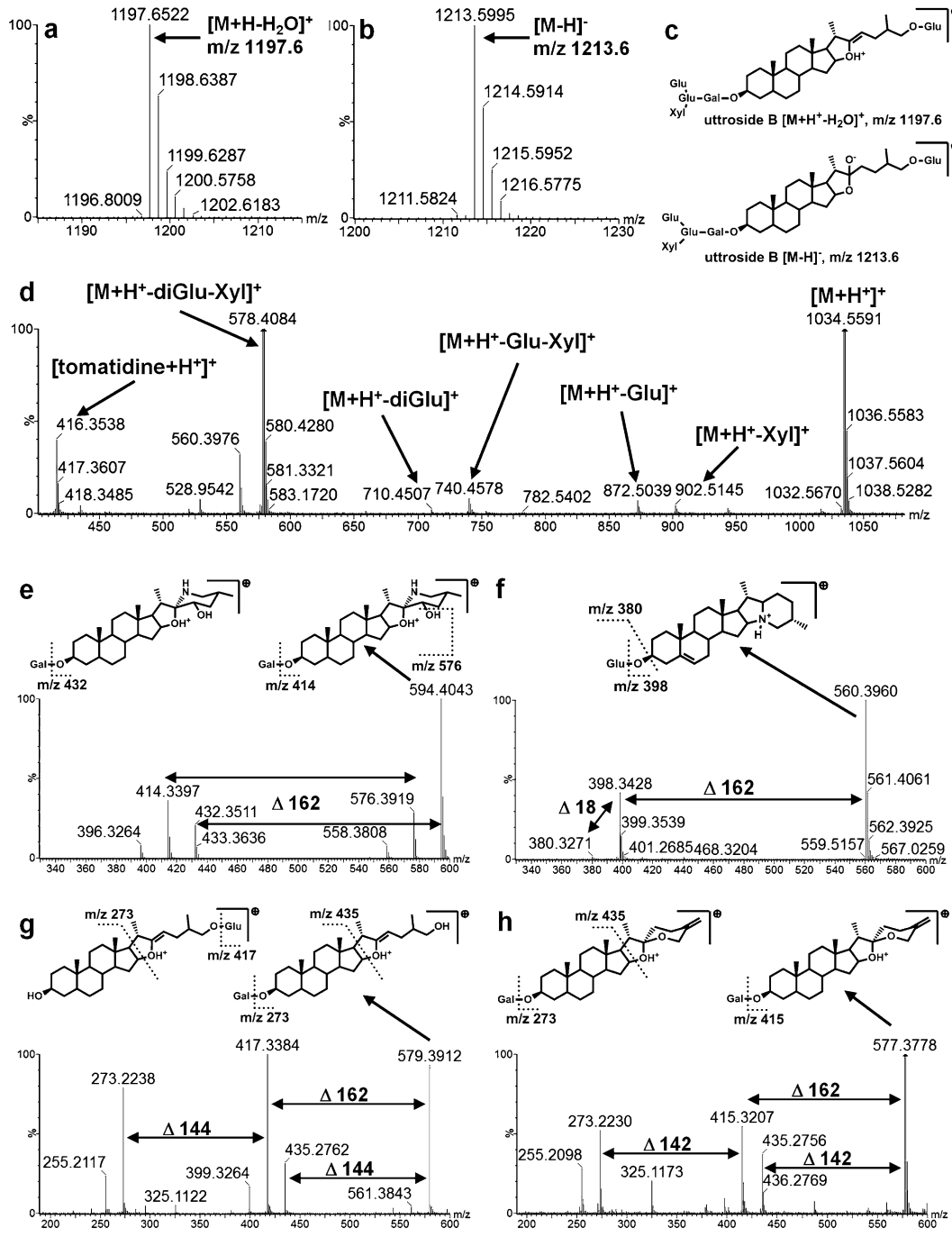


Fig. 2 Mass spectrometric characteristics of steroidal saponins and steroidal glycoalkaloids. (**a**) Molecular ion of uttroside B from *Solanum nigrum* detected in the positive ionization mode. (**b**) Molecular ion of uttroside B detected in the negative ionization mode. (**c**) structures of detected molecular ions. (**d**) fragmentation pattern of the C3-glycosyl-chain of α-tomatine from *Solanum lycopersicum.* Characteristic fragmentation patterns and structures of: (**e**) hydroxyl-tomatine as an example for tomatidine-type aglycones, (**f**) α-solanine from *Solanum tuberosum* as an example for solanidine-type aglycones, (**g**) uttroside B from *Solanum nigrum* as an example for furostanol-type saponins, and (**h**) a putative hydroxylated spirostanol-type saponin from *Capsicum* sp.; *m/z*: mass to charge, *M* molecular ion, *Glu* glucosyl, *Xyl* xylosyl, *Gal* galactosyl

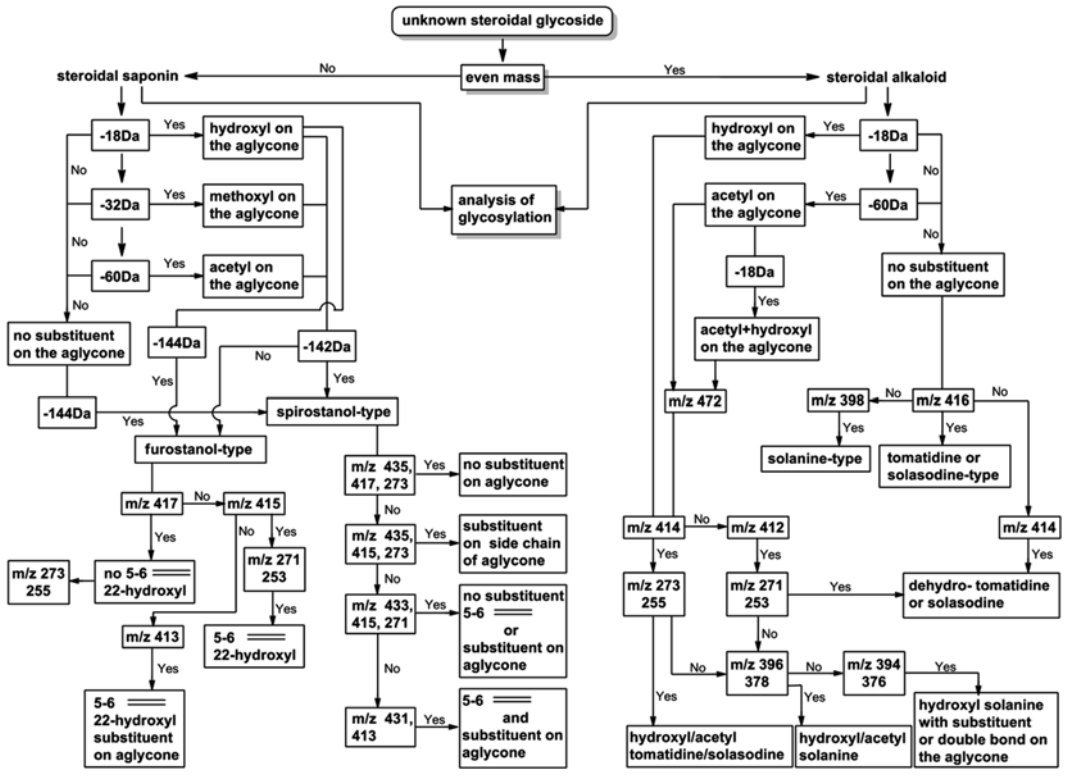


Fig. 3 Scheme for putative identification of steroidal saponins and alkaloids via interpretation of characteristic mass fragmentation patterns. Alkaloids and saponins can be first differentiated according to the occurrence of odd or even (one nitrogen containing compounds) molecular ion masses. Besides analysis of glycosylation pattern (mainly from C3) hydroxyl-, acetyl- and methoxyl-substitutions lead to specific neutral losses of 18, 32, and 60 Da (most common aglycone substituents, in rare cases also other groups can occur, e.g., carboxyl). Discrimination between different aglycone structures can be done either according to specific neutral losses (e.g., −144 Da for furostanol-type saponins, −142 Da for spirostanol-type saponins) or by analysis of the aglycone mass and masses of fragmentation products of these aglycones. *m/z* mass to charge, *Da* Dalton, = double bond

is determined to 3*hexose + 1*pentose. Loss of xylose as well as glucose directly from the molecular ion indicates that both sugar residues are located at the end of the glycosyl chain.

6. Identify characteristic neutral losses, besides sugars, that provide a hint to the substitution of the aglycone. As shown in Figs. 1 and 3, hydroxyl, acetyl and methoxyl substituents are common for steroidal glycosides. Neutral losses of 18, 32, or 60 Da from the molecular ion or a difference of 16 Da between positive and negative ionization (*see* **step 1** of Subheading 3.4) (Figs. 2 and 3) are characteristic for these substitutions.
7. To distinguish between different types of aglycones (Fig. 1), analyze the mass and the fragmentation pattern of the carbon backbone. In case of steroidal saponins, two major types are

formed in *Solanaceae* species, furostanol- and spirostanol-type saponins. Dependent on the substitution pattern determined before, specific neutral loss of the side-chain can be used for differentiation. If the putative compound does not have substitutions on the aglycone, a loss of 144 Da is characteristic for the spirostanol-type, resulting in fragment masses of *m/z* 435 and 433 which are the ions for rings A–D (Fig. 1) plus a hexose moiety at C3, without or with the 5–6 double bond. If the molecules are substituted on the aglycone side-chain a characteristic loss of 142 Da is observed, resulting in fragment masses of *m/z* 435, 433, or 431 depending on the number and position of the substitution (Figs. 2h and 3) [15]. For confirmation of the putative assignment, marker ions *m/z* 417–413 (aglycone without sugars) and *m/z* 273/271 (ions of rings A–D) can be used to determine position of the substitution and the possibility of a double bond in the molecule. Furostanol-type compounds typically loose water by loss of the free 22-OH group (Fig. 2a–c). Hence water loss and loss of the resulting side chain (−144 Da, Fig. 2g) is a clear indication for the furostanol-type. As described for spirostanol-type saponins the number of substitutions and saturation of the aglycone can be determined by analysis of the fragments of the aglycone (Fig. 3).

8. Steroidal alkaloids show a similar fragmentation behavior, although loss of the side chain is not observed as in the case of the saponins. When compounds are not substituted, aglycone masses of *m/z* 416 and *m/z* 398 (carbon backbone without sugars) are detected for tomatidine/solasodine-type and solanine-type alkaloids, respectively (Fig. 2d, f). If *m/z* 414 is observed as the dominant aglycone-fragment ion for non-substituted molecules together with fragment ions of *m/z* 271 and 253, the compounds are putative dehydro-tomatidine/solasodine-type alkaloids.
9. In the case of substituted compounds (e.g., acetyl or hydroxyl), elemental compositions and hence the detected ions, can be identical for different types of structures (for example hydroxyl-solanine and dehydro-tomatidine). To assign the correct structure, combinations of mass fragments have to be taken into consideration as well as their relative abundance.
10. These ions include typical aglycone masses like *m/z* 414/412 (hydroxyl-tomatidine after water loss/hydroxyl-dehydro-tomatidine after water loss, but also possibly hydroxyl-solanine) in combination with further fragments of these, for example *m/z* 273, 271 (A–D ring fragment), *m/z* 255, 253 (A–D ring fragment minus water). Characteristic ions of solanidene-type molecules are for example *m/z* 396/394 (desaturated solanidene) and *m/z* 378/376 (desaturated solanidene minus water).

Although these last mentioned ions can occur in tomatidine/solasodine-alkaloids' spectra, too, higher relative intensity compared to *m/z* 414, 412 supports a solanidene-type aglycone (*see* **Note 16**).

11. The scheme shown in Fig. 3 summarizes the differences in fragmentation behavior that can be used for and efficient putative assignment of unknown steroidal glycosides structures (saponins and alkaloids) according to the proposed stepwise analysis strategy of mass fragmentation spectra.

4 Notes

1. Prepare stock solutions with a concentration of 1 mg/mL in MeOH for all standards except tomatine and tryptophan. Tomatine can be dissolved in MeOH in a concentration of 0.5 mg/mL by sonication. Tryptophan is dissolved in 80 % MeOH supplemented with 2 % FA by sonication. It is recommended to start with a higher percentage of water, because of the poor solubility of the compound in MeOH. A QC-mix is prepared by combination of equal amounts of all stock solutions (final concentration 83 μg/mL, tomatine 42 μg/mL) and is stored at −20°C. Prior to injection this mix is further diluted 1:10 with MeOH.
2. All plastic materials have to be resistant to 75 % MeOH/0.1 % FA.
3. Ion mobility separation technology can be used to separate isobaric compounds (co-eluting compounds, e.g., isomers) and hence to reduce noise and complexity of MS and MS/MS spectra that is due to overlapping isotope and fragmentation patterns.
4. The ratio of MeOH/ddH_2O should be adjusted according to the biological material used. For tissues such as tomato fruit that have a high water content, addition of three volumes of MeOH was found to be optimal to obtain a final MeOH/water ratio of about 75 %. Furthermore some tissues are rich in organic acids, hence no additional acidification is necessary [10].
5. Pre-cool spatula in liquid nitrogen to avoid sticking and thawing of material on it.
6. A solvent/biomaterial ratio of 3:1 is used normally. Depending on the tissue and the abundance of the compounds of interest this ratio can be changed (e.g., 2:1 to increase concentration of trace metabolites).
7. To control system stability we use a 12-compound standard mix (8 μg/mL each, tomatine 4 μg/mL). The compounds are selected substances known from plants and cover the polarity

range of the chromatographic run from 2 min (L-phenylalanine) to 19.5 min (kaempferol) as well as a broad mass range (e.g., *m/z* 179.03 for caffeic acid, *m/z* 1,078.5 for tomatine $[M-H+FA]^-$ in negative ionization mode). This allows its use for stabilization of retention times and control of mass accuracy. Furthermore, the column performance over time can be monitored by comparison of peak resolution and shape of QC-mix injections on new and old columns.

8. Injection of higher volumes is not recommended because it leads to bad peak shape and loss in resolution. For analysis of low abundant compounds, concentrate the extract, e.g., via lyophilization.
9. Compare QC-mix samples to control stability of retention times, intensity, and mass accuracy.
10. An example of a script for peak detection and processing is shown below. Different parameters thereby depend strongly on the quality of the raw data or on the general specifications of the analytical setup. Most of these parameters are defined in line 1, including mass accuracy (20 ppm), peakwidth (5–20 s), prefilter (three consecutive scans have to display an intensity higher than 15 IPS), the signal to noise threshold and the integration method. For example, mass accuracy could be lowered when using an instrument with a higher resolution. When using HPLC instead of UPLC peaks will be much wider. Also further steps depend on the quality of the entire dataset, because grouping (group) and retention time correction (retcor) is done for masses that appear in all or many samples (chromatograms). In cases of large retention time shifts from sample to sample; this will result in a low number of peaks in the final output. The last line in the given example script defines the given output after peak picking and processing of the dataset.

Example-Script for XCMS Peak Picking

- Name = xcmsSet(method="centWave", ppm=20, peakwidth=c(5,20), prefilter=c(3,15), snthresh=10, integrate=1, mzdiff=-0.001)
- Name = group(Name, bw = 5, mzwid = 0.06, minsamp = 1)
- Name = retcor(Name_cor, method = "loess", plottype = "none", span = 2)
- Name_cor = group(Name_cor, bw = 4, mzwid = 0.06, minsamp = 1)
- Name_fill = fillPeaks(Name_cor)
- diffreport(Name_fill, 'SG1', 'SG2', 'Name', sortpval = TRUE, 200, metlin = -0.05)

11. Typically an XCMS output table contains between several hundred and several thousand mass traces. If this number is significantly lower, the parameters need to be adjusted in cases when in the chromatogram more compounds/masses are clearly observed. In the table generated each line represents a mass trace and the intensity in each sample for this mass is given. The replicate to replicate reproducibility (e.g., WT1, WT2, WT3) can be checked by comparison of these values. Differences thereby can be due to different extraction efficiency between samples, but also to chromatographic variations between injections. This variance can lead to missing of a mass trace at a retention time in the chromatogram of a sample in which the peak is shifted and hence does not fulfill the criteria defined for XCMS. Analysis of the extracted ion chromatograms allows control of integration borders as well as alignment of mass trace peaks between the samples.
12. Steroidal alkaloids typically contain one nitrogen atom; hence they show an even mass. This observation is of course also true for compounds containing three or five nitrogen (or odd mass for two nitrogen atoms). Therefore this differentiation of glycoalkaloids and saponins is preliminary and putative assignment always has to be validated by calculation of elemental composition and analysis of the fragmentation pattern.
13. Filtering of data for masses higher than the possible aglycone removes all mass peaks that originate from fragmentation that already occurred at low collision energy and can reduce the complexity of the XCMS output table significantly.
14. For high abundant compounds, like for example uttroside B from *Solanum nigrum* [16] or α-tomatine from tomato [4], analysis of the collision energy ramp data can be sufficient for putative identification, since here the signal to noise ratio is sufficient to assign all fragments to the compounds. For many low intensity signals, mass fragments of co-eluting substances interfere with the compound specific ions, e.g., for two differently glycosylated but co-eluting alkaloids it is not possible to determine the origin of the sugar residue loss. In these cases only MS/MS data can lead to a putative identification.
15. Rare substituents on the glycosyl-chain include for example organic acids (e.g., ferulic acid, coumaric acid). Differentiation between losses of sugar residues and acids can be done either according to their accurate mass or by looking for glycosylated aromatic acids in the product ion spectrum, that are formed via alternative fragmentation with the steroidal molecule part as a neutral loss. Another indication for the presence of aromatic substituents is a higher double bond equivalent than usual when calculating the elemental composition using MassLynx 4.1 elemental composition calculator.

16. Although shown to be a powerful tool for putative assignment of steroidal glycosides the presented method of mass spectra interpretation cannot completely solve the structure of newly detected molecules. The correct order of sugar residues can only be partially determined and the exact position of substituents/double bonds on the aglycones becomes not completely clear. However, such information is most helpful in discriminating between the (often) large numbers of possible structures obtained for a given calculated elemental composition.

Acknowledgements

We thank A. Tishbee for operating the UPLC/qTOF instrument, I. Rogachev and S. Meir for assistance in XCMS analysis, mass spectra interpretation and putative assignment of alkaloids from potato and tomato and the European Research Council (SAMIT-FP7 program) for supporting the work in the A. A. laboratory. A. A. is an incumbent of the Peter J. Cohn Professorial chair.

References

1. Vincken JP, Heng L, de Groot A et al (2007) Saponins, classification and occurrence in the plant kingdom. Phytochemistry 68:275–297
2. Nohara T, Ikeda T, Fujiwara Y et al (2007) Physiological functions of solanaceous and tomato steroidal glycosides. J Nat Med 61:1–13
3. Friedman M (2002) Tomato glycoalkaloids: role in the plant and in the diet. J Agric Food Chem 50:5751–5780
4. Itkin M, Rogachev I, Alkan N et al (2011) Glycoalkaloid metabolism1 is required for steroidal alkaloid glycosylation and prevention of phytotoxicity in tomato. Plant Cell 23:4507–4525
5. Itkin M, Heinig U, Tzfadia O et al (2013) Biosynthesis of antinutritional alkaloids in solanaceous crops is mediated by clustered genes. Science 341:175–179
6. Mintz-Oron S, Mandel T, Rogachev I et al (2008) Gene expression and metabolism in tomato fruit surface tissues. Plant Physiol 147: 823–851
7. Smith CA, Want EJ, O'Maille G et al (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem 78:779–787
8. De Vos RC, Moco S, Lommen A et al (2007) Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat Protoc 2:778–791
9. Moco S, Bino R, De Vos R et al (2007) Metabolomics technologies and metabolite identification. Trends Analyt Chem 26: 855–866
10. Rogachev I, Aharoni A (2012) UPLC-MS-based metabolite analysis in tomato. Methods Mol Biol 860:129–144
11. Brodsky L, Moussaieff A, Shahaf N et al (2010) Evaluation of peak picking quality in LC-MS metabolomics data. Anal Chem 82: 9177–9187
12. Kuhl C, Tautenhahn R, Neumann S (2010) CAMERA: collection of annotation related methods for mass spectrometry data. R package (version 1.4.2.). http://msbi.ipb-halle.de/msbi/CAMERA/
13. Katajamaa M, Oresic M (2005) Processing methods for differential analysis of LC/MS profile data. BMC Bioinforma 6:179
14. Lommen A (2009) MetAlign: interface-driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. Anal Chem 81:3079–3086
15. Li R, Zhou Y, Wu Z et al (2006) ESI-QqTOF-MS/MS and APCI-IT-MS/MS analysis of steroid saponins from the rhizomes of *Dioscorea panthaica*. J Mass Spectrom 41:1–22
16. Sharma SC, Chand R, Sati OP et al (1983) Oligofurostanosides from *Solanum nigrum*. Phytochemistry 22:1241–1244

The AP2/ERF-type Transcription Factor GLYCOALKALOID METABOLISM 9 Regulates Cholesterol and Steroidal Alkaloid Biosynthesis in the *Solanaceae*

Pablo D. Cárdenas[1,2], Prashant D. Sonawane[1], Jacob Pollier[3,4], Robin Vanden Bossche[3,4], Efrat Weithorn[1], Lior Tal[1], Sagit Meir[1], Ilana Rogachev[1], Alain Goossens[3,4], Saul Burdman[2], Asaph Aharoni[1*]

[1]Department of Plant and Environmental Sciences, Weizmann Institute of Science, Rehovot 76100, Israel

[2]Department of Plant Pathology and Microbiology, The Robert H. Smith Faculty of Agriculture, Food and Environment, The Hebrew University of Jerusalem, Rehovot 76100, Israel

[3]Department of Plant Systems Biology, Flanders Institute for Biotechnology (VIB), B-9052 Gent, Belgium

[4]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium.

[*]Corresponding author: asaph.aharoni@weizmann.ac.il

ABSTRACT

Steroidal alkaloids (SAs) are plant defense compounds produced in the *Solanaceae* family and considered as anti-nutritional factors in the human diet. Recently, a multi-step pathway for SA biosynthesis was proposed starting from cholesterol up to the glycosylated SAs (i.e. steroidal glycoalkaloids; SGAs). Here, we discovered that *GLYCOALKALOID METABOLISM 9* (*GAME9*), an APETALA2/Ethylene Response Factor (AP2/ERF), regulates SGA biosynthesis. *GAME9* is closely related to transcription factors regulating the biosynthesis of nicotine (a pyridine alkaloid) in *Nicotiana tabacum* and terpenoid indole alkaloids in *Catharanthus roseus*. Downregulation of *GAME9* in tomato resulted in a considerable reduction of the main SGA α-tomatine levels in leaves. Conversely, overexpression of *GAME9* caused an increase in α-tomatine in tomato and α-chaconine and α-solanine in potato, together with an altered sterol composition. Altered *GAME9* expression affected genes involved in the biosynthesis of SGAs and the upstream cholesterol precursor pathway. Some but not all of these genes are direct targets of GAME9, hence, likely additional transcriptional regulator(s) that act either downstream or interact with it may be involved. Our findings provide insights into the transcriptional regulation of SA biosynthesis and means for manipulation of these defense metabolites in *Solanaceae* crops in which high SGA levels affect product quality.

INTRODUCTION

Steroidal alkaloids (SAs) and their glycosylated forms (steroidal glycoalkaloids; SGAs) are nitrogen-containing toxic compounds occurring primarily in the *Solanaceae* and *Liliaceae* plant families (Heftmann, 1983). This class of metabolites is produced in *Solanaceae* vegetable crops such as potato (*Solanum tuberosum*), tomato (*Solanum lycopersicum*) and eggplant (*Solanum melongena*). Although SGAs contribute to plant resistance to a wide range of pathogens and predators, including bacteria, fungi, oomycetes, viruses, insects and animals (Milner et al., 2011), some are considered as anti-nutritional compounds to humans due to their toxic effects (Eich, 2008; Roddick, 1996).

In potato, α-chaconine and α-solanine comprise >90% of the total SGA content in the tubers. Nevertheless over 50 different SGAs have been identified in a variety of potato wild species and commercial cultivars (Itkin et al., 2013; Shakya and Navarre, 2008). In tomato, α-tomatine and dehydrotomatine are the major SGAs in green tissues, while esculeosides are predominant in the red ripe fruit (Fujiwara et al., 2004; Itkin et al., 2011; Mintz-Oron et al., 2008). About 100 SAs have been reported in different tissues and developmental stages of tomato (Iijima et al., 2013; Moco et al., 2006; Itkin et al., 2011; Schwahn et al., 2014). Explored to a lesser extent, α-solasonine and α-solamargine are the two major SGAs found in eggplant (Mennella et al., 2012; Sánchez-Mata et al., 2010). Early studies of SGA biosynthesis in potato reported on the characterization of three glycosyltransferases (SGTs)- SGT1 (a galactosyltransferase; GAME1), SGT2 (a glucosyltransferase) and SGT3 (a rhamnosyltransferase; GAME2)- that are involved in the addition of sugar moieties on the aglycone solanidine, leading to specific synthesis of either α-solanine or α-chaconine (Moehs et al., 1997; McCue et al., 2005, 2006, 2007). In tomato, the first gene reported to be involved in the synthesis SGAs was *GAME1*, encoding a glycosyltransferase homologous to the potato *SGT1*, which catalyzes the galactosylation of the aglycone tomatidine (Itkin et al., 2011).

Recently, Itkin et al. (2013) reported a set of *GLYCOALKALOID METABOLISM* (*GAME*) genes that participate in the core pathway producing SGAs in both potato and tomato. Consequently, an elaborated pathway for SGA biosynthesis in the *Solanaceae* family, starting from the precursor cholesterol up to the SGAs, was proposed (Itkin et al., 2013). Extensive functional characterization suggested that cholesterol undergoes several hydroxylation, oxidation, transamination and glycosylation steps to generate SGAs. The *GAME* genes were found to be located physically close to each other in the genome and thus organized in a form of metabolic gene

clusters. In tomato six *GAME* genes are positioned in a cluster on chromosome 7, whereas two other genes are located next to each other on chromosome 12. Furthermore, two additional genes, encoding cytochrome P450s (P450s), not belonging to these clusters, were also associated with SGA biosynthesis (*GAME7* and *GAME8*). In potato, four SGA-related genes are located on chromosome 7 and two on chromosome 12 (Itkin et al., 2013). In tomato, the *GAME* genes include P450s [GAME7, GAME8, GAME6 (chromosome 7 cluster) and GAME4 (chr. 12)], a dioxygenase (GAME11; chr. 7) involved in the hydroxylation and oxidation of the cholesterol skeleton and a transaminase protein (GAME12; chr. 12) required for the incorporation of the nitrogen atom into the SA aglycone. Finally, glycosyltransferases (GAME1, GAME17, GAME18 and GAME2; chr. 7) required for generating the sugar moieties that decorate the SA aglycone were also among the clustered genes.

Cholesterol, produced through the cytosolic isoprenoid mevalonate pathway is a key precursor in the biosynthesis of SGAs in the major families of plants producing this class of specialized metabolites. In sharp contrast to other kingdoms, the pathway leading to cholesterol biosynthesis in plants is yet unresolved. In the first part of the known mevalonate pathway, acetyl-CoA undergoes multiple enzymatic reactions to yield squalene. These reactions include two committed steps: the conversion of 3-hydroxy-3-methylglutaryl CoA to mevalonic acid, catalyzed by a family of 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR) enzymes, and the synthesis of squalene by squalene synthase (SQS) (Nes, 2011; Ginzberg et al., 2012). Squalene is subsequently oxidized to form 2,3-oxidosqualene, a branch point intermediate metabolite leading to multiple triterpenoids, including α-, β-amyrin, lupeol and cycloartenol. The latter metabolite represents a key branching point between the pathways forming the C-24 alkylated phytosterols (e.g. campesterol, β-sitosterol) and cholesterol (non-alkylated sterols).

Very recently, research related to SGA biosynthesis advanced our knowledge regarding the pathway to cholesterol formation in SGA-producing *Solanaceae* species. Sawai et al. (2014) demonstrated that STEROL SIDE CHAIN REDUCTASE 2 (SSR2) exhibits $\Delta^{24(25)}$ reductase activity that converts cycloartenol to cycloartanol in the first committed step towards cholesterol formation. Hence, SSR2 directs the pathway towards cholesterol and SAs instead of to alkylated sterol biosynthesis (Sawai et al., 2014). On the other hand, STEROL METHYLTRANSFERASE1 (SMT1) directs the pathway towards C-24 alkylsterols by adding a methyl group at the C-24 position of the cycloartenol side chain (Diener et al., 2000). Overexpression of a soybean *SMT1* in

potato plants therefore increased the metabolic flux of cycloartenol into alkylated sterols at the expense of cholesterol (Arnqvist et al., 2003).

In contrast to the intense research related to structural genes of the pathway, the transcriptional regulation of SGA biosynthesis and its cholesterol precursor pathway is utterly unclear. Some transcription factors have been identified that regulate the biosynthesis of other classes of alkaloids in different plant species (van der Fits and Memelink, 2000; De Sutter et al., 2005; Shoji et al., 2010; Todd et al., 2010; De Boer et al., 2011; De Geyter et al., 2012; Yamada et al., 2011; Yamada et al., 2015). A major class of transcriptional regulators mediating plant alkaloid biosynthesis is the one represented by the APETALA2/ethylene response factors (AP2/ERF) family members. The AP2/ERF transcription factor ORCA3 regulates the biosynthesis of terpenoid indole alkaloids (TIAs) in *Catharanthus roseus* (van der Fits and Memelink, 2000). *ORCA3* gene expression is induced by jasmonate and is regulated by direct binding of the basic helix-loop-helix (bHLH) transcription factor CrMYC2 to the *ORCA3* gene promoter (Zhang et al., 2011). Close homologs of ORCA3 in *Nicotiana tabacum* present in the *NIC2* locus were associated with nicotine levels in the tobacco leaf and have been used extensively in breeding of low-nicotine tobacco lines (Hibi et al., 1994). Specifically, the *NIC2* locus comprises at least seven ERF transcription factors that regulate the expression of structural genes in the biosynthesis of nicotine. In the *nic2* mutant, this *ERF* gene cluster is deleted, resulting in a low-nicotine phenotype (Shoji et al., 2010). Genes present in the *NIC2* locus include *ERF189* and *ERF221* (also known as ORC1; De Sutter et al., 2005). Overexpression of *ERF189* and *ERF221/ORC1* was sufficient to stimulate nicotine biosynthesis in tobacco plants (Shoji et al., 2010; De Boer et al., 2011). Members of the ERF family of transcription factors can recognize three different GC-rich boxes in the promoters of target genes (Shoji et al., 2013). For instance, ERF189 binds a P-box (5'-CCGCCCTCCA-3') in the promoter of the *putrescine N-methyltransferase* (*PMT*) gene involved in the formation of the pyrrolidine ring in nicotine (Shoji et al., 2010). ORCA3 can recognize a CS1-box (5'-TAGACCGCCT-3') in the promoter of the TIAs biosynthetic gene *strictosidine synthase* (*STR*). Finally, the Arabidopsis ERF1 functions as an activator of transcription mediated by a GCC box (5'-AGCCGCC-3') (Fujimoto et al., 2000).

In this study, we characterized *GLYCOALKALOID METABOLISM 9* (*GAME9*), an AP2/ERF transcription factor that regulates the biosynthesis of steroidal alkaloids in *Solanaceae* plants. Based on previously reported quantitative trait loci (QTLs) associated with SGA content in potato and co-expression analysis, we found that *GAME9* is part of an *ERF*-gene cluster existing in

potato and tomato. Analyses of transgenic plants as well as promoter binding assays suggested that GAME9 exerts its activity through interaction or activation of additional regulatory proteins as well as by directly controlling the levels of the upstream cholesterol precursor pathway. The findings provide insight into the transcriptional regulation of SAs in *Solanaceae* plants as well as a base for engineering plants with modified levels of these anti-nutritional compounds.

RESULTS

Initial evidence that *GLYCOALKALOID METABOLISM 9* is associated with steroidal glycoalkaloid biosynthesis

In a previous study, we discovered an AP2/ERF type transcription factor (Solyc01g090340 and Sotub01g029510, in tomato and potato, respectively) displaying a similar expression pattern to genes in the tomato and potato SGA biosynthetic pathway (Itkin et al., 2013). To further characterize the involvement of this putative regulator (named GLYCOALKALOID METABOLISM 9; GAME9) in the control of SGA biosynthesis, we carried out combined co-expression analysis using potato and tomato transcriptome data. A total of 1,260 and 168 genes were co-expressed with *GAME9* in tomato and potato, respectively (Figure 1; Supplemental Table 1). Thirty seven homologous genes were co-expressed with *GAME9* in both potato and tomato (Figure 1A; Supplemental Table 2). Among the co-expressed genes, we found all those previously associated with SGA biosynthesis in potato including *GAME2*, *GAME11*, *GAME6* and *GAME1* on chromosome 7, and *GAME12* and *GAME4* on chromosome 12. Similarly, in tomato, *GAME11*, *GAME6*, *GAME17*, *GAME1* and *GAME18* on chromosome 7, and *GAME12* and *GAME4* on chromosome 12 were co-expressed with *GAME9* (Figure 1A). Genes encoding HMGR and SQS, involved in the synthesis of isoprenoid precursors in the mevalonate pathway, were not co-expressed with *GAME9* in either species (Supplemental Table 1). Interestingly, sterol biosynthesis related (phytosterols and cholesterol) genes, *SSR2* and *CYCLOARTENOL SYNTHASE* (*CAS*) were co-expressed with *GAME9* in tomato (r-value≥0.73) while only *SSR2* was co-expressed with the potato *GAME9* gene (Supplemental Table 1). When examined across 19 different tomato tissue types, *GAME9* was highly expressed in leaf and flower buds. In fruit tissues, it was expressed early, predominantly in the immature stages of development (Figure 1B) while displaying some, albeit relatively low level of expression in petals and root tissues. The expression pattern of *GAME9* was analyzed using RNA *in situ* hybridization. In 13-day-old tomato shoots, *GAME9* was expressed in both young leaves and throughout the vascular system. *GAME9* expression was also detected in mature leaves, mostly in the outer layers of the blade (Figure 1C).

An *AP2*/*ERF*-gene cluster containing *GAME9* is part of a major QTL linked to SGA content in potato

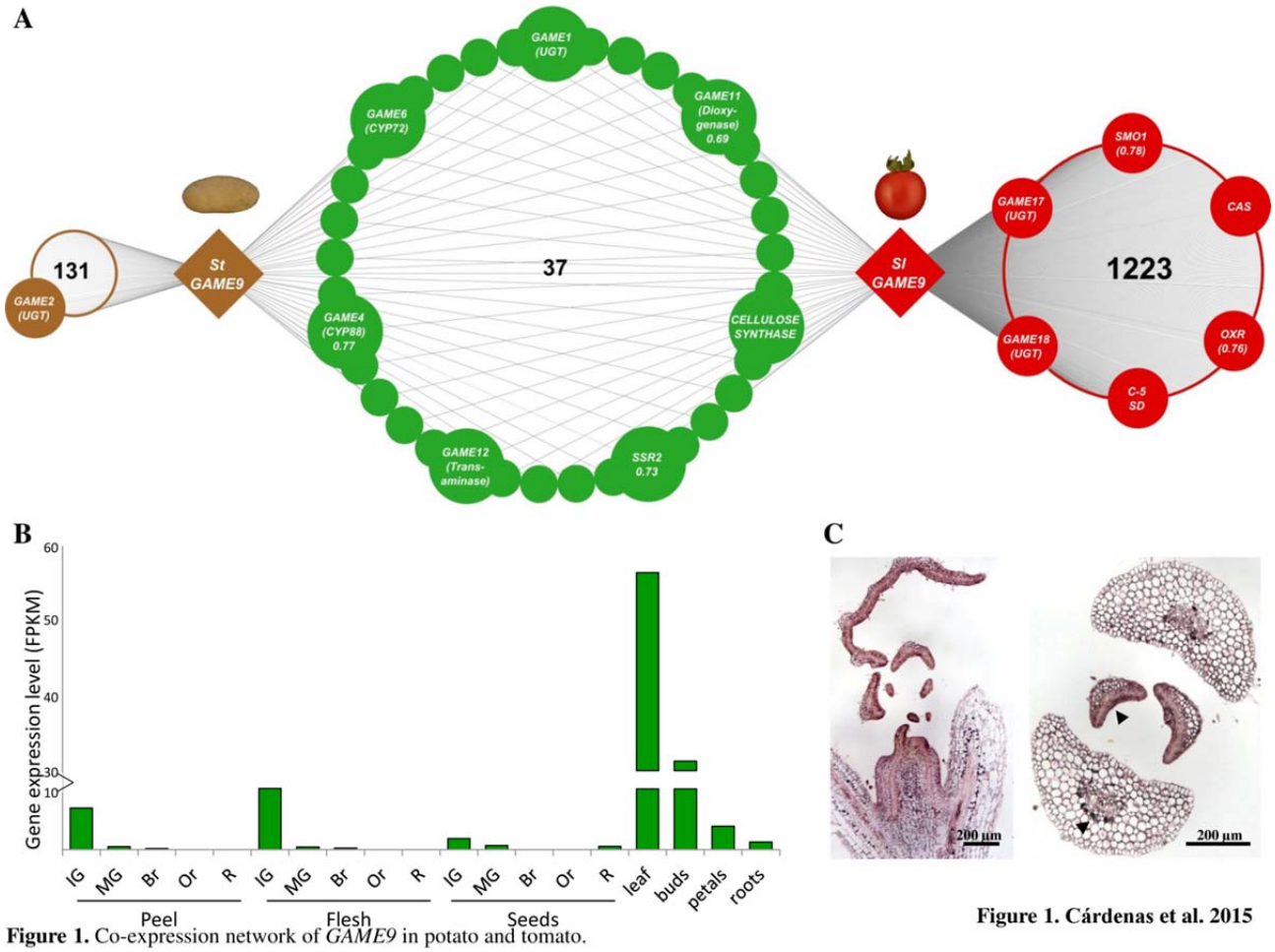


Figure 1. Cárdenas et al. 2015

Figure 1. Co-expression network of *GAME9* in potato and tomato.

(A) Using RNA-Seq transcriptome data from potato and tomato, we found that *GAME9* was co-expressed with most of the SGA biosynthetic genes. Thirty seven shared homologs were co-expressed in both potato and tomato (see Supplemental Table 1 and Supplemental Table 2). *C-5 SD*: *Δ(7)-STEROL-C5(6)-DESATURASE*, *SMO1*: *METHYLSTEROL MONOOXYGENASE 2-2-LIKE*, *OXR*: 3-*β HYDROXYSTEROID DEHYDROGENASE*, *CAS*: *CYCLOARTENOL SYNTHASE*.

(B) Expression profile (from RNA-sequencing) of *GAME9* in different tomato tissue types and developmental stages. IG: immature green, MG: mature green, BR: breaker, Or: orange, R: ripe.

(C) *In situ* mRNA hybridization showing *GAME9* expression in wild-type tomato shoot meristems. Both longitudinal and transverse sections are shown. Arrows indicate higher expression in the vascular system and the outer layer of the leaf blade. Scale bars: 200 µm.

Identification of QTLs linked to total SGA content in potato tubers has been of high interest in breeding of new potato cultivars. Sørensen et al. (2008) reported a highly significant QTL on chromosome 1 that explained a major proportion of the SGA content in potato tubers (both in dark and light exposed tubers). Considering that *GAME9* is located on chromosome 1, we suspected that it might be associated with this earlier reported QTL. The potato QTL was flanked by the simple sequence repeat (SSR) markers STM5136 and STM2030 (Sørensen et al., 2008). We used these SSR markers to identify the corresponding chromosomal region spanning 6.6 Mbp on chromosome 1 of tomato [between markers TG21 and TG59 (Figure 2A)]. Interestingly, in both species, *GAME9*

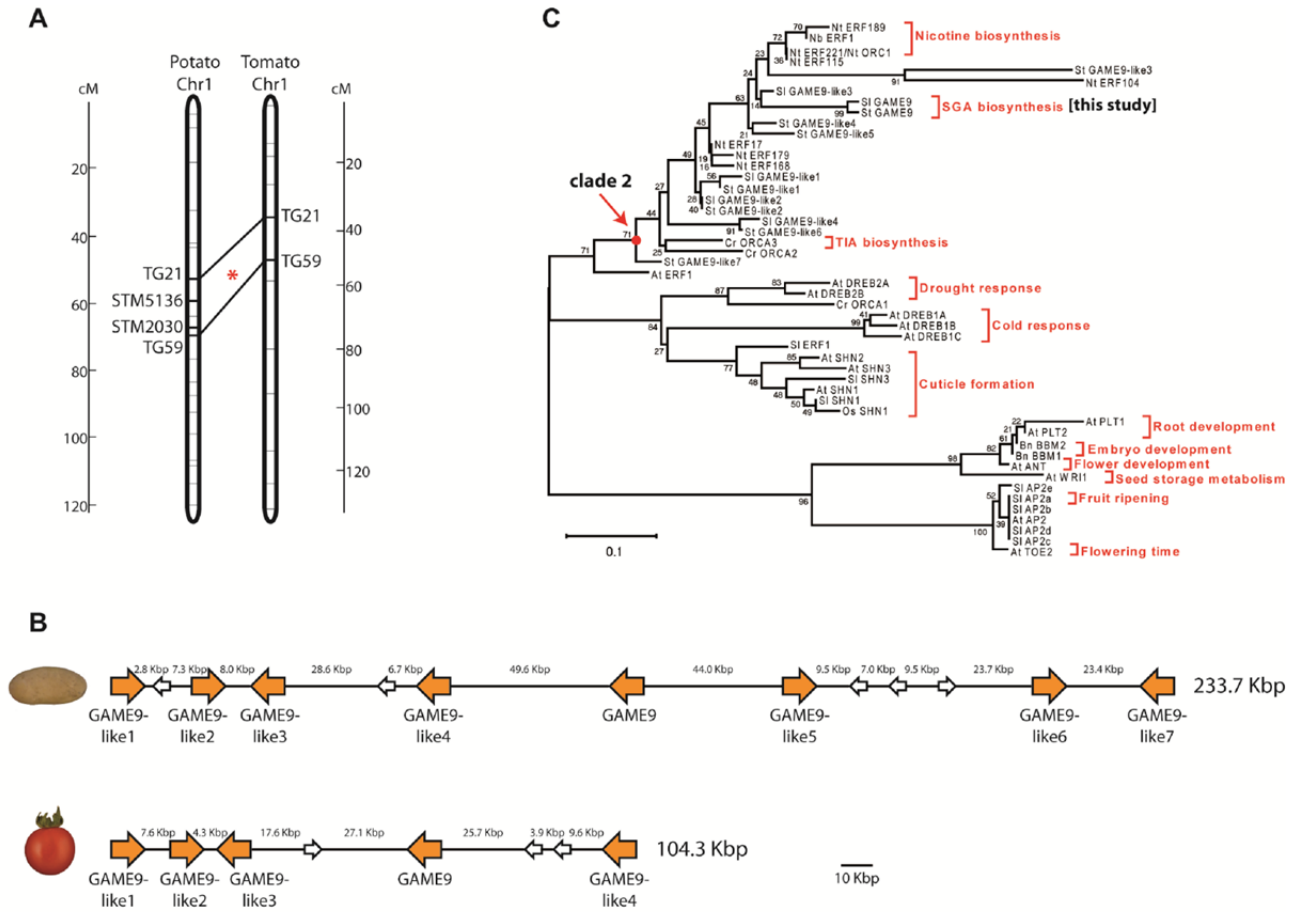


Figure 2. Cárdenas et al. 2015

Figure 2. *GAME9* is part of an *ERF*-gene cluster in the *Solanaceae* and is related to other alkaloid-associated regulatory genes.

(A) A major QTL involved in the synthesis of SGAs is present on potato chromosome 1 (Sørensen et al., 2008). The QTL is flanked by the SSR markers STM5136 and STM2030. Using a comparative map viewer we identified the corresponding region in tomato (markers TG21 and TG59). *- *GAME9* was found to be located in this QTL region in both potato and tomato.

(B) Schematic presentation of *GAME9* and *GAME9*-like genes in chromosomal regions of potato and tomato. In these regions, we found clusters of *ERF* genes spanning a ~230 kbp region in potato and a region of ~104 kbp in tomato.

(C) Phylogenetic analysis of GAME9 and other related AP2-family proteins from tomato (Sl), potato (St), tobacco (Nt), *N. benthamiana* (Nb), periwinkle (Cr), rice (Os), *Brassica napus* (Bn), and Arabidopsis (At). The evolutionary history was inferred using the neighbor-joining method in MEGA6 (Tamura et al., 2013). The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. Accession numbers can be found in Supplemental Table 7.

was located inside these QTL regions, and moreover, as part of a cluster of AP2/ERF transcription factors. In potato, a cluster spanning ~230 kilobase pair (kbp) genomic region includes *GAME9* together with seven *GAME9*-like transcription factors, whereas in tomato, a region of ~104 kbp contains *GAME9* and additionally four *GAME9*-like genes (Figure 2B).

Phylogenetic analysis showed that GAME9 and GAME9-like proteins are part of the ERF IXa subfamily (Nakano et al., 2006) divided earlier by Shoji et al. (2010) into two seprate clades. GAME9 and the GAME9-like proteins are part of clade 2 that includes ERF189 from tobacco, one of the AP2/ERF proteins in the *NIC2* locus involved in the synthesis of the pyridine alkaloid

nicotine. The same clade also includes ORCA3 and ORCA2, both transcription factors involved in the synthesis of TIAs in *C. roseus* (Menke et al., 1999; van der Fits and Memelink, 2000; Figure 2C). Two other members of the ERF IXa subfamily clade 2 are also involved in the control of nicotine biosynthesis, namely, the tobacco ERF221 (ORC1; De Sutter et al., 2005) and the *Nicotiana benthamiana* ERF1 (Todd et al., 2010). Thus, *GAME9* represents a potential third case in which proteins of this clade control the biosynthesis of different types of alkaloids.

Altering *GAME9* expression impacts the levels of principal SGAs in tomato and potato

To provide additional evidence regarding the role of *GAME9* in SGA biosynthesis, we generated transgenic tomato lines in which *GAME9* was silenced (*GAME9*-RNAi) or overexpressed (*GAME9*-Ox). Transgenic potato lines overexpressing *GAME9* were also generated. *GAME9* gene expression was examined by qRT-PCR in both wild-type and transgenic potato and tomato lines (Figure 3A). *GAME9* transcript levels were significantly higher in *GAME9*-Ox lines from potato and tomato, and were decreased in the *GAME9*-RNAi tomato lines. SGAs profiling was carried out on extracts of tomato and potato leaves and skin of potato tubers by Liquid Chromatography Mass Spectrometry (LC-MS). In leaves of potato *GAME9*-Ox lines, the levels of α-solanine and α-chaconine increased between 3.5-4.6 fold and 2.8-4.2 fold, respectively as compared to wild-type plants (Figure 3B). Likewise, in tuber skin isolated from the same potato lines, we detected an increase in α-solanine levels (up to 1.2-2.6 fold) and α-chaconine (up to 1.2-2.1 fold) (Figure 3B). In tomato, the level of α-tomatine was significantly reduced in *GAME9*-RNAi lines, while it was increased up to 1.5-2.8 fold when *GAME9* was overexpressed (Figure 3B).

The effect of altered *GAME9* expression on levels of mevalonate pathway intermediates and its branches in potato

We envisaged that regulation of SGA contents by GAME9 is achieved, at least partially, by regulating the flux through the mevalonate pathway and its branches. These include C-24 alkylated phytosterols (e.g. campesterol and β-sitosterol), non-alkylated sterols (primarily cholesterol, which is the precursor for SGA biosynthesis), and the triterpenoid branch. Gas Chromatography Mass Spectrometry (GC-MS) was employed to profile the various metabolic intermediates in leaves of the four *GAME9*-Ox lines. Overexpression of *GAME9* in potato resulted in a dramatic decrease in levels of cycloartenol and cycloartanol, early intermediates in cholesterol biosynthesis (Figure 4 and Figure 5). Cholesterol itself showed a slight, but significant increase in leaves of the *GAME9*-Ox

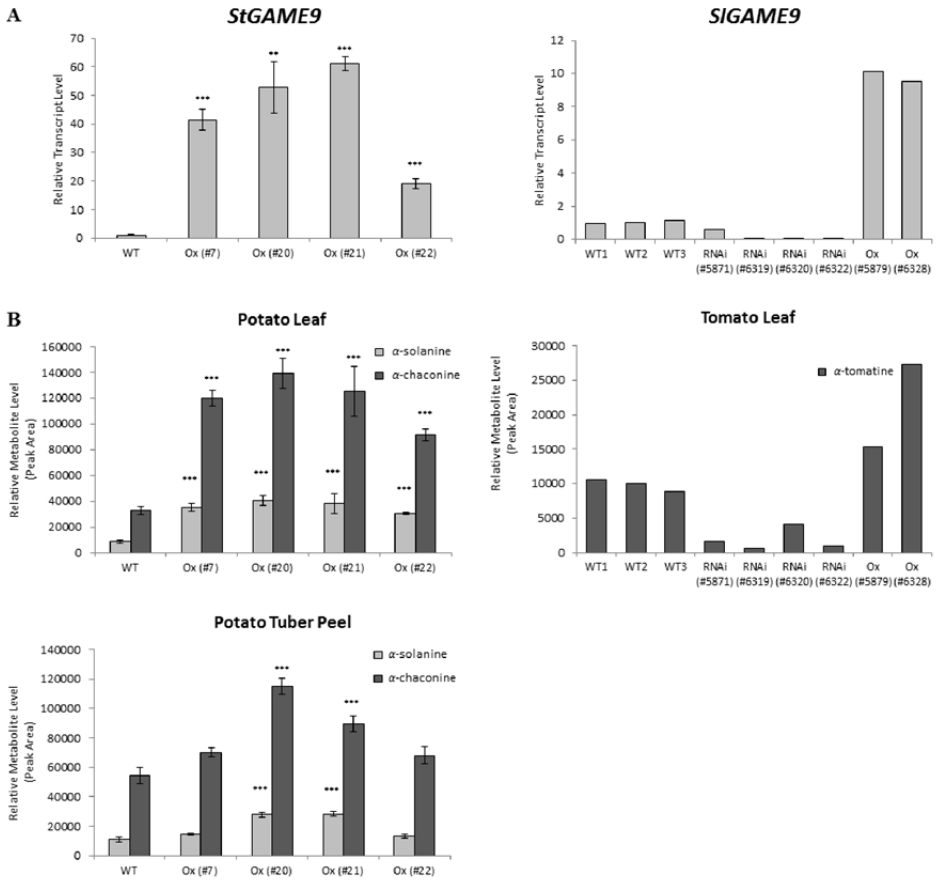


Figure 3. Cárdenas et al. 2015

Figure 3. Changes in *GAME9* expression in potato and tomato results in altered levels of the predominant SGAs.

(A) *GAME9* gene expression (qRT-PCR) in *GAME9*-Ox and *GAME9*-RNAi lines in potato and tomato.

(B) Levels of *α*-solanine and *α*-chaconine in leaves and peel of potato tubers of *GAME9*-Ox lines and levels of *α*-tomatine in leaves of *GAME9*-RNAi and *GAME9*-Ox tomato lines.

Values represent means ± standard errors (n=3). Student's *t*-test was used to assess whether the transgenic lines significantly differ from wild-type plants: ***P value < 0.001.

lines (Figure 4). Interestingly, *β*-amyrin and campesterol contents were also increased, yet, *β*-sitosterol was detected in levels similar to those in leaves of wild-type plants (Figure 4 and Figure 5). These observations point to increased flux to cholesterol as well as to a certain part of phytosterol (i.e. campesterol) biosynthesis due to *GAME9* overexpression.

Quantitative Real-Time PCR analysis of transcripts associated with SGAs, mevalonate and downstream pathways in the *GAME9* altered plants

We used quantitative Real-Time PCR (qRT-PCR) to examine the expression level of SGA biosynthesis genes and those in the mevalonate and downstream pathways (towards triterpenoids, phytosterol and cholesterol biosynthesis) in the *GAME9* altered plants. In potato, *GAME9*

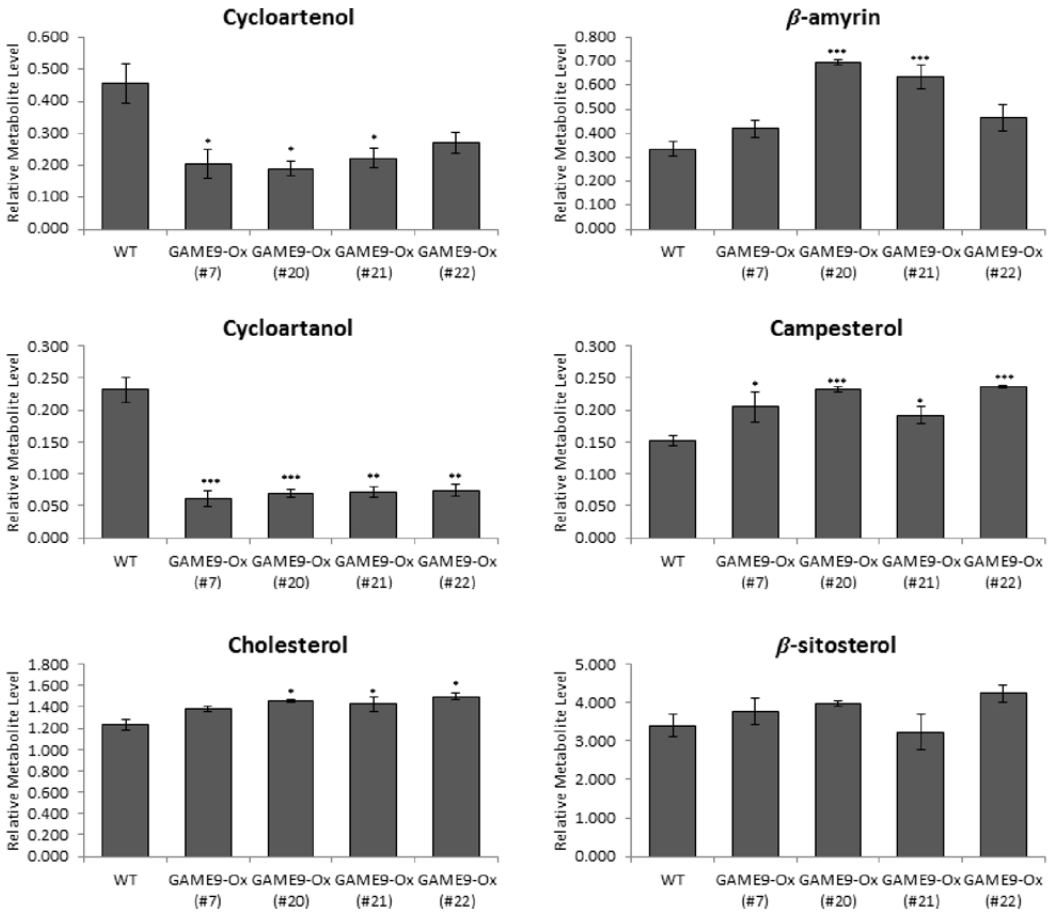


Figure 4. Cárdenas et al. 2015

Figure 4. Cholesterol and other sterols levels in potato *GAME9*-Ox lines.

Relative abundance of six sterols in leaves of four independent *GAME9*-Ox plant lines as compared to wild-type (WT) measured using GC-MS. Epicholesterol was used as an internal standard. Relative metabolite levels are expressed as ratios of peak areas compared to internal standard. Values represent means ± standard errors (n=3). Student's *t*-test was used to assess whether the transgenic lines significantly differ from wild-type plants: *P value < 0.05; **P value < 0.01; ***P value < 0.001.

overexpression did not change the expression of genes involved in the synthesis of triterpenoid precursors (i.e. *HMGR*, *SQS*), *β*-amyrin (*TRITERPENOID SYNTHASE 1*, *TTS1*, Wang et al., 2011) and campesterol/*β*-sitosterol (*SMT1*) (Figure 5; Supplemental Table 3). However, genes acting downstream to 2,3-oxidosqualene, towards the formation of sterols, including *CAS* and particularly *SSR2*, encoding the enzyme performing the first step in the biosynthesis of cholesterol from cycloartenol, were upregulated in the *GAME9*-Ox potato lines. Similarly, in the same lines, the *GAME* genes responsible for the synthesis of the solanidine aglycone (*GAME11*, *GAME6*, *GAME4* and *GAME12*) and the subsequent glycosylation (*GAME1*, *GAME2* and *SGT2*) were all significantly upregulated (Figure 5; Supplemental Table 3).

In tomato, we found that when *GAME9* was either silenced or overexpressed, expression of *HMGR* and *SSR2* was significantly altered (down and upregulated in the *GAME9*-RNAi and

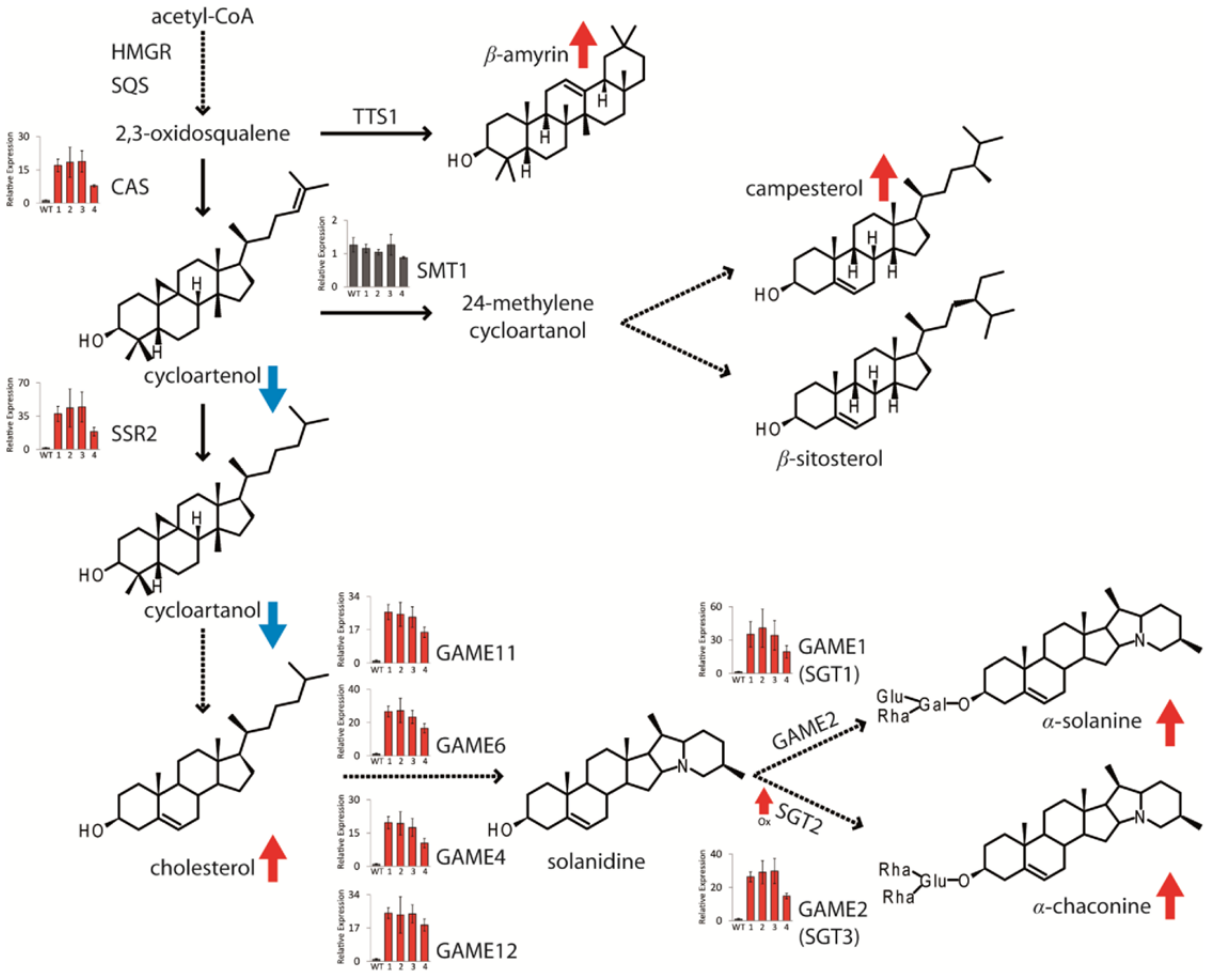


Figure 5. Cárdenas et al. 2015

Figure 5. Expression of genes involved in the synthesis of SGAs and sterol precursors in potato leaves derived from *GAME9*-Ox lines determined by qRT-PCR and RNA-Seq analyses.

Schematic view of the sterol and SGA biosynthetic pathway. Dashed arrows represent multiple biosynthetic reactions whereas solid arrows represent a single step. The graphs next to each gene name show expression levels in wild-type and four *GAME9*-Ox independent lines (lines 1:#7, 2:#20, 3:#21 and 4:#22) determined by qRT-PCR. According to RNA-Seq data, HMGR, SQS, TTS1 were not altered, and *SGT2* was overexpressed in GAME9-Ox lines. Arrows next to each compound represent an increase (red) or decrease (blue) in potato *GAME9*-Ox lines (see Figure 3 and Figure 4).

GAME9-Ox lines, respectively). However, expression of *TTS1* involved in the triterpene *β*-amyrin formation was not affected. Altered expression of *GAME9* did not affect *SQS*, *CAS* and *SMT1* expression levels (Figure 6; Supplemental Table 4). Finally, 7 out of 8 examined *GAME* genes involved in the synthesis of the SGA aglycone tomatidine and its glycosylation were altered in expression, at all times correlating with the *GAME9* transcript levels.

Transcriptome changes in *GAME9* overexpression and silenced lines

To obtain a more global picture of genes that are downstream of GAME9 and to understand more precisely the metabolic pathways under its control, we performed RNA-sequencing (RNA-Seq) in

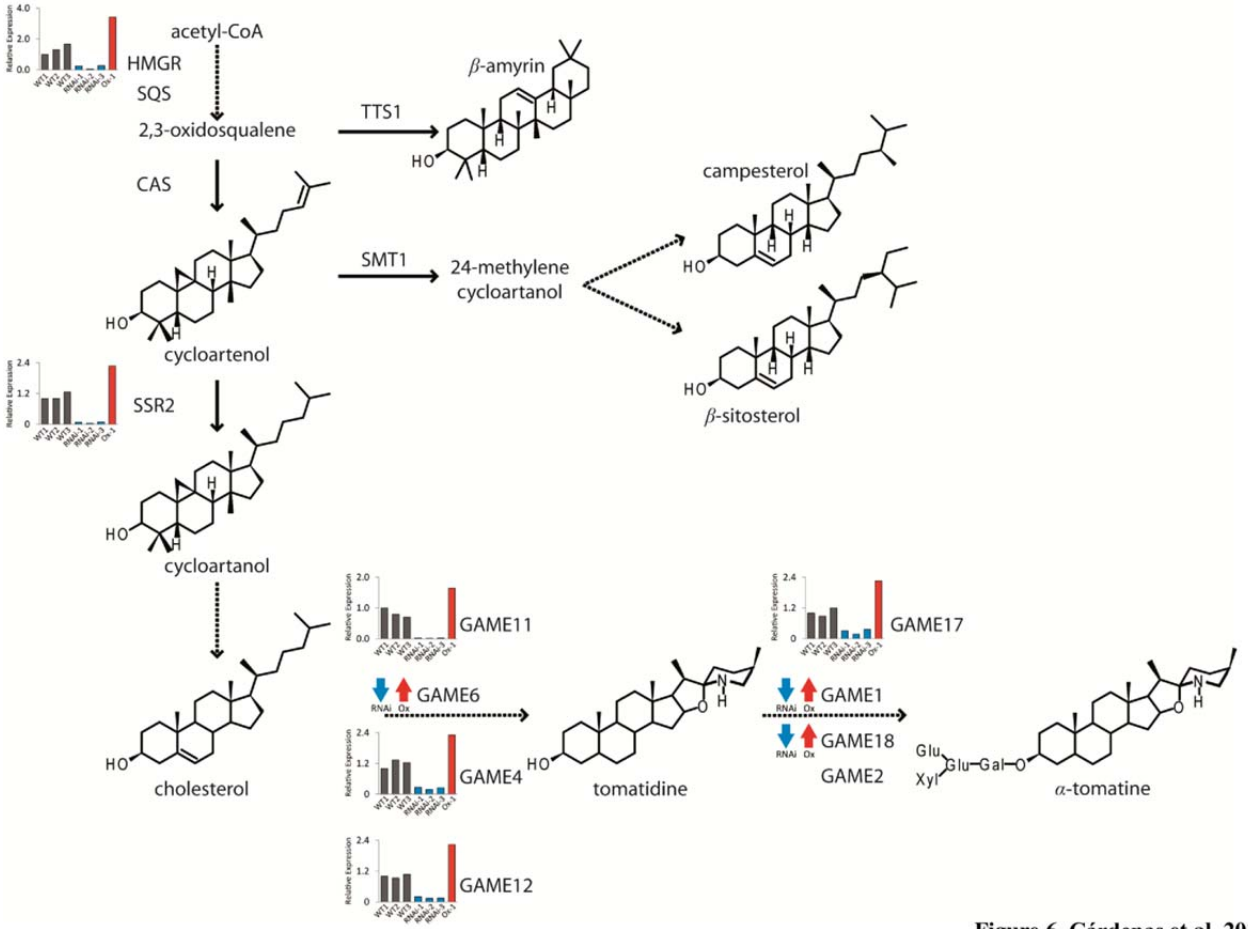


Figure 6. Cárdenas et al. 2015

Figure 6. Expression of genes involved in the synthesis of SGAs and sterol precursors in tomato leaves derived from *GAME9*-RNAi and *GAME9*-Ox lines.

A schematic view of the sterol and SGA biosynthetic pathways. Dashed arrows represent multiple biosynthetic reactions whereas solid arrows represent a single step. Graphs next to each gene name show expression levels in three wild-type (WT1, WT2, WT3), three *GAME9*-RNAi (lines RNAi- 1:#6319, 2:#6320 and 3:#6322) and one *GAME9*-Ox plant (line Ox-1:#6328) determined by qRT-PCR. According to RNA-Seq data, SQS, CAS, TTS1, SMT1 and GAME2 were not altered in *GAME9*-Ox and *GAME9*-RNAi tomato lines. For other genes, expression is represented based on RNA-Seq data, with arrows next to each gene depicting an increase (red) or decrease (blue) in *GAME9*-Ox and *GAME9*-RNAi lines, respectively.

leaf tissue of *GAME9*-RNAi and *GAME9*-Ox tomato lines and wild-type. Transcriptome analysis was also conducted on leaves of potato lines overexpressing *GAME9* and wild-type ones. Silencing of *GAME9* in tomato resulted in 931 genes that were downregulated [fold change $\log_2$ (RNAi/WT) < -0.5; Supplemental Table 5]. When *GAME9* was overexpressed, 1,002 genes were upregulated in tomato [fold change $\log_2$ (Ox/WT) > 0.5]. *GAME9* overexpression in potato, led to upregulation of 1,829 genes [fold change $\log_2$ (Ox/WT) >0.5; Supplemental Table 5].

A concise set of 27 genes (including *GAME9*) was found in common between the down- and upregulated genes in the *GAME9*-RNAi and *GAME9*-Ox tomato lines, respectively (Table 1). Among these, we found a significant representation of SGA biosynthetic genes (*GAME*s), explicitly

those located in the metabolic gene cluster in tomato chromosome 7 (*GAME*s *1*, *6*, *11*, *17* and *18*; Supplemental Figure 1; Itkin et al., 2013). This gene set also contained an additional gene in the SGAs cluster on chromosome 7, a sequence with homology to cellulose synthase family proteins (Solyc07g043390). Transcript of the *CELLULOSE SYNTHASE* sequence was also found to be significantly co-expressed with *GAME9* in both tomato and potato (Figure 1; Supplemental Table 2). Four genes out of the 27 could be associated with sterol metabolism, possibly phytosterol or cholesterol biosynthesis. Recent work reported one of the four genes, namely *SSR2*, a sterol side chain reductase catalyzing the first committed step towards cholesterol formation in the *Solanaceae* (the conversion of cycloartenol to cycloartanol; Figure 5; Sawai et al., 2014). The three additional genes include homologs of a *Δ(7)-STEROL-C5(6)-DESATURASE* (*C-5SD*), *METHYLSTEROL MONOOXYGENASE 2-2-LIKE* (*SMO*1) and a *3-B HYDROXYSTEROID DEHYDROGENASE* (*OXR*) (Table 1). Out of these 4 sterol metabolism associated genes, *SSR2* was co-expressed with *GAME9* in both potato and tomato, while the other three (i.e. *C-5SD*, *SMO1* and *OXR*) were significantly co-expressed with the *GAME9* transcript in tomato (Figure 1). Finally, among the 27 genes set we found a homolog of the *E3 UBIQUITIN-PROTEIN LIGASE RMA1H1-LIKE*. Apart from being significantly co-expressed with *GAME9* (Supplemental Table 1), this gene is related to an ERAD-type RING membrane-anchor E3 ubiquitin ligase reported to control the activity of 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) (Pollier et al., 2013), the rate-limiting enzyme in the mevalonate pathway leading to cholesterol and subsequently SGAs formation.

A set of 466 genes was found in common between the upregulated genes in the *GAME9*-Ox tomato and potato lines (Supplemental Table 5). Among them, we found a high representation of *GAME* genes located on chromosome 7 both in potato and tomato (*GAME*s *1*, *6* and *11*). Interestingly, *CELLULOSE SYNTHASE* is inside the chromosome 7 cluster of metabolic genes and was upregulated in tomato and potato in the *GAME9*-Ox leaves. The *SSR2, Δ(7)-STEROL-C5(6)-DESATURASE* and the *E3 UBIQUITIN-PROTEIN LIGASE RMA1H1-LIKE* were also among the genes upregulated in both the potato and tomato overexpression lines (Supplemental Table 5).

GAME9 transactivates genes involved in cholesterol biosynthesis

To study the GAME9 transactivation capacity of upstream regions of the putative target genes, we performed transient luciferase expression assays in tobacco protoplasts. Altogether, we assayed a total of 18 different putative promoter regions (ranging in size from 1200 bp to 2700 bp) of known tomato SGA genes and those putatively involved in the synthesis of SGAs or the mevalonate and

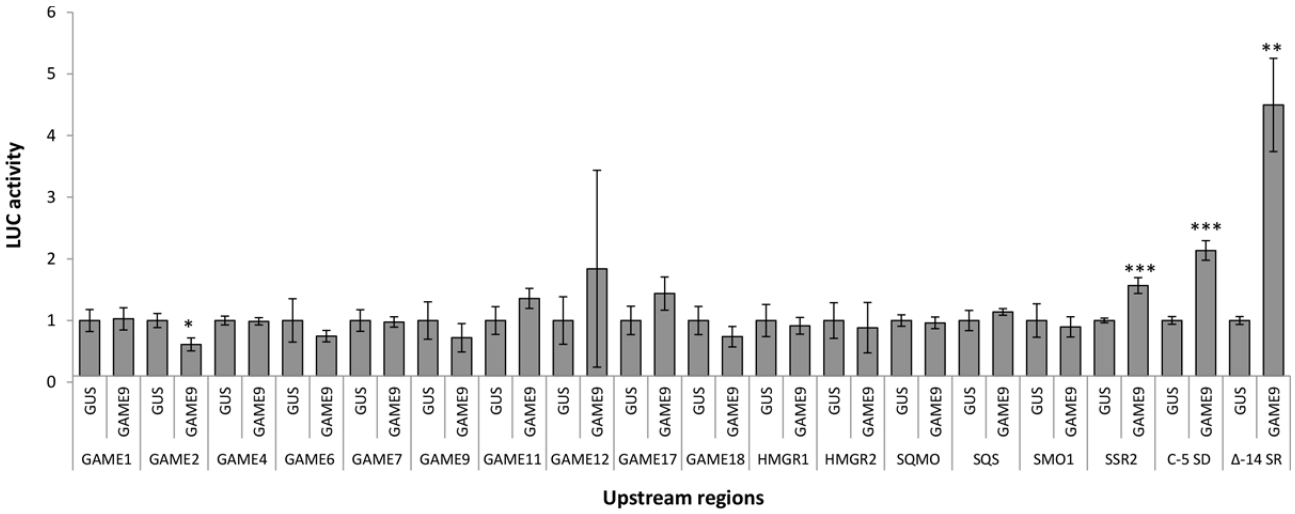


Figure 7. Cárdenas et al. 2015

Figure 7. Transactivation assays of putative downstream gene promoters by the GAME9 transcription factor.

The capacity of GAME9 to transactivate 18 different promoters of candidate downstream genes was evaluated in protoplasts prepared from tobacco Bright Yellow-2 (BY-2) cells (n=4). A validation experiment (n=8) was performed, confirming the transactivation of the *C5-SD* and *Δ14-SR* promoters (Supplemental Figure 2). Values in the y-axis are normalized fold-changes relative to protoplasts cotransfected with the reporter constructs and a pCaMV35S:GUS (GUS) control plasmid. For the normalization procedure, see Materials and Methods. Details for each promoter are provided in Supplemental Table 6. Student's *t*-test was used to assess whether the transgenic lines significantly differ from wild-type plants: *P value < 0.05; **P value < 0.01; ***P value < 0.001. *GAME1*: UDP-galactosyltransferase; *GAME2*: UDP-xylosyltransferase; *GAME4*: CYP88D; *GAME6*: CYP72A; *GAME7*: CYP72A; *GAME9*: Ethylene responsive transcription factor; *GAME11*: 2-oxoglutarate-dependent dioxygenase; *GAME12*: Aminotransferase-like protein; *GAME17*: UDP-glucosyltransferase; *GAME18*: UDP-glucosyltransferase; *HMGR1*: HMG CoA reductase 1; *HMGR2*: HMG CoA reductase 2; *SQMO*: Squalene monooxygenase; *SQS*: Squalene synthase; *SMO1*: Methylsterol monooxygenase 2-2-like; *SSR2*: Sterol Side Chain Reductase 2; *C-5 SD*: Δ(7)-sterol-C5(6)-desaturase; *Δ-14 SR*: Δ(14)-sterol reductase.

cholesterol precursor pathways (Supplemental Table 6). After an initial screening, we found that GAME9 in this assay did not transactivate the putative promoter regions of any of the core SGA pathway *GAME* genes acting in between cholesterol and *α*-tomatine (Figure 7). However, transactivation was detected for the promoters of SSR2 and two genes encoding Δ(7)-STEROL-C5(6)-DESATURASE and Δ-14 STEROL REDUCTASE, both putatively involved in the synthesis of the cholesterol (Figure 7; Supplemental Figure 2). These experiments support that GAME9 likely requires additional factors to control SGA production as discussed below.

DISCUSSION

Alkaloids represent one of the three major classes of specialized (or secondary) metabolites produced by plants with more than 20,000 of them being reported in thousands of species to date (Buckingham et al., 2010). The steroidal alkaloids produced by most members of the *Solanum* genus in the *Solanaceae* family are known primarily due to the toxicity of the major potato SGAs α-chaconine and α-solanine to mammals. In recent years, the investigation of these plant defense molecules in the two *Solanaceae* model plants, tomato and potato, has accelerated significantly. The major emphasis of the research has been on SGA identification and profiling as well as unraveling their biosynthetic pathway (McCue et al., 2005, 2006, 2007; Shakya and Navarre, 2008; Itkin et al., 2011, 2013; Iijima et al., 2013; Cárdenas et al., 2014). In this study, we identified GAME9, an AP2/ERF-type transcription factor, which regulates the biosynthesis of steroidal alkaloids in tomato and potato, and likely in other *Solanaceae* plants producing SGAs (e.g. eggplant). It appears that GAME9 belongs to a separate clade of AP2/ERF transcription factors together with proteins regulating the biosynthesis of other classes of alkaloids in other species namely, the pyridine alkaloid nicotine in tobacco and TIAs in *C. roseus*. This raises thought-provoking questions regarding the specificity of transcriptional regulation of alkaloids in plants and its molecular evolution as discussed below.

GAME9 and GAME9-like proteins are part of a clade of AP2/ERF transcription factors that regulate the production of different alkaloid classes in plants

As in the case of its homolog *ERF189* located in the *NIC2*-locus in *Nicotiana tabacum*, the *GAME9* gene is positioned inside a cluster of several similar, *GAME9*-like genes, both in potato and tomato. In the tobacco *NIC2* locus, seven highly similar *ERF* genes were shown to positively and specifically regulate the expression of structural genes involved in nicotine biosynthesis (Shoji et al., 2010). When these *ERFs* genes were used to rescue nicotine content in a *nic2* background, they showed some functional redundancy. However, ERF189 was able to recover nicotine content to the wild-type levels (Shoji et al., 2010). Similarly, it appears that ORCA3, involved in regulation of the TIA biosynthesis, is also positioned inside a cluster of similar genes (Kellner et al., 2015). As we did not investigate the GAME9-like proteins, we cannot exclude functional redundancy between cluster members (5 in tomato and 8 in potato). Yet, *GAME9* was the only gene in this cluster that

was co-expressed with other SGA genes and is thus likely to play a key role in the regulation of SGAs in both tomato and potato.

Phylogenetic analysis showed that GAME9 and GAME9-like proteins are part of clade 2 of the ERF IXa subfamily (Nakano et al., 2006). Hence, GAME9 represents a third case in which members of this clade control the biosynthesis of different types of alkaloids. It is important to note that the unique plant family of AP2/ERF-type transcription factors typically contains more than 100 proteins in a single genome (Nakano et al., 2006). Hence, we postulate that the phylogenetic data described above indicated that clade 2 members associated with alkaloid biosynthesis likely share a common ancestor. Possibly, an existing transcription factor acquired new downstream gene (promoter) targets. Nonetheless, it cannot be ruled out that these homologous regulators evolved through separate evolutionary events.

An intriguing question is therefore what could be the common feature linking between the different events that evolved these clade members to control pathways of alkaloid biosynthesis? Chemical structure and similarity in biosynthetic routes is not the case, as pyridine-, terpenoid indole-, and glyco-alkaloids are not related in these aspects. The common presence of nitrogen in these molecules is a non-specific feature, which cannot explain this phenomenon either. Common biosynthesis of nicotine might explain the recruitment of an ancestor nicotine pathway regulator for controlling other, newly evolved alkaloids pathways. Nevertheless, although endogenous nicotine has been detected in several plant species apart from tobacco including tomato (Sheen, 1988), we could not find any report describing its accumulation in *Catharanthus* plants.

Genetic affiliation might be a common thread between these species as *Catharanthus* (*Apocynaceae* family) belongs to the Gentianales order, which is related to the *Solanaceae* family members' tobacco, tomato and potato that are part the Solanales order. Both orders are positioned in the Asterids clade. Interestingly, the *Apocynaceae* and *Solanaceae* together with the *Liliaceae* and *Buxaceae* are the only four plant families known to produce glycoalkaloids. Another common feature that unites these three transcriptional regulatory systems could be that an old regulator was recruited in order to possess a common expression pattern required to generate molecules with related defense activity. The temporal and spatial activation of these pathways at the transcriptional level as well as the function of these alkaloids might therefore be an important evolutionary link. In relation to this, all three pathways are known to be part of a chemical defense response that is mediated by the stress hormone jasmonic acid (Choi et al., 1994; Aerts et al., 1994; Baldwin et al., 1994; Ohnmeiss et al., 1997).

Recently, Hatlestad et al. (2015) reported a similar phenomenon in which highly related R2R3-MYB transcription factors regulate the biosynthesis of structurally unrelated plant pigments, the anthocyanins and betalains (produced from phenylalanine and tyrosine, respectevaly). The authors suggested that a novel pigment pathway has coopted an ancestral, anthocyanin transcription factor, therefore employing different chemistry for providing similar functionality (e.g. in pigmentation of flowers, fruit and epidermis).

Biosynthesis of cholesterol, the precursor pathway for SGAs is under *GAME9* control

The precursor for SGA biosynthesis is cholesterol, which undergoes several hydroxylation, oxidation, transamination and glycosylation steps to generate the SGA chemical diversity (Eich, 2008; Ohyama et al., 2013; Itkin et al., 2013). While still far from being resolved, cholesterol biosynthesis in plants is predicted to be a multi-step branch from cycloartenol in the mevalonate pathway. Recently, the first committed enzyme in the cholesterol pathway, SSR2, was described in potato and tomato (Sawai et al., 2014). Several studies demonstrated the tight crosstalk between the cholesterol and C-24 alkylsterol pathways in SGA-producing plants (Arnqvist et al., 2003; Ginzberg et al., 2012; Sawai et al., 2014). The SSR2 reaction is therefore a junction for controlling flux towards cholesterol and downstream to SGA biosynthesis. The enzyme SMT1, catalyzing the alternate branching reaction in which cycloartenol is trans-methylated to 24-methylenecycloartanol, is not less important in maintaining the balance between the two pathways (Diener et al., 2000). Our results with *GAME9*-Ox plants in both tomato and potato showed that GAME9 is most likely involved in regulation of *SSR2* but is not associated with *SMT1* expression. Apart from *SSR2*, three additional genes including homologs of those encoding a *Δ(7)-STEROL-C5(6)-DESATURASE*, *METHYLSTEROL MONOOXYGENASE 2-2-LIKE* and a *3-β HYDROXYSTEROID DEHYDROGENASE* could be associated with cholesterol biosynthesis as their expression was affected very significantly in the tomato *GAME9*-altered transgenic lines. Out of these 4 genes, the promoter region of *Δ(7)-STEROL-C5(6)-DESATURASE* and the one of an additional candidate gene showing homology to a *Δ14-STEROL REDUCTASE* (that is also predicted to act in cholesterol biosynthesis) were transactivated by GAME9 in our transient assays.

It appears that genes encoding enzymes in the mevalonate pathway, upstream of the SSR2-SMT1 branch point, are also under some level of control by the GAME9 transcription factor. This was evidenced in the tomato *GAME9*-Ox and *GAME9*-RNAi lines that showed a significant change in *HMGR* expression. In potato *GAME9*-Ox lines, a dramatic increase in *CAS* but not in *HMGR*

expression was observed. Expression of the gene encoding SQS, an enzyme downstream HMGR in the mevalonate pathway, was not altered in either the tomato or potato transgenic plants. Yet, it cannot be ruled out that altered expression of the mevalonate pathway genes may be a secondary response to altered expression of the downstream SGA pathway genes, and thus reflect not a direct regulatory effect of GAME9.

It is hence apparent that GAME9 control of SGA biosynthesis is not restricted to the *GAME* genes of the core pathway between cholesterol and *α*-tomatine, but it includes the upstream biosynthetic genes of the cholesterol and possibly of the mevalonate pathways. This is likely crucial for ensuring the flux of precursors in times of SGA production and to maintain the homeostasis in the interface between the cholesterol pathway and the essential phytosterol biosynthetic pathway. Likewise, the *Catharanthus* ORCA3 was shown to activate several TIA biosynthetic genes as well as some primary metabolism genes involved in the synthesis of TIA precursors (van der Fits and Memelink, 2000).

GAME9 action as part of the transcriptional regulation of SGA biosynthesis in tomato

It was previously reported that group IXa ERFs proteins from several plant species possess similar but diverse DNA-binding specificities and that each can differentially bind to multiple GC-rich sequences (Shoji et al., 2013). At least three different GC-rich boxes can be recognized in promoters of these transcription factors target genes: a P-box (5'-CCGCCCTCCA-3'), a CS1 box (5'-TAGACCGCCT-3') and a GCC box (5'-AGCCGCC-3'). We performed transactivation assays by testing combinations of GAME9 and upstream regions (i.e. putative promoters) of core SGA biosynthetic genes, mevalonate and cholesterol (i.e. SSR2) pathway genes as well as of other candidates altered in both *GAME9*-Ox and *GAME9*-RNAi tomato plants. The results suggested that GAME9 does not activate the core SGA pathway genes in spite of finding that putative promoters of genes including *GAME9*, *GAME17*, *GAME7*, *GAME4*, *CELLULOSE SYNTHASE*, *E3 UBIQUITIN LIGASE* and *SSR2*, contain a P-box in the upstream region (Supplemental Figure 3). Nevertheless, GAME9 significantly activated the SSR2 upstream region as well as those of two genes possibly involved in the biosynthesis of cholesterol, namely, a *Δ14-STEROL REDUCTASE* and a *Δ(7)-STEROL-C5(6)-DESATURASE* (Sonawane and Aharoni, unpublished data). Both *Δ(7)-STEROL-C5(6)-DESATURASE* and *SSR2* genes contain a P-box in their upstream regions.

Apart from acting directly on the above pair of promoters, GAME9 might be acting indirectly through an intermediate transcription factor that by itself directly activates the promoters

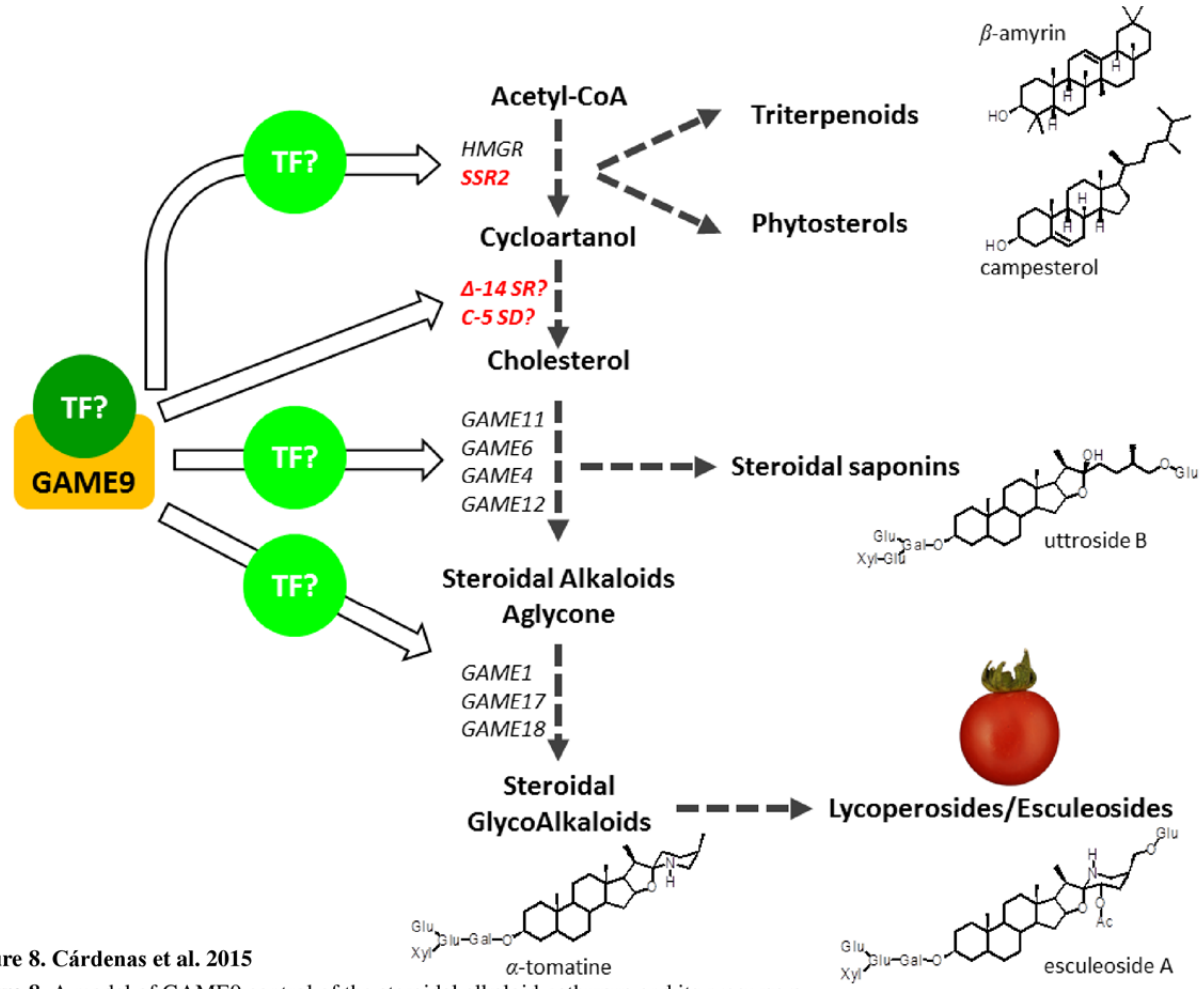


Figure 8. Cárdenas et al. 2015

Figure 8. A model of GAME9 control of the steroidal alkaloid pathways and its precursors. GAME9 activates the synthesis of sterols precursors and SGAs in potato and tomato. GAME9 might activate genes across the pathway by directly binding the promoters, forming a complex with a second transcription factor or activating a downstream transcription factor. Promoter transactivation assays showed that GAME9 might be partly regulating SGA biosynthesis by directly activating SSR2 and the putative cholesterol gene candidates, *Δ14-STEROL REDUCTASE* (Δ14-SR) and *Δ(7)-STEROL-C5(6)-DESATURASE* (C5-SD) marked in red.

of core SGA genes (Figure 8). In a different scenario, GAME9 requires an interacting factor and co-binding of both regulators to the promoter region of target genes in order to permit target gene activation. Such an interacting factor might be the MYC2 protein, a jasmonate signaling component shown to take part in activating the tobacco nicotine and *Catharanthus* TIA biosynthetic pathways together with the *NIC2* locus protein ERF189 and the ORCA3 protein, respectively. In the current model of transcriptional regulation of tobacco nicotine biosynthesis, when the bioactive jasmonate is perceived (i.e. JA–Ile) an active MYC2 is liberated (Shoji and Hashimoto, 2011). The *NIC2* locus ERF proteins recognize the GCC-box and activate structural genes in cooperation with NtMYC2 that recognizes the G-box element. NtMYC2 also induces, directly or indirectly, the *NIC2* locus *ERF* genes (e.g. ERF189). In *Catharanthus*, MYC2 and ORCA3 factors likely act in a

transcriptional cascade to regulate TIA biosynthetic genes and no evidence is available suggesting direct interaction of both proteins on the promoters of biosynthesis genes in a cooperative manner (Zhang et al., 2011).

The major potato SGAs are considered anti-nutritional factors for humans and their levels in tubers of commercial potato varieties is limited by law (Eich, 2008; Roddick, 1996). One approach to select low alkaloid potato lines is the identification of associated QTLs and carrying out marker-assisted selection. GAME9 is likely the gene underlying the major QTL on chromosome 1, reported by Sørensen et al. (2008) to explain 75% of the variance in SGA content among tubers. Hence, the identification of GAME9 in this study provides a platform for the generation of *Solanaceae* crops with modified levels of SGAs. Furthermore, GAME9 provides a starting point for the elucidation of signaling and transcriptional regulatory networks that mediate constitutive and induced SGA biosynthesis in the *Solanaceae*.

METHODS

Plant material and generation of transgenic plants

Tomato plants (*Solanum lycopersicum*) *cv*. MicroTom and potato (*Solanum tuberosum*) *cv.* Desiree were grown in a climate-controlled greenhouse at 24°C during the day and 18°C during night, with natural light. The *GAME9*-RNAi construct was created by introducing a *GAME9* fragment to pENTR/D-TOPO (Invitrogen) (by *Not*I and *Asc*I) and further transfer of the resulting plasmid to the pK7GWIWG2 (II) binary vector (Karimi et al., 2002) using Gateway LR Clonase II enzyme mix (Invitrogen). The *GAME9*-Ox constructs were generated by introducing the corresponding tomato and potato *GAME9* coding sequences into pDONR221 using the Gateway BP Clonase II enzyme mix (Invitrogen) and then transferred to the pJCV52 binary vector using Gateway LR Clonase II enzyme mix. Constructs were transformed into tomato and potato as described previously (Itkin et al., 2011, 2013). Primers used in this work are listed in Supplemental Table 8.

Co-expression analyses

Co-expression analyses were done as previously described by Itkin et al. (2013). Briefly, the tomato *GAME9* (Solyc01g090340) and its potato ortholog (Sotub01g029510) were used as 'baits' in co-expression analyses, resulting in lists of co-expressed genes (r-value ≥ 0.8) for each bait, separately and shared homologs between the two species. The analyses were performed using tomato RNA-Seq transcriptome data from different tissues and organs (flesh, peel, seeds, roots, leaves, buds and flowers) and developmental stages (20 experiments in total; Itkin et al., 2011) and potato RNA-Seq transcriptome data from different tissues and organs (40 experiments in total; Xu et al., 2011). The co-expression network was visualized with the Cytoscape program (Shannon et al., 2003).

In situ RNA hybridization

In situ hybridization was performed as described previously (Pekker et al., 2005). Tomato shoot meristems were collected 13 days after germination. *GAME9* transcripts were detected with a full-length cDNA DIG (digoxigenin)-labeled *GAME9* antisense probe.

Phylogenetic analysis

Literature search was performed to identify functionally characterized proteins belonging to the ERF family of transcription factors. Amino acid sequences were aligned using ClustalW2 (Larkin et

al., 2007). A phylogenetic tree was built using the neighbor-joining method (Saitou and Nei, 1987) implemented in MEGA6 (Tamura et al., 2013). The analysis involved 50 amino acid sequences and evolutionary distances are in units of number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. Accession numbers for sequence data used in this tree can be found in Supplemental Table 7.

Preparation of plant extracts and metabolite analysis

Profiling of phytosterols in potato tissues was carried out as described previously (Itkin et al., 2011). Briefly, after extraction and saponification, compounds were derivatized with MSTFA and analyzed with GC-MS. Preparation of extracts for SAs analysis was performed as in Itkin et al. (2011) with the following modifications: potato and tomato extracts were diluted 80 and 50-fold, respectevaly, before injection. Compounds were analyzed in MRM positive mode using a UPLC-TQ-MS (Waters), equipped with Acquity BEH C18 column and Triple Quadrupole MS detector. Mobile phases A and B, column temperature and flow rate were set as described in Mintz-Oron et al. (2008). For potato samples, α-solanine and α-chaconine were isocratically eluted at 20% B for 10.5 min., the column washed with 100%B for 3.5 min. and re-equilibrated at 20%B for 1 min.. The following MS parameters were applied: capillary voltage 2.7 kV, cone – 61 V, collision – 65 eV. Relative quantification was done using the TargetLynx program (Waters), using the sum of two MRM transitions for α-solanine (868.5>398.4, 868.5>706.5) and α-chaconine (852.5>398.4, 852.5>706.5). For tomato samples, the following linear gradient was applied for α-tomatine analysis: 15% to 30% B over 5 min, 30% to 50% B over 10.5 min, 50% to 100% B over 0.5 min, held at 100% B for a further 1.5 min, then returned to the initial conditions (15% B) in 0.2 min and conditioning at 15% B for 1.3 min. MS parameters: capillary – 2.72 kV, cone – 60 V, collision energy – 40eV. MRM transitions were set as 1034.5>416.3 and 1034.5>578.3. The first transition trace was used for α-tomatine quantification.

Quantitative Real-Time PCR

RNA isolation was performed by the Trizol method (Sigma-Aldrich). DNase I (Sigma-Aldrich)-treated RNA was reverse transcribed using a high-capacity cDNA reverse transcription kit (Applied Biosystems). Gene-specific oligonucleotides were designed with Primer Express 2 software (Applied Biosystems). The *TIP41* gene (Expósito-Rodríguez et al., 2008) was used as an

519 endogenous control for tomato samples and the *NAC* gene (Ginzberg et al., 2012) was used for
520 potato. Oligonucleotides used are listed in Supplemental Table 8.
521
522 **RNA-Seq library preparation and sequencing**
523 RNA-Seq libraries were prepared as described by Zhong et al. (2011) with minor modifications.
524 Briefly, 5 µg of total RNA was used for poly(A) RNA capture using Dynabeads Oligo $(dT)_{25}$
525 (Invitrogen), fragmented at 94°C for 5 minutes and eluted. The first-strand cDNA was synthesized
526 using reverse transcriptase SuperScript III (Invitrogen) with random primers and dNTP, whereas the
527 second-strand cDNA was generated using DNA polymerase I (Enzymatics) using dUTP. After end-
528 repair (Enzymatics), dA-tailing with Klenow 3'-5' (Enzymatics) and adapter ligation (Quick T4
529 DNA Ligase, NEB), the dUTP-containing second-strand was digested by uracil DNA glycosylase
530 (Enzymatics). The resulting first-strand adaptor-ligated cDNA was used for PCR enrichment
531 (NEBNext High-Fidelity PCR Master Mix, NEB) for 14 cycles. Indexed libraries were pooled and
532 sequenced.
533
534 **Transient expression assays**
535 Transient expression assays in *Nicotiana tabacum* protoplasts were performed as described
536 previously (De Sutter et al., 2005; Vanden Bossche et al., 2013). Briefly, protoplasts prepared from
537 tobacco Bright Yellow-2 (BY-2) cells were transfected with three different plasmids. The first
538 plasmid (reporter plasmid) contained the firefly luciferase (fLUC) gene under control of the
539 investigated promoter; the second plasmid (effector plasmid) contained the ERF transcription factor
540 GAME9 driven by the cauliflower mosaic virus 35S promoter (pCaMV35S) and the third plasmid
541 (normalizer plasmid) contained the renilla luciferase (rLUC) under pCaMV35S control. After
542 transfection and overnight incubation, the protoplasts were lyzed and both fLUC and rLUC
543 activities were measured with the Dual-Luciferase® Reporter Assay System (Promega). The fLUC
544 activity is a measure of the activity of the investigated promoter, whereas the rLUC activity reflects
545 the transfection efficiency. For normalization, the fLUC value of each independent transfection was
546 divided by the corresponding rLUC value. For screening and confirmation experiments, 4 and 8
547 transfections were performed for each promoter-GAME9 combination, respectively, and the
548 obtained normalized fLUC values were averaged and compared relative to the values obtained from
549 transfections with an effector plasmid containing the GUS gene.
550

ACKNOWLEDGMENTS

This work was supported by the European Union Seventh Framework Program FP7/2007-2013 under grant agreement number 613692–TRIFORC (AG and AA partners) and the Israel Ministry of Agriculture cooperation with Italy grant (for AA; agency reference number: 261-0929-13). We thank the Adelis Foundation, Leona M. and Harry B. Helmsley Charitable Trust, Jeanne and Joseph Nissim Foundation for Life Sciences, Tom and Sondra Rykoff Family Foundation Research and the Raymond Burton Plant Genome Research Fund for supporting the AA lab activity. PDC thanks Becas Chile Program (CONICYT, Chile) for PhD financial support. AA is the incumbent of the Peter J. Cohn Professorial Chair. JP is a postdoctoral fellow of the Research Foundation Flanders (FWO).

AUTHOR CONTRIBUTIONS

PDC designed and performed the research, analyzed metabolomics and transcriptomics data, and wrote the article. PDS performed the promoter cloning and assisted in the RNA-sequencing and metabolomics. JP, RVB and AG performed and analyzed the promoter transactivation assays. EW assisted in the co-expression and RNA-sequencing data analysis. LT performed the *in situ* hybridization experiments. SM and IR assisted with metabolomics data analysis and operated the LC-MS and GC-MS. SB designed part of the research and wrote the article. AA designed the research and wrote the article.

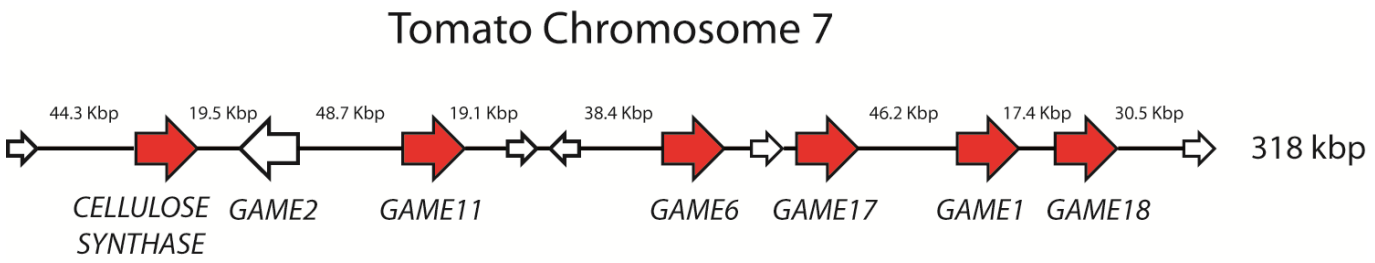
Table 1. Details of shared homologs in transcriptome data of *GAME9*-RNAi and *GAME9*-Ox tomato lines.

Solyc ID	Annotation
Solyc01g090340	Ethylene responsive transcription factor (GAME9)
Solyc07g043490[a]	UDP-galactosyltransferase (GAME1)[b]
Solyc07g043480[a]	UDP-glucosyltransferase (GAME17)[c]
Solyc07g043420[a]	2-oxoglutarate-dependent dioxygenase (GAME11)[b]
Solyc07g043460[a]	CYP72A (GAME6)[b]
Solyc07g043500[a]	UDP-glucosyltransferase (GAME18)[c]
Solyc02g069490	Sterol side chain reductase 2 (SSR2)[b]
Solyc02g086180	Δ(7)-sterol-C5(6)-desaturase (C-5 SD)[c]
Solyc06g005750	Methylsterol monooxygenase 2-2-like (SMO1)[c]
Solyc01g073640	3-*β* hydroxysteroid dehydrogenase (OXR)[c]
Solyc10g008410	E3 ubiquitin-protein ligase RMA1H1-like (UBL)[b]
Solyc07g043390[a]	Cellulose synthase[b]
Solyc09g018660	Unknown protein
Solyc03g051690	Alcohol dehydrogenase (fragment)
Solyc10g074660	Unknown protein
Solyc12g032960	Photosystem I P700 chlorophyll a apoprotein
Solyc09g074480	Unknown protein
Solyc09g009070	Repressor of silencing 1
Solyc01g017320	Photosystem I P700 chlorophyll a apoprotein A1
Solyc06g009950	Photosystem I P700 chlorophyll a apoprotein A2
Solyc03g078030	Unknown protein
Solyc02g036320	Photosystem I P700 chlorophyll a apoprotein
Solyc05g007830	Expansin 2
Solyc04g007480	Unknown Protein
Solyc06g074590	Ycf2
Solyc01g007390	Unknown protein
Solyc12g056290	Unknown protein

[a]Genes localized in cluster in chromosome 7 associated to SGA biosynthesis.

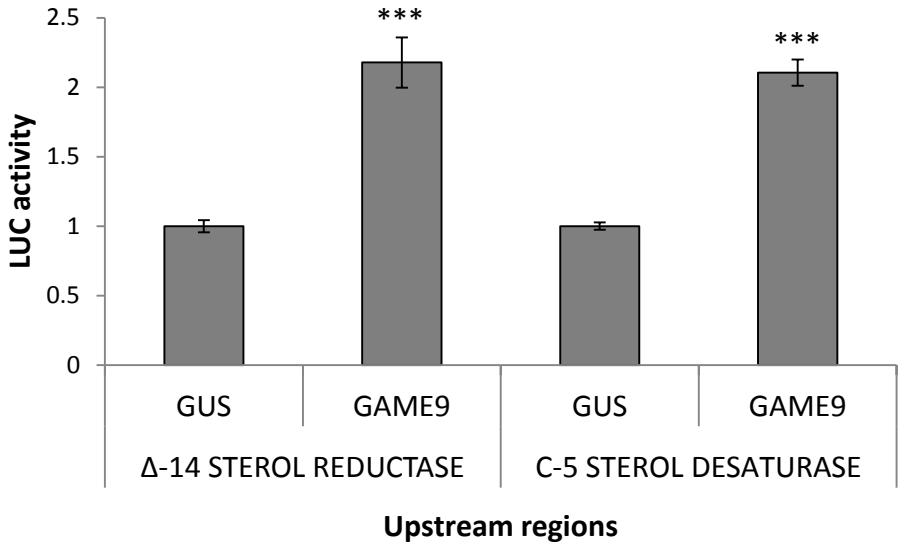
[b]Genes co-expressed with *GAME9* in potato and tomato.

[c]Genes co-expressed with *GAME9* in tomato.



Supplemental Figure 1. Schematic map of altered SGA-related genes in tomato *GAME9*-RNAi and *GAME9*-Ox transcriptome.

Among a set of 27 common genes in the *GAME9*-RNAi and *GAME9*-Ox tomato lines, we found a significant representation of SGA biosynthetic genes (*GAME'*s) located in chromosome 7 (*GAME*'s- *1*, *6*, *11*, *17* and *18*) and a *CELLULOSE SYNTHASE*. Altered genes in the region (spanning ~318 kbp) are highlighted in red.



Supplemental Figure 2. Cárdenas et al. 2015

Supplemental Figure 2. Validation of transactivation assays of putative downstream gene promoters by the GAME9 transcription factor. Validation experiment (n=8) was performed, confirming the transactivation of the *C5-SD* and *Δ14-SR* promoters. Values in the y-axis are normalized fold-changes relative to protoplasts cotransfected with the reporter constructs and a pCaMV35S:GUS (GUS) control plasmid. For the normalization procedure, see Materials and Methods. Details for each promoter are provided in Supplemental Table 6. Student's *t*-test was used to assess whether the transgenic lines significantly differ from wild-type plants: ***P value < 0.001.

GAME4 (Solyc12g006460)

Promoter

>SL2.40ch12:941703..938704
TCATGACGTATGACACGTCCCTTGAATCATAATTCCTCAAAATGATGCATCCTCTTTAATATCTTATACTTCCTCAATTCACATAAAACTTGTCACAAC
CATATCTCATAATAAATGCAAATGACATGTTCACACAATAAATTTGAAGATAATCATTTTTATAATAATGCACAATTTACACATTACAATTGTGATAG
AACATCAATGATGAATCAACAAACTTTTCAAATCATTCTCTACATATCACACACCAATAATCAATTACATTGTCTTTTATTAAACTTATTCAATATGGA
CAACTCAGCACATTACCAAGTCTCATTACTCCTCAACACATACCACAAGAAAATAAACAATTCTATACACTTACCGAGGGTTGAGAAATTCACTTGCC
TTAAACAATCTAATAATTAATTTGAGACTTGAACCTTCCTCTTTCGTTGGGCTTCCAAATTAATGCAATCTATTCGCATAAATAATCACAATGAGATTT
CTAAACTAACAACATCCATAATACTACATGTATATCCTAGACCTAAAACCTCACCAAATTCTATAAAGTTTCTAATCTTGATGCCAATTAATACGTCAA
TTCGCATAAATTTTAAATTTTAAGCTTAGAGTTGAGTCTCAACTTTCTCCAAATATAATATAATTAGTGATATTCATATAATATAATACATCTAATATTT
AATTAGATATCTCAATATATTAAACTTAATTAACATGTGATAACATATAAAAAAAAACATCAAAACCCCATGTGAAGTCTGGCGCTAATCTACCACCA
AAACAAATATGCATAATTTTCTTGCCCTAATTGAAATAATATGACATCATTCCTTAAAAAAAATTGACAATCACCATGCTATGATTGCAAACCAATGTG
ATCTTGTGAACAATAAAATTAATATATAATTTTTAGCGGTAATAAATAAATATATTAATAAAAAATGTTAAAATCTTTATCGGTATTAATTAAATGTC
ATTAGATCCAATATATTTAAAGCCTTTAGGTCATATAAAATGAGTATTAAATGCCATTAAAAATGAAAATTTAACAATAATTTCTAATTAATTACCGC
TAAAAATTCACACAAAAATAAACAATTGAATAGCCCATTGCTCGAATACGGTGCTTTACTCCAAGTCACACTTAACTCCATTTAATTTGATTCACTTAT
ACCTATGGGTCAAAGTGATAGAGGATATTAAAAACCAACTGAAATAATTATTAAAAAAATACTCATCAACTAAATATTTTTAGTCAAACTTGAATAA
TTTAAATTTTTAAATCAACTTTATTGACTAATCAAAATGACTCGGGACAACTATCGAAAGAAATAAAAATATTTATTTTTGAATTTTTTTGAAAATGGC
CTTTCGAGTCGTTATATTTAGACCAATAACAAAAAAAAATCACAAAAAGCAATAAACATAACATTTGATTGTTTCTTTTATTTTTTTAATTAATAATAT
TTGCATAAATTATGTGCAAAATTTTATATTATTTTACAGAAGGATATAATATGGTATGAAAAATGGGTTTTATAAGTAAATATGCTTAAAAGCTAAAA
TAACAAAATTAATAGTGTGATTGACTAAATTTTTTTAGAAAAATATTATAGAAAATTACTTTTTATCAAAATATTTTTGTATATTAACAATTTTTGTTGA
CTAATTTTTTTTTTTGATTAATATTAGAGTTCTATTAGTGCTTGATCAAGATTTCAAAATTAATTTTGTTGTTCCAATGATAAAATGTGATTGATATTTT
AAACTATTTAAATAGTTTTTTTTGAATCTGATACTACTAATATTAAATTTAAAGCCAACAACCCCGCAAAATAAATCTTCACGTGACCAATGCATCAA
GTGGAATTTATTAAAATTAAATACGACAAGATATGTAGAAAATTATTGTTACAAAAAATTAGAAAGAAACAAAGTCCATCTTTGAATTCTCTCACATA
AAATATTGGGACCCAACGTTAGAAACATTATTGTTATGGCAAAAGAAAAAATCGGTTATCATCGACGTATCTGATCCAGTAGATAAGATTATTTCTC
TCGTTAATCGTAGGTTTTGAATTTGATTTTTTAGTATTAAAAAAAATTTTTGGTAAGAAATGTCTTTCCCTATCCTTCTAATGGAATCCTATGTAGCGC
GAATTCGAAATAGTCTAATTTCAATTTAATTATTAAACACAAAATGAAAATAAAAACAAAAATCCTCATTTTACTAATAATCATAATAATGATGTAGA
AATTATTAGCTAATCCCAAAATATTAATCATAAATAAAATTTAACTAATAATCAAAATAAATATTATATATAAACCAACAATCAAAACAAAAACTCAT
AGTATAAGACAGTACAACAAGTGTAAAGGTTGTGTAAAAGCTTTTAGGTTTCTAGTCAAATAACTATGATTTTTCTTCAAAAATATCTTAAAGAATT
ATATATATAATCAATATATGATTTTTCATTTTTTTTGGTTTATTATTATTTTAGTTTTTCCCTTTCTTTTAATTTTTTCCATACCTTTTGTCCGTAATCATG
GACCATATCATTATTTGTTAACTGAAATGACAAAAATAAACCAAAATTTAACAGCTAACATAAATAAGTTTTTCTCAAAATTCAACAACACATTTAAA
CCTTTTCATTTAAAAAAAAACAATTTCTTTAATAATTTTACAAAATCGAAAATATATTTAGTTTATCAATAATGAAACCTTTTAACGAAAGTTGGGGGG
GCTAAATTATATTCACGTGATTGTATTTTTTTTTTTAATTGCACTTCTATTAAGGAATAAAATTATAAAGATTTAACTATCTTTGAAAGTAAAAAGGTT
AGTATTGTCTTTTCACATATTTGGTCTCATCATTATTGTTGCCATATTTTTTTAGATAAATAAATGTAGCCATTTACTACCTCCAATCAAATTCACAAAT
CCATCCAAATTATTCAAAGAAAAAAAAATAGAGAGAGAGAACA

cDNA

ATGGATTTCTACAATTTAGCCTTGTTCTTCATAGCTTTAATACTTGGAATTTTCACATTTTATGCCATATTAATGAGAATAAATGGTTGGTATTATGCA
ATCAAATTTTGTTCAAACAAATATAACATCCCAAATGGTTATATGGGTTTGCCATATTTTGGTAACACACTTTCTTACTTCAAAGCTTCAATGTGTGGT
GATCCAAAATCATTCATTGATTTCTTTGCTACTAGGTTTGGAGAAGGAGGAATGTATAGGGCATACATATTTGGGAAGCCAACAATTATGGTGACA
AAGCCAGAAATAATTAGAAAAGTTTTGATGGATGAAGAGTATCTTGAAAGAGGTTTGCCTAATTATATGAAAAAATTAATTGGATTAACAACTTCG
ATAGAAGAAGACAAATATTTTCGTAGATTAACAGCACCAGTAAAAAGTCATGGATTATTATCTGATTATTTCGATTATATCGATAAAACTGTGAGTT
CTACATTAGAGAAATACGCTACTACGGAAGAACCTGTTGAGTTTCTTCATAAAATGCACAAGCTTACGTTTGAGGTGTTTATGAGACTTTTAATTGG
TGATGAAGTTAATCAAGAATTATTTGATGAAATGTTTGAGGAGATTACTGCTGTAATTAGTGGTGTTCACAATTTGCCAATTAATCTCCCAGGATTT
GCTTATCATAAGGGACTCAAGGCTCGAAAAGTACTAGGAGAGGTATTTAAAAAATTAATTGATGAAAGAAGAGAAGCCATGAAGGATGGAAAATC
AATGCCAAAGGCAAACATAATTGATATGTTGTTATCAAACAACAATCAAGATTATGAAGCAAACATGTTGAGTGACAAGAAGATCATTGAAATCCT
AGTTTTGTTTTCATTTGCTGGTTTTGAACCTGTTGCTCTTATGTCTGTCAAGGCAATTTTCCACTTACAAAAACATCCACATTTCTTGGAAAAAGCCAA
AGAGGAACAAGAGGAAATAGTAAAGAGAAGAGCATCTTCAAATGCTGGACTTAGTTTTGATGAAATTAGACAAATGACATTTGTTAGTAAGATAA
TTAATGAAACGTTACGTATAGCTACTGATCAGTCGGTATTCCTTAGAGACACAAGTACTACTTTTAACATAAATGGGTACACCATACCCAAAGGGTG
GAAGTTTTTTGCAGTTGTATGGAATATTCATATGAATCCTGATGTTTATGTTCAACCTAAGGAATTTAATCCTTCGAGATGGGATGATATTGAAACTA
AGCCAGGCATTTTTCTACCTTTTTCAATGGGCCCCAAATCATGCCCAGGATCCAATTTGGCCAAGCTTCAAATTTCAGTAATTCTTCATTATTATCTTC
TTCACTACAGGGTTGAGCAAATTAATCCAGAGGCTAGATGTTATCCTCCTGAAAATTGTCTTGTGAAATTCAAGAAGCTATCGATCTCTAGTAATGG
TAATTAA

GAME7 (Solyc07g062520)

Promoter

>SL2.40ch07:62440567..62443566
GGTAGTTAGAAGTGTCCAGTGGGTCAAAGTCCCACATTAGTTGGAGAATGAACTGGTGGTCTGCTTATATGAACTTGAGTAATCTTCATCTCATGA
ACTAGCTTTTTCGGTTGAGTCAGACTCCAATATTATACTGGGGTCTTGGATTGAGTCTGACGGCCAAACATCTCTATTATCAAGAAGAACAAACAAT
GATTGTGATTACTGAAGCTGATTTAATTGAGGATGATGATATAAGCTAAACATAATCCAAATTAAACGCGTTCTTAGTTATACATATGCTTGCAGGG
AAAAAGTCTTTTATTTGGTTGTGTCCAAAACCACAAGTACTAATCATGGATCCGGAGCTTATAAAGGAAGTATTCTCGAAAAATTATTTTTATCAAAA
GCCTCATGGAAATCCATTTGCAGCGTTATTGGTACAAGGACTAGCAGCCTATGAGGAAGACAAATGGGCCAAACATAGAAAAATTATCAATCCCGC
TTTCCATCTAGAGAAGCTAAAGGTTTGTCAGACGTTTCTTCGTTTTTCCTCCTTTAACGTATGACTATTTCTAGTAAATTCATCATTTGTCAAGACTCTT
CTACAGTTCTATGTAAGTTCATAGGCTTGGACGGGGCATTTAGTCTATGTAAGTTGATAATGTCTTAGATATATTACCGTAAAATGAATCTTTTTTAT
ATTAAAGAGATTTGTAATATTGCTGAACGTGTTTCACCTTAATTGTTGTTTGAACGCTCTTTCTTTTTATTAAACTAGATGGTGATGGTATCACTATA
TCTGTATTCCTAATTCTGTTTCAAGTGTTGTGTCATCGTTATATGTTCCCTTACATCATTTTCTCGAACAGCATATGCTTCCAGCTTTTCACTTGAGTTG
TAGTGAGATGTTGAGTAAATGGGAAGACGCTGTTGAAGGCTCTTGTGAGATAGATGTATGGCCTCACCTTCAGCAATTAACTTGTGATGTGATCTC
TCGTACAGCTTTTGGTAGTAGTTACGAAGAAGGTAGAAAGATTTTTGAACTTCAAAGGGAACAAGCTCAGCATTTTATCGACACTGTACGCTCAGTT
AATATCCCAAGATGGAGGTAAATTAATTTGCCTGAGAATCGAAAAATTGTAGCCAAAGCTATGTTGTTCGGACTCTTAAAAATGATGATGGTTGCAT
GTTGGATCTTCCAAATATTAGTGCATTTCTATAGAATCCGACATGGTTGCAGGAACTATTTTGGAGAGTCCGAACAACATAGAACCAAAGTTCATCA
GGAGATGCACATAGCTTTTACTTCATCTAATGCATTTTGGTTTTCTGTCTTATAATATAATTTCAAAGATTTTTGCCAACAAAGAGGAACAGAAGAAT
GAAGGAAATAAAAAACGAGGTTCGGACCTCAATGAAAGGTATTATCGATAAAAGAATGAAGGCAATGAAAGCAGGTCATGATGCCGATAATGAG
GATCTGTTAGGCGTATTGCTTGAATCAAATTTCAAAGAAATTGAACAGCACGGAAACAAGGATTTTGGAATGAGCATCGAAGAAGTCATCGAAGA
ATGCAAGTTATTCTATTTTGCTGGCCAAGAAACTATAGCAGTCTTACTCGTGTGGACTCTGGTGTTGTTAAGTAGGTATCAAGATTGGCAGACACAG
GCCAGAGAAGAAGTCTTGCAAGTGTTTGGGAGTCGGAAACCAGATTTTGATGGATTAAATCATCTAAAAGTTGTGAGTATTCTGTTTTTTTTAATTC
TTGGTTTAGTTCATGTGCCATGAGTTCCTTCTCTCATTGTTATTGCTATTGCTTGAGAGCTTTGTGATACTAATGTTTCATTGTCCACGCACCAGATTC
TAAGCACTGGGCGTGAGGTGGGGTGTTAAAAGATGGAGCAGATTCCTTATAAAGCTTGGTCAATCCTCCTCTCTTTGAGCTAGCTTTTGCGGTGTG
AGTTAGACCTAAGACCTAATTTCACAATTACAACTTGTAACTTACAAGCTAACGATGATCTTGTACGAGTCATTAGATATAACTGAAGAAGAAGCCT
TTTTTTTGCAGGTAACGATGATCTTGTACGAGTCACTAAGGCTATATTCCCCATTATTTACACATAACAGACAGGCAAGTGAAGACACAGTATTAGT
AGAATTGTGTCTACCAGCTGATGTGCTAGTTTCATTGCCCACGATCATATTGCATCATGACAAAGATATATGGGGTGAAGACGCAAACAAGTTCAAT
CCAGAGAGATTTAGAGAAGGAATATCAAGTGCAACAATGGGTAAAGTTACATATTTTCCATTTTCATGGGGTCCAAGAATATGCATTGGACAAAAT
TTTGCAATGTTAGAGGCAAAGATGACTCTCTGTATGATCCTACAAAGCTTCTCTTTTGAATTGTCTCCATCCTACACACATGCGCCTCAACCCCTACTT
ACCATTCATCCTCAGTATGGTGCTCCTCTTATTTTCCACAAACTATAATTTCATATTACGTGAGGTATCTTTCACTTGTACGACTATACTGGATATGTTA
TTGTTCAGACAAACTCAATAACTCTTGTTTAAACCTTTTATCTGTACCAGAGAATAAATAACATGTTGCAAATTATATACAAATCATCTCAAGAACCA
ATGAGAAGCTTATTGTTATGACAAATTTAGATATATTCTACAAACACATTAAGCTAAAGAATTTGTGGAAAAAATAACTGTTTTTTTTTTTGGTCACA
ATAAAAACATGTGACATGCTAGTAAGCAGTAAGCACAGCCGGCCTTTTTGGCAATGAAACCACTACATACACTTGTACACTAATTCTTTTTTCAT
TTTTTAAACTAAAATTATGAATGTCAGAATACGAAATAACAAGTAAAAACAAACGAAAATGATTAAAAATATAAGAGAAAATCAGAGCCACTCATT
GGTGTCTTTGCTCATCATTGTTGACGACGTCGTTTGTTAGTTCCCAGAGACATAGTATATATAAATTAATACTCCTCGTTTTTT

cDNA

ACACACAAATTCAACTCAAACAAAATGGACGAAATTCAAATATTAGTAAGAGTGTGTTGTTCTGCCATAGCAATTGCTCTGTTTGTGTGTCTGTGGA
AAGTACTAAATTGGGTTTGGCTCAATCCGAAGAAGTTGGAGAATTTGTTGAGGAAACAAGGGCTAAATGGGAACTCTTACAAAATATTGTATGGG
GATTTGAATGATTTTATTGGAATGATTAGGGAAGCTAGTTCAAAACCTATGAATTTGTCTGATGATATAGCTCCAAGATTGGTGCCTTTCTTTCTTGA
GACCACCAAGAAATATGGGAAAAAATGTTTTATATGGTTGGGTCCAAAACCACAAGTAATAATCATGGACCCTGAGCTTATTAAGGAAGTACTCTC
GAAAACCTATTTGTATCAAAAGCCCGGTGGAAATCCATTTGCAGCATTATTGGTACAAGGACTAGCAACCTATGAAGAAGATAAATGGGCCAAACA
TAGAAAAATTATCAATCCCGCTTTCCATCTAGAGAAGCTTAAGCATATGCTCCCAGCTTTTCACTTGAGCTGTACTGAAATGCTGAGCAAATGGGAA
GATGCTGTTCCACTTGGCAGCTCTCGTGAGATAGATGTATGGCCTCACCTTCAGCAATTAACTTGTGATGTAATCTCTCGGACAGCTTTTGGAAGTA
GTTATGAAGAAGGTAGAAAGATTTTTGAACTTCAAACGGAACAAGCTCAGAATTTTATCGACGCTGTACGTGAAGTTTATATCCCAGGGCGGAGAT
TTTTGCCAACAAAGAGGAACAGAAGAATGAAGGAAATAAAAAACGAAGTTCGGACCTCAGTTAAAGGTATTATCGATAAAAGAATGAAGGCAATG
AAAGCAGGGAATGCCGATAATAATGAGGATCTGTTAGGCATATTGCTTGAATCAAATTTCAAAGAAATTGAACAGCACGGTAACAAGGATTTTGG
AATGAGCATCGAAGAAGTCATCGAAGAATGCAAGTTATTCTATTTTGCTGGCCAAGAAACTACATCAGTCTTACTCGTGTGGACTCTGGTGTTGTTG
AGTCGGTATCAAGATTGGCAGACACAGGCCAGAGAAGAAGTCTTGCAAGTGTTTGAGAGTCGGAAACCAGATTTTGATGGATTAAATCATCTAAA
AGTTGTAACAATGATCTTGTACGAGTCATTACGGCTATATCCCCCATTGATAACACTTAATCGACAGGTTAACGAAGACATAGTATTAGGAGAATTG
TGTCTACCAGCTGGTGTACTAGTCTCATTGCCAATGATCTTATTGCATCATGATAAAGATATATGGGGTGAAGATGCAAACGAGTTCAATCCAGAGA
GATTTAGAGAAGGAATATCAAGTGCAACAAAAGGTAAAGTTACATATTTTCCATTTTCATGGGGTCCAAGAATATGCATTGGACAAAATTTTGCCAT
GTTAGAAGCAAAGATGGCTTTGTGTATGATCCTACAAAGCTTCTCTTTTGAATTGTCTCCATCTTACACACATGCTCCTAAATCCTTAGTAACCATGC
AGCCTCAGTATGGTGCTCCACTTATTTTGCACAAACTTTAGTTTTGTAATGTAATGTAACAAGAATTAATATAAAGACTCAAATAAGGGGTGGAGTT
TGCAATATTAAACAAAGGGAC

GAME9 (Solyc01g090340)

Promoter

>SL2.40ch01:75793630..75790631
TTTTCATTGATTTAAAAAGGTACGAGCATACTAAAGTTCAACTCAATTCTTTAAATGGTCACATATATCATAAAATTGACCACTAAAATTATTTCTATT
TCTTTTCTTCCATTTTTCATTTATAATACTTTAAAAATTTTCACCTTTGTATTTCGTATCGTTTTAATTTTTACTTTAAAATTTATCTAAAAATCATTTTTAC
TTTATTTAAAAATTAATGTTTGATCATGAAAATTTAAAATACTACTTTTCAACTTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
TATATATATATATATATATATATTGTTTTAGAGATTCATTTTACTAAATTAATCTTATTAATTACATTATTAAAAATTATGGTAATTATTAATATCTTAA
GTAGTTTTGATAAGAATATTATGTTTTAATTTTTTTTGTTAAAACAAACAATACAGAGTAACTGAATACTTTTATATTAAAATAATCTATTTTTTTAACA
AAAAATACATTTAAAATATTTTTTTTATCATAACTACATAAGATACTAATAATTAACATTATTTTCTTAAATGTTATAAATAAAACTAATTTAGTAAGAT
GGATCTGTAATATAATACAATTTAAAATAAATTTAAAGGGAAAAAAACAAAAAATAGAAAAAAAAACTGGAAAGAGAAATAAGAGGCAAGGAAT
ATAAATAAGTGTAGGCCTACGAATTCAAATTTAATAATGATAAAGTACTCCCTCATATAGCTACGGTCAGTTGTGTTTCATCCACAATAAAATGTTTC
AAAGTATTCTCTCCATCTCAATTTATGTGACACTTTTTAAATTTTGAGGCTCAAACAAATATTTTTTATCGTAAGTATTTATTATCTTTTAAATATTTTG
AAATGTCAACTGTTATGACTTACGTACTTTTTAGATAAAGGACGTAAAGTCACTACTATAGTTATCCTAAATTTTTAAAAACACATCTAAATTTTGCAA
TCGTTTTATTACTCACTAAACAATATTTAAGTGATATCATTATATCATTTTTCAATCAACCCAAATTTAAAGAAATAAATATATTCATATACCATCATTA
CATGAGACGTGCTAATTTCATTAAAAATATAATTTTCTTTCTCTCTCTATGACTTATTTTTATTTTGATTTTGATTTCAATTTCAATTTCACTTCATCCCCT
CCAACCCCCCTAAAAAAAACTTTTAAGTTGATCCATCCCTCCATTTTTTGCAAACCCTTCACCTCACGCCGACCCACCATGTCAAGTTGATCCACCTCC
CATTTTTTATTTTTTTTCTTCTTCATCTCCAATCAAATACTTTATTTTAATTTTTCTAAATAAAAAATCGTATAATCTTCTAGACTTTATTGAATTATTGAT
GACAAAACTAATTATTTTTCTAAAAAAATTTAAAATTTTGAAATTACTCTTTCTAATATAACTTCAAAAATAGAAATGATAATATTAAAATAATTGGTG
AATTCCCCAAAAATTGAAAGAACTTGATTGAAGTCGGATAAAAAAACTATATAATACCTCCTAATCATCTTGAAATCTAATAGGTTTATATGTAAGA
AAATATTTAGATAAAATTTTAAAATCAAAGAGAATAAAAGTAATAGTAGAAAAAAAAAGATGAAGTTGTGGTAAAAATCGTGAAATTAAGAGTGA
GAAGTTTAAATGGAGATGACAATAATTTTTTTTAAGGAGAAAGAAGAAAGAGAAAAAATAACGAGGAAAAAAGTTAAAATCGTAAGAAAATAAAT
TACTAGTAAAAATAAAATTATATTGTCTATCAAGTGGACAAAAAAATATGTAAAAATATTTGTTCATGATTGTTTTATGGAGGTAATAGGATGCC
CACATAGTCAAGGTAAATTTTATCATGGTAGAGACCCAATAATGTGTGAGTCTCCATTTATTTTTTTGACACATGGACAAGTGGGCCATTTATGTATT
TTTCTAAATACATAAATAGTAAAGAACTTGATTGAAATCGAATTTTTAAAAAACTATATAATATCTCCTAATCATCTTGAAATCTAATGGTTATTAAAT
TGTTCGGATCGTAAGAAAATATTTAGATAATGTTTTAAAATTAAAGAGAGTAAAAGTAATAATCCCTCCGTTTCGAAAAGGATGACCTAGTTTGACT
TGAAACGGAGTTTTAGAAAAAAAGAAATTTTTTTAATCTTGTGGTTCTAAATTAAAGTTCTGTCAAATGTACCAAAATACCCTTTAATCATGTGGCCT
TAAGCATGCCATGTGGAAAAATTAAAGTTAAAGTGTTGCTAAAAAAGAAAAAGCAGACATTCTTATTAAAAAGGAAATACAGTCATTCTTTTTTAAA
CGGAGGGAGTAATATAAGGAAAAGGAAGGTGAGAAGTTTCAAAGAAGATGACAATAAATTTTTGAACAAGAAAGTGAAAAAATAAAAGAAAAAA
GGTGAAAAATAATAAGAAAATATCTTTTATAATAAAATACTTTAAAAAAATAAAATTATATTAGTTATTTTAGGTTTGTCCACTCTTATGTCAAGTGG
ACAATAAATAATAGAATTATAATAATATCAGTTCAAGATTGTTCAGTGAGATAATAAAATTAACGTGTATTTTTGAAAATTTAAAACAATTTCACGTG
TCAGTTTATACATTTTTTTCCTCTATTTTTTAAAGTAGTATATAAAAAAGATTCCATGTAAAAAAATTTTTTGATTAGTCTTAAATTATTTGATTATGGG
AAAAGAAAAAACGAAAAGTGTCGTATAGACGGGGTGAGATGGAAGTAGTAGTATATATAAGTAGCACTAGAGTTCCACTAC

cDNA

TAATATTACATTACATTCATCACATATTATTCATCCAAAACAAGAATGAGTATTGTAATTGATGATGATGAAATCTTCTCTTTACCTAGCCTTGATGAA
CTTGAATCCATCACACATCTTCTTTATGACGACGATTCCGATTTTTTCGAAACTCTTTCCCCAATGAGTTTAGATGTTACAACATTATTGCCTAATATTC
CTACCTCCAATTCAATTGAATCCCCCGTAACACCGGAGGAAACAAAAGAACCATCTGTGGCGTGTGAGGACGCGCCACAAGATTGGAGGCGGTTC
ATAGGGGTGAGGCGGAGGCAGTGGGGCACGTTTTCAGCCGAAATAAGAGATCCAAATAGGAGAGGAGCGAGGCTGTGGCTCGGAACTTATGAG
TCCCCGAGGGATGCAGCATTAGCTTATGACCAAGCCGCTTACAAGATTCGGGGAACCAAAGTTCGGCTTAATTTTCCTGACCTGATTGGCTCGGAC
GTACCTATGCCACCTAGAGTAACGGCTAGGCGTCGTACACGCTCACGCTCACGCTCACCCGAGCCATTAACAACTTCGTCCTCGTCATCCTCATCATC
CTCGTCCTCGTCCTCGTCCTCGTCGGAAAATGGAACGAAGAAAAGGAAAATAGATTTGATAAACTCAATAGCAAAATCCAAATTACTTTGTGGGAT
GGATTTACAAATGTTAATACAAATGTGAGAAAAGAGCAAAGGTTTATTTTTCTTCGTTTGACAATTAAGTACTACGTCGTATAATTAATAGACTCATC
AAGGTCATTGTGTAAATGCACTTCTTTCACGACCTTCTCCTTTATGAGATTGTTATGAATTTTACATTATTTCCTTTATCAACTATATATTTATCGTTTTC
ATACGCGGTGGAGTTCATCTGAATTTCTCTTTCTAAGGTTATATATAGAGAAGGATGTTGAATTTTTCGTCTTCTTTTTTTTATTAAATAAAAAATCTA
TCTTCTACATCAG

Cellulose synthase (Solyc07g043390)

Promoter

>SL2.40ch07:54427095..54430094
AATTAGCAACTCGGTAGCCTTGGAAAGATAATTAAAAAACCTTTCATTTGACATCTTGTAGCCCACCTAACTGTCCAAGTAATAAAAGTTATGATCAT
TTGTTGACCCAAAATTCATACTTATTATTTTCAAAAATGGTTCTTTCTTATTCTAACTCTTTCAAATAGCCGGGTATTACAAATATATCAAAGAGAAAT
ACTCTATAAGTCATGTTAGATATTACTTTTTAATTAAACTAAAATACTATGAATAAATGTCGTATTTTTAAGAATGTCTTTGTTAAAGTATATTATACTT
TTGTTTATTCAATAACATTGCACAAGTAATAGGAGCATTTTGATAGTTTTCACAACTTATGATTCTTTTTATGAGTAAAAAATACAATTTAAAAGTTGA
AATTACTTGCCAGAAGTCCAAAATAAAACTTTCGTTTCATCTCATATGATTTCCTTTCATAATTTCATCTCACATGATTAAATGTTCCTTAATTTTTTTTA
CATTGTTATACAATCTTTATTCAAAATTATACTAGAGAAAATTGCAAAGATTGCTAAGACTAAAAACATAATCACGCAAAATAACTATAGCTTGCTTC
ACTACGATATGTAGCTACATGTTAAAAGTGAAGAGAGTTGAGGGAGATTAAGAGAGGGATAAAAGAGATCAGCGAGATGTAGGGCGTAAAGACA
AAACAAATGAATTGTACTTGTATATTTGCTGGATACATGTTTATATATAACTAGTATGTATATTGTGTATTCCAAAACGATTTGTATAAATGAATAAA
TTGTATCGGCAAAAATAATTATATCACTGTTGTATCAATGAATATAATTATATCAATGTTATATCAGAGAATACAATTATTCAATATTATATCAAGTTT
CACATTTTGTATTTGTCTAATTGATTGCTAACATTTGTGTCAATGTTATATAACGAATAAATGAATTTTTGCACTTTGTATCACTAAAAAGATCTTTAT
AGTTGACATTTATCATTTTTATAATGTATAATATTTATAGATAGTTATGAGTGCTGAAAATTATTTCATGAAAAATTACATATTTTCTGAAAAAGTAAA
CTATTTTTAATTATAATAATAATAATCTTTTTGAACAGGCAAATTTTTAATATTTTATTGTTTATAAAATTAATGAGTAAATAATATTATAGTTTTACTT
TATCCTTAAGAATTATTAAACTTGAAATATATATAAAAAGTTAAGTAAATAATTCATATTCATTGGTCTAATTAAACAATCAAAATGAATCAATTCATT
TTCATTGATTGAAAATTTGACTTTAATAGTACAATAAACAAGGATACAATAATAAATTTACCTAAATTTTTAAGAAATCTAAAATCTAAAATTGTAA
TAGTTAATATTTCATCTGTAATTACTTGTTCACTTTTATTTTTTAGTTGTCCCAAATTACTTCTCTATTTGACAAACTAAGAGTGGATATTTTTTTACTGA
TGTCCTCAATAAATTATTTTAAAAAAGATATAATTTTTTTAAATTATAATTTTTATTTGATCTAATTTATATTTAATAAAAGTAAAATAATAAACTCATT
ATATTAATTTTCTTAATAGAAATGACAATTCAAAAGCATATAATTTAGAGATAATTTAGAACATGTGCACTCATACGAGCGTTGCAATTACATTTAAA
AATTCCGACACGTTTCAGCGATAAAGGTTAAGATATTTTGTCCGACGAGTTAAAATATGACATTAATTTATACTTCAGGGTTTAAACCAGAGAGTTT
GTAAAAATTGTGGCCAAATGTTGTCATATTTATTTTTAGCAGTCACTCACACCATTTCATTTTTTCCTTATATTGTCATGTCTCAATGTGGCATAGTAG
GACTTCATAGAAATTTCGATGATCAAATAGATCAAGGATCCGTTTGGTCCCAAGAATATTTAAAAAAAACTTGAAAAATAAATTAGAAAATGGTGTT
TGATTATATAAATTGACAGATGTATTTATCAAATATCCGAAGTTTATAAATACTAAAAAAAAACTTAGTATTTAAGCCAAGTTTGAATATTTGGAAAT
TTTAAAAGATTCAAGTTTTGATATGAAATATTCATGTTTTTAAAATTAAATTTAATATATTTGAAGATGCTATAAAATTATTAAAAAAATATATAATAA
TTGACAATTCAAAATATTTAAAATAAGTATGAAGAAATTAGATCAAAGATAAAAATAAATAAATTAAATCTAGTAATCCAAAAAGTATCATTTAAAT
TGGGATACATAAATTAATAGGCTCAGAGATCGTTTGGTAGCTTGGTTAGAATTATATAAGTAACGTAGAGATAGTTATTAATGAATCTATACATTAT
TTTAAGCAAATATTAGTTACGTAAAGATTAATCATTTATGTGACTACACATGTATTTATTATTTCATCTTGTATCATGTATAAAATAGTACGTATATAT
TTTCATAACCTATATATGTATCAATTATGCGATGTCTAAATTATCAACCAAATATCATATTAATTCTTATAAATAAATAACTCATTTCTTCTTTACCTAC
CAAACATTATACTATTTTATAAAAAAATATTATATAAATAACCTATTTTTCATCCAACTACCAAGCAACCCATCAACTCATGCTCTCTTGGTTGAATTTA
CCTACACCTTTTTCTTGATAATTCGTCCAAAAAAGAATATTATGTTACTATAATAAAAAATAATTTAATTTTAATTTTTTTATATATTTTTAATTAAATAA
TTTATTATTACACAAATATTTAATAATTATTTTAAATTTTAAACTTTAAAAGTATGTATCAAATACCACCTAAATTAAGACGAATAAAATATTTATTACT
CGGTGAAAACATAATTGATTGTGAACCGTAATGGTTAATTGTTATATATCAAAAATCATTTTAATAAGTTGAAACCTATTAAGGACCAATAAAGCAT
TATCAAATACATTAAGGATCATTTTGGTATTTTTCTATTT

cDNA

TAGTGCAACCAAAAATATTGAAGCCATATATAGGTGTGGTGAAAGAGCAAGAGAGGGGAGAGGGAGTGCAGAGAGAAACCACACCATGAAAAA
AACCATGGAGCTCAACAAAAGCACTGTTCCACAACCTATCACCACCGTATACCGACTCCACATGTTCATCCACTCAATAATCATGCTTGCATTAATAT
ACTACCGTGTATCTAATTTGTTTAAATTCGAAAACATTCTCAGTTTACAAGCACTTGCTTGGGCGCTCATCACTTTTGGTGAATTTAGTTTCATTCTCA
AGTGGTTCTTCGGACAAGGTACTCGTTGGCGCCCCGTTGAACGAGATGTTTTCCCTGAAAACATTACTTGCAAAGATTCCGATCTACCGCCAATTGA
CGTAATGGTATTCACTGCCAATCCTAAGAAAGAGCCAATTGTAGATGTCATGAACACTGTGATATCCGCAATGGCTCTTGATTATCCCACCGATAAA
TTGGCTGTGTATCTCGCTGATGATGGAGGATGTCCATTGTCGTTGTACGCCATGGAACAAGCGTGTTTGTTTGCAAAGCTATGGTTACCTTTCTGTA
GAAACTATGGAATTAAAACGAGATGCCCAAAAGCATTTTTTTCTCCGTTAGGAGATGATGACCGTGTTCTTAAGAATGATGATTTTGCTGCTGAAAT
GAAAGAAATTAAATTGAAATATGAAGAGTTCCAGCAGAAGGTGGAACATGCTGGTGAATCTGGAAAAATCAATGGTAACGTAGTGCCTGATAGAG
CTTCGCTTATTAAGGTAATAAACGAGAGGGAGAACGAAAAGAGTGTGGATGATATGACGAAAATGCCCTTGCTAGTTTATGTATCCCGTGAAAGA
AGATTCAACCGTCTTCATCATTTCAAGGGTGGATCTGCAAATGCTCTACTTCGAGTTTCTGGAATAATGAGTAATGCCCCCTATGTACTGGTGTTAGA
TTGTGATTTCTTCTGTCATGATCCAATATCAGCTAGGAAGGCAATGTGTTTTCATCTTGATCCAAAGCTATCATCTGATTTAGCCTATGTTCAGTTCCC
TCAAGTCTTTTACAATGTCAGCAAGTCAGATATTTATGATGTCAAAATTAGACAGGCTTACAAGACAATATGGCATGGAATGGATGGTATCCAAGG
CCCAGTGTTATCTGGGACTGGTTATTTTCTCAAGAGGAAAGCGTTATACACAAGTCCAGGAGTAAAAGAGGCGTATCTTAGTTCACCGGAAAAGCA
TTTTGGAAGGAGTAAAAGGTTTCTTGCTTCATTAGAGGAGAAAAATGGTTATGTTAAGGCAGATAAAGTCATATCAGAAGATATCATAGAGGAAG
CTAAGATGTTAGCTACTTGTGCATATGAGGATGGCACACATTGGGGTCAAGAGGCAAATTGGTTATTCATACGATTGTCATTTGGAGAGCACTTTTA
CTGGTTATCTATTACACTGCAAAGGGTGGACATCTACTTATTTGTATCCAGACAGGCCATCTTTCTTGGGTTGTGCCCCAGTTGATATGCAAGGTTTC
TCATCACAGCTCATCAAATGGGTTGCTGCACTTACACAAGCTGGTTTATCACATCTCAATCCCATCACTTATGGTTTGAGTAGTAGGATGAGGACTCT
CCAATGCATGTGCTATGCCTATTTGATGTATTTCACTCTTTATTCTTGGGGAATGGTTATGTATGCTAGTGTTCCTTCTATTGGCCTTTTGTTTGACTTC
CAAGTCTATCCTGAGGTACATGATCCGTGGTTTGCAGTGTATGTGATTGCTTTCATATCGACAATTTTGGAGAATATGTCGGAGTCAATTCCAGAAG
GGGGATCAGTTAAAACGTGGTGGATGGAATACAGGGCATTGATGATGATGGGAGTTAGCGCAATATGGTTAGGAGGATTG

GAME17 (Solyc07g043480)

Promoter

>SL2.40ch07:54590070..54593069
AATATGACATCAAACGGTTATATCTTCCTTATACGAGTTTAAAATTAACAAATTCGATGACGTTGTAAAGAAGGTTTCAACATCTTTCATTTGATATG
TTATAGGCCAACTCATCATATACCAAAAATTATGATTATTTAAAATTGACCAAAAACTCTTATTTTAAGCTTAATTAGCAAAATTTCAAATTTTGCTAT
TGCTATTTCTAATTTAGTTCTTTCAGAAAAATAAAAATCTTCTTAGGTCCTCTTTAATAAAAATCTTTTAGGACCATAAAAAAAAATAATGGGGCTCAT
TTTTCAAAATATAGCTGTCGAACATAAATAATGGAGCTTTGATTTAACGtacagcataCATAGATCTGCCAACATAGAAAGATTTTCCTAGTATAAAATT
TTACATATTTTAATAAAATATTTAATTTTAATTTTAATTTTTTCATTTCATCTTTTATAAGATATTTTCAATTTTTTTTGTTTGTTTGTATTTAGTCAAATA
TCTTCCCGTGATTGAGAAGTATTCAATTTACATGAAACAGACTCTAATGAATTCTCATTTCAAGGAGTGGACCAGAGTTCAGAATTGACAAAAATAA
AGTTAATGCCACCCCTTGTTGGACAAATCACGTATACTACAAAATTATTTCATTTTCCTTGTTTATTTTTGATTTTATATTTAATTCTTTTTTACTTTTAT
CATTTTTAACAAATAATAATATTTTTTCCTTATTACGTTCTCAATTTATTTATAAAATTAATGTTTTTGAAAAAGTAAAACCTTGTTAATCAGTAAATTG
ATTTTGAAATATTATCAATTAAAAGAGATAAAATTGTAAATTCACTATGTCAATTATTATTTTTTTGAAAAAAGTTTTGAAATATGTTAAATTTTTAAC
CAAAATTATTATAACAATTCCAAACTTTGAATATAATTTTTTAAGATGTATTCAAGTAAAAAAAATGAAGAACTAAATAGGGTATGAAATATCAAAT
AATAGAAGTAAATTTAATTTTTTGGTTAAATTTATTGTGGATAAATCAAACTAGTATTACTTTTTCTGTTCATTATTTTAAGGAATTGTATTTGTTATAT
ATTATATATAAAATCACGACTTACCCCTTTATAGCAATCCCACTAAAGTGGTTTGTTTCTTTGGTGCACTATACGTGTATATGATTTTATGGCATATCA
CCTGAGGGTCTTGGAGTTGTGTGATAACATACTGCATTGGTCATTACGATGAATATGTTATTGGATTATAATGTGATGAATAATTATGTTTCAAAGT
TATATGAAGCAAAGCTTTACTCATGGTCTCTTATCTTATTTATATGATCGTGAGAGGTTATGGGGCTCCACATGATCATTGCACTTATAGGTTGAGGA
TGAGATTATTGATAAGTGTTTTTTATGTTATGTTTATATGATCTTTTGATTATGAAATGTTCACATGACTTTATGATAATATGTCTTCGCTTATGATCAT
GCTTTATGTTAATGGAATCCTTAACTAGTATATGTTCTATGGTATGTGCATTGAAAAGGGAGAAAGCACCAACTCTTTTAAAGTCCAATTATATATAT
ATATTTATATATATGAAATAAACATTTTAACTTCTTTTAAGTTAAAAATAATAACTATCTATTTTTTAAAAAAAAATAAAAATAAAAATCACACAACAC
ATCCATTTGCGTCAAAGGAAAAGGTATGTAAAATTAAAAAAACAAGTAAATGTGAAAAGAGAAAATGTAAGGTAGTTCTCTAAAATTTCTAAAAGA
TAATTTAATTATGGGGGTGTTTTATCATATTAATACGAAATTTTAAAAAATAAATATCATCATACACTAATGACTGCATGACACGTGATAATTAATTT
AAAATATAATTATATTTATCTTAAATTTTGCTTTATCCTTTTCTTTTTTATTTTTTCTCTTACTTTTTGACTAATTTTCAATTTTTTCTTTAATTTTAAATTT
ACCTTCCCACTTCACCCATCTCCCACGTCAGAACTGACCCGCTCCCCTCATTTTATTTCTTTTTTTCTTCAACTCCCATACTGCATATAGGTGTCACTCTC
TCTACATGTGAACAAACCACACAAGACGCAACATGAATTTAACTCCTTTTCATTATTTATTCTTTTCATTTTTCAAGTATATATTTGGATGAAAATCAT
ATCTTATAATTAAAAATTTATTTATTATTTTAAATCATCACACATGACTTCTATTTAATTAATAAAAAGATATGTATTCACTCAACACATCCATTCGCCC
CCCACACCCCACTCCCAGCACCTATTTAACATACAGTACACCCATTATGCGCACGTGAAAAGGTAGTAAATTCTTCTTCTTCTTCTATATAATAAAAAA
TAAAAACTAAATTTTTAAAATTTTATTTGTGTCATATAAATTGAAACAGTGAAAATAATGTATATTATTAGTGACAATAATTAATGATAAACGACATT
TGATTGAATTTTAAATTTCACACGTGATACATAGATTGAGATTATTATTTGACAATTAAAATTTAAATTAACAAATTACAATAACTGACAACTAAATTT
TTATTTTTATTAAAAGTTTGATTGAATTTCAAAATTTCATGTGTCGAATAAATTGAGATAGAAATAGTTATATATATATATGTATATCAATTGAAAACT
ATTTTAGAAGTATTATAAATCATAATGATTAATATATAATTTAAAGTAGTTAAAATTTAACATGATTCAAAATACATAAATTAATGGACTTATTTATGA
TATTTTCTTCTCAACACATTGAAGTATATATTTGGAAAAATGATGAATCTAAATTTTAGGCTGTCTGTTTAAATTATTATTTTCATATGCAAACATTAA
AAATACGGTTGTTGGTATTCGGTAGTTGTTACTTCATATCTATATATCCTTGTTTAGAGTAACACTCCAATAAGTCTTCAAAAAAAAAAAAAAATACT
AATAGTTATACTCCACAACATCTTCACAC

cDNA

ATGGATAACAAAGATGATGTAGCTGTTGTGATAGCTCCATTACATCTTCAAAGCCATCTCCGACAACTTCTCCATTTTGCATGTAGAATTTCCTCCTA
TGGTCTACCGGTCTACTATCTCGGCTTAGCCGCCTCTAACCGTGAAGCTAGGCAACACTCGACTACCTTAAATCCTTGTGACATAGAGAAAATCCAC
TTCCATGACCTCCAAATTCCTAATGAAGATCCTCCTCCAGATATTCCCATGCATATTTGGCATGCTTCTATGCGAACACGTGAACCCATTGCTTCTTTC
CTTGGAGATATCTCCTCAAAGGCAAGACGAATTGTTGTTGTTCATGATCTTTTAATGTCCTATAATGTTCAGGATATTTCTTCATATCCCAATGGAGA
ATCCTATGTATTTCATTGCCTACCTATTTTCGATATGTATTGTAGTCACTATGCCCCAGCTGCAGGATTCCCTATTCCACTTGAAGAAGCGCTTCTTAA
AAGGCTACCTTCCAATGATGGATGTTACAACCCTGAGGATGTGGAGCACCATGCTAAGTATCTAAATCATTGCATGGGCAAGAATGCTGGTGATAT
CTTTAACACAAGTCAAGTAATTGATGGTACTGCTATTGTTGATTCCATGGCAAATCTTGCCACCATGCAAAACAAAAAGTTATGGGCATTAGGACCA
ATTCTCCTTACACAAGATCATGATAAGGTCCAAACCAAACACTTCTGTTTGGATTGGCTTAACAAACAGCCTCCTAAATCAGTTATTTATGTGTCTTTT
GGAACATCTACTTCATTTTCAGCTGAACAAATCAAGGAGCTAGCAATTGGTTTAGAGCTAAGCAAACAAAAGTTCATATGGGTACTTAGAAATGCT
GATGTAGGAGATCCTGTTAACAACAAATGCGAAGAGAATAGATCAGGAACACTTGAGTTGCCAGAAGGATTTGAAGAAAGGGTAAAAGGGGTTG
GCTTAGTGGTGAGAGAATGGGCACCTCAACAAGAAATCTTCGCTCATCCCTCTACTGGTGGCTTCATGAGTCACTGTGGATGGACTTCGTGTTTAGA
AAGCATTATTGAAGGTGTGCCAATAGCCGCTTGGCCTATCCAATTTGATCAACCAAAAAATGCTTTTCTAGTCGCGGAAGTGCTAAAAATAGGGCTC
CATGTTAGGGAGTGGGAGCAGCGGAATGAGTTAGTGACTGCTTCCACCGTTCAAAATATAGTTTGCAAATTAATGGCATCGGAAGAAGGTGATAT
GATTAGGAAAAGAGCACAAGAACTAAGGACGGCTGTAAAGCGGTCCACGGAGGAAGGTGGTGTTTCTCGAATGGAGTTGGATTCTTTCATCGCGC
ATATCACAAGATAG

SSR2 (Solyc02g069940)

Promoter

>SL2.40ch02:33940599..33937600
ATTTATTAAATTAAATTTTAATTCATAAATTATTTTACCTAATATGCCTCCAGATCGTTAGCAGTTTCATCTGATTAACAAAAGGAATCTTGAGCATTA
AATAATGAACACGTGTAAAAATTTTATGAAATTTTAATCAAAGTGAATTTTATGAAAAGTTTATGTTATTCATAATGGTTTGAAATATAATAATACGT
CTTTTACTTTATAAATCTTAACAAGAACGGAGAAAAATTATTCAGTAAATATACGATATTATTATTTATATATATTTATTAATGGAGAAGGAGTAAGA
CATTCTAATAAGATTTTTCTTTTAAAATAATAACTCGAAAATGCCTCTCAACCAATTGTAAAACAAAACTCCAAAATCTAACTTGCACGGTAGATTCG
ATCTGGGATAAGAATCTAATTTTAAAATAATTTACAAAATAACTTTTTAATCTTATAAAGCGTGATAATAATTAGTATTGGTTCAAAGTTAATTTCTTT
TTATATAAATAAATTATTTCATATAAAACAATGACAAAATAATCACATAACGTATTTTAATTTCTCTCAGCTCATCATTTGCACTCAATAGATTCTACG
TTTACGATATCCTATCTGATTACTCATAATTCTTATAGTTATTTTTTCTAAAATGCAAATTATAAAAATTATTTTATGATAAATTATTTATCATATCAATA
TGAAATAAAATCATAATGTATAACACTAATTATATAATTTTTTAAATTTTTAATTTTTTATTTAAAATATAAAATTAATGTATTCTAATTTTAAAAATTG
TCCAAATTAACTTTCGGAAAGATACTATTATATAATTTTTTGTGGCAGACAAATCATTTATTGTTTCACCAAAATAATGCACTTGGAAAGCGGGCCAA
AGGGGTCACGTGGTTGTCAGCGGACGATAAGTTACTTTGTCTTTTTCCCTATTTTTTATTTTCCAATATATTAAAAATGAAGTTTTATATTTAGTTTTCT
AATTTAACTTATTAAGATAAAAATATCTTTTTAATATTAAGTAGCATCTATAATTATTTTGAATATTTTAAAACTAAACCTCACTTAATGAGTAATTATT
ATGGTAACAGTATTTTATCATCACGATTTGATCCTAATTGATGCAAAACCATTTTTTCATGTTTATTTTAATTGTTTCTGTAATAATATGTTACTTATAA
GAATAATAATACCCGTTATAATGTTAAATCATCCGTAAAATTATCTCTTTATACTACACATATAAAACTATAGTGAATAATATTTTATGAAAAATATAT
TATATATAAGTAAAATAAATAATACTAAGTGTTTTAAATATGAGTATCTAAAAAAATAATATTTAAATAATATTCGTCAATTATATATAGGATTGCTT
CTTTAGTGTTTAGTTTAATTAACAAAAATATTGCTTCTTTGTTATATCAAAATTATTAAAACTAATTCTTTGTCATACTTTTCTTCAAAAATAAAATCTA
AATAGTACCTCACTTTAATAATCATGTAAATTTCAAATGATTATTATTAGTATAATGCCATTTAACGATATATAAAGAATAATTAACGACTCTTCAAAA
GTGATACTTTTGTACTCTTTCTAACAATTTTAGTTAAGTAGTGGGGGCGTTTGATTGGTAGGCTAAATGGTGGGCTAAAGATATAAATCACACTTAC
AAAATGATTTTGGATTTTTTTATATTCATTTCATCATTTAAATATATATAATTCATATAAGACGGAGAAATATTGAAAATATTCATAAATTTAGCAAGT
TTTTAAATTATGTTATAAGATTCAAGATGTACTATTTAAATTATAATTAATAAAGTAAAAATATTATTTTTAATAATTTTTACATATAATTTAAAAGTGT
ATTTAATACTTAGCTTTTTACGATTTTATCGATGTCCTCTGCCTACCAACCAAATAACATGTTTTTCCGTAAATACCCCTATTGTTGTAAGTTCTCTA
TAAAGCACACTATCATCTTTAACTAAACCATCTCTCATTTTGGGGAGTAAACTCTGCAATCCAATTATTAGGTACTCATTTTTCATCTCTTCTCAATTTT
TTCTCACTATGATTATTATTTTGTTAAATTTAAGCTAATTTTATAGCTATTTATGCATGAAAAATCATAATTAGTGTTAAGAATTCGGTAGTTAATTAT
GCTTTTCTTTTTGTACAATGCATGACATAGTTAGAGTAGAACAAATTATTTTGTTGCTCATGAGTCATGTTTCACGCATGCTGAAGTTTTTGGGTTTAA
TTTTTTTTTTGAAGTTAGTTTTATAAACCAATGTGTTGAAAATAATTTTCTGTAAAGAACAAATTATATTGTTGCGGAATCAGTTGATAATTTTGAAGT
GAAAGAGCATAATCTTATTTTTTGGTAATGCTCCAATAGAAGCATATACTAGTAATTTTGTTAGTTTTGCATATCAAAGAAAAATGTTATTAGTTGAA
AAATTATGTTGTATATATAGTCAAATTATAAATATTTAAAGCTTGGATATTTTTGACATAAGTTAGGTTTGAACTTGCTCTGTTTTCGTTTACTACTGT
TCTTCCCTAAATTTGAATAAATTATTGAGAAAGCGGAATGGATTATTTCTCTTGAGTTTTTTTTTTTTCTGAATATTTCTATGTTTTTTTTTGTGAATAT
GGAGGTAAATTTGTCAAAAAGTTTAGATGATTTTTGGAAGACAACTTAGATTAGTGATGTTTAACGATTAATTAGTATAAGTTGAGTGACCATTCAG
ATATTCAACTCCATATTTAGTCTATTTTACCTTTACATCAGTTGATTTTACAATGATAATTAGTAACAAGAAAATGGCCATGGTGCATAAATTTCAACT
TTTAGTTTCAAGATTGTACTGAACTTGGCGTTTTTCGCGGGATGCTGTTGTTGTTAATCTGACTGTCAATAGCATTATGTTTTTGGGAGAGGTTTCAT
ATTGATGCTATTTGTGCCTTAAAGTTTGTTAA

cDNA

GTAGTGCTCAATTGGTATGTGATTGATGATAAGTAGTACATTTCTTGGCTGTGGTTCTTAACTTGGATTCGTCGTTGTAACAGTTGTTTGTCCTCAAG
TTTTCAAGATGTCGGATGCTAAGGCCCCCGTGGCCACTGCTTACCCTAAGAGGAAGATCCAGTTGGTGGACTTCCTTCTTTCGTTCCGATGGATCAT
TGTCATCTTTTTTGTCCTTCCATTCTCGTTCTTGTATTACTTCTCGATATATCTAGGGGATGTGAAATCCGAGAGGAAATCTTACAAGCAACGCCAGAT
GGAACACGATGAGAATGTTAAAGAGGTTGTGAAGCGTCTTGGCCAGAGGAATGCAGAAAAGGATGGTCTTGTATGCACAGCCAGACCTCCTTGG
GTGGTTGTTGGAATGAGAAATGTCGACTATAAACGTGCTCGTCATTTTGAAGTTGATCTTTCAAAGTTTAGAAATATACTTGATATTGACACGGAGC
GGATGGTTGCTAAAGTTGAGCCTCTAGTCAATATGGGCCAAATGTCAAGGGTCACTATCCCAATGAATCTTTCCCTTGCAGTTCTCGCTGAGCTCGA
TGATCTTACCGTTGGTGGTTTGATCAATGGGTTCGGGGTTGAAGGAAGTTCTCACATATTTGGGTTGTTCTCTGACACTGTTGTAGCACTTGAGGTT
GTTCTAGCTGATGGAAAGGTTGTTAGAGCTACAAAGGACAACGAATATTCTGATCTTTTCTACGCTATTCCGTGGTCTCAAGGAACATTGGGGCTTC
TTGTTTCAGCTGAAATCAAGCTTATACCAGTTGATCAATACGTGAAACTTACCTACAAACCTGTAAGGGGTAATCTTAAAGAGCTTGCTCAGGCTTA
CGCGGATTCTTTTGCACCTAAAGATGGAGATCAGGACAATCCTTCTAAAGTTCCTGAGATGGTAGAAGGCATGATTTATGGTCCAACCGAAGGGGT
TATGATGACCGGTATGTATGCTTCGAGGAATGAAGCCAAACGAAGGGGTAATGTAATCAACAATTACGGTTGGTGGTTCAAACCATGGTTTTACCA
ACACGCTCAAACCGCACTAAAAAGAGGGGAATTTGTTGAGTACATTCCAACTAGGGACTACTACCACAGACACACGAGATCGTTGTATTGGGAAG
GTAAACTAATTCTTCCATTCGGTGATCAGTTCTGGTTTAGGTTCCTCTTAGGATGGCTCATGCCACCAAAGATTGCTCTGCTCAAAGCCACACAAAGT
GAGGCTATTAGAAACTATTACCATGACCATCATGTCATTCAAGATCTCCTTGTTCCTCTTTACAAGGTCGGTGATTGTCTCGAGTGGGTTCATCGCGA
GATGGAGGTATATCCAATTTGGCTGTGCCCACACAGAATTTACAAGCTGCCAGTGAGACCAATGATCTACCCTGAACCAGGATTCGAGAAACACAA
AAGGCAGGGTGACACCGAATATGCACAAATGTATACTGATGTTGGTGTGTACTATGTTCCGGGAGCAGTCCTGAGGGGTGAGCCGTTTGATGGTT
CAGAGAAATGCCGACAACTTGAGCTTTGGTTGATAGAGAACCATGGATTTCAGGCTCAATACGCGGTGACTGAGCTGACAGAGAAGAATTTCTGG
AGGATGTTTGATAACGGTCTGTACGAGCAATGCAGAAGGAAGTACAAAGCTATCGGAACGTTCATGAGTGTGTACTATAAATCGAAGAAAGGAAG
GAAGACAGAGAAGGAAGTGCAGGAAGCTGAGCAAGAGAAAGCTGAACAAGAGACTCCTGAATTGGTTTCTAGTATGAATTTTCGCCTTTTAAATA
AGCAGCTACATTGAACAAGTATTTTACTAGTAGTAAGGTCTGCTTGCCTGTTCCAAATTCAAAAAATGTTTCAAATAATAATGGCTAGTGCAGTTCA
AAACTCTCAAGTTTCATTTCT

E3 Ubiquitin ligase (Solyc10g008410)

Promoter

>SL2.40ch10:2535790..2538789
AAAGAATTACTATGTTCGTTTTTATTTGTTTGTCCTATTTTAACTTGATTCTAAATTTAAGAAAGTAAAATAATTTTTGAATTTTGTGATCTTCAATTAA
AAAAAATGTGAAATGTATCAAATTGTCAATTAAGTTTTTAATTTTAAACTTGTCGGGTTAAAAGTTAGAATTAACAATTTGTTAAAAAAGAAACGTTA
AAGGAAGAATAAGATGATTCAGATATGCAAAGAAGAGGTAGTGAATACGCCAATTAGGAGGTTGGCCAAAAAAACATTAATCGCTTTCTTGTTTGT
AGCTACTTAGTAATTTTAAAAATGATAGAATTATCAACAAAGAACACAGATTTTCTAGAACACAGTTTTGAAGTTGCAATTTAATTACAGATCTTAGT
AATTGTAGTATTTATTAGACAAAAATACAAAAAAGTAATTACAAGTGTAAAGTACAATACTTAGTAGTTAATTATTAGAGAAAGTCAGAGTATGTCT
CTTAACGATTAGGAAACTGCTGAGTAGCCCCAAAAAACTAAAAGGCGGTGAGAAAACCCTAAATTAATATTGAAGAAAGGAAAAAAAATTAGATG
TATCCCGAATTATCGTAAATGGTATGCAGATATTCTTCGTCATACTTTGGAGGATATTGGTGCCCTAGCTGTCCAGAAATTACAGCATACATACCCTT
TATACTAACGGACATACACATGTCATAATTTTATCCACCGATATGACATTTATCGACGGATAAAACTGCACCACGTGAGCATATTTAGTCTTTCGTTA
GAGTAAAGGGCATACATGTTCTAGTTTATGGATGACAGAGATACTAATATCCCGAAAATATGACGAAAGATGCTTACATACCATTTATGATAGTTTG
ATGTATATTTATATTTTTCCCCTTAAAGAAATTGGGTGATAAAGACTAATAACTTTGACTTTAATACTTAATTACATACTAATTGTCAGGTTGGATAAG
GTCTTAGCTGAAAAATTAAGGTAAGACTTCAATTACTGGACACCTGACATCATGGGCGGATTAAAATATAACAACGGATTTGAGCTTTGAAAATAA
TCTCTCGTAGTGTTCAATCCAACCTACGTCTTCATGTTCATAAACATGATTAGTGTCGAATGACATTATCGGAATAATATCTAAGATGAATGAAGGA
GGATCCTCCATAAACATATCAAAATTTTGATTTTCTTTTTTTAAAAAAATTATGAGCAAAGTAATGAAATAGAAATAATAATACCAAGAAGGAATGT
GAAATTCTTTGATCGGAGAGATAAATAAAATTCAAATTCAAAATTAGAATGATAAGATATGAGAGGAAAAATATGGCGTGATTTCTGAAATATTGG
AGAGAGAAACATGGGTTGGGTTAGAAGATGGGAGGTTATGTATCTGGGAGTTAGGTTTAAGAGGTAAAAAAGAGATTTTAGTAATATTTTAAAAA
AGAAGAGAATAATAGAAAATAAATAAATTAATGATTTGTATATAAGTAATTTTTTCAAAATAAATTCTTGTGATTCGATTTTTAATAATTGAATAAGT
CTCCTTGTATATATCCAATTATTTGAACAATTCATTTATTTAAAATTTAAATTTAATATTTATATTTTTTTAAATAAAAAGTTTTGTCTATTCAAAAAGG
AAGCAACAAAAGTTTTGTTGAGTAACTGAAGGTTGCATAGGTAAATTATATTGAAAGACATATTATATGATTTGACTGATTAACCTACATATTCAAA
AATTAAAAAAAAACTTAAAAATTATTGAGAAGATTTCCTCATAAACAAACTCTTTAAAAACTACATTATGGGTGCGTTGTTTGGATTGAGAAAATAA
GTCTTTAATTTATATTAGTTCTAATTTTTCCTCCAACAAAGGATTAACTCAAATTGGAAGAAAATTACTAAAAACATTTATGATATACTTTGACTTTCTT
TTAAGAGAAAAATAAAAATCAAATAAAGTAAGAAAATTAGGGCTAAACGCTAACAGATCTTTCATGCCTCAAAATTTTAATTTCTTCTAATATAAATC
ATTTCATTAAAGATAAAATGAATATTTTTTAAAATTAAATTATCAGTAAACATATAAATATATTGTTAAAATATATTATTTATTTAAATTGACTAAAAA
GAATATATCACATAAATTAAGACAGTTAAGTTGACGTATAACACCGGCTTTAGTCATATTTAACACATGACCTAATGGCTTGTATATACTTACCAAAG
TGTCATACTCAATAGCACACCATTCACTTTATGACTATTTTTAGCCCCTCCATCTTTTTGGGACTTTATATTAATCAACATATTGATTAGAAATGTTCTT
TTTTAGTGGGATATTCTGATGCTTGTATCCAATTGAGGTATTTTCCTTTTTCCTTCTTTGTGTCTATATTTCTAATTCACTTATATCGAAAAAGATTTTG
ATCTTTTTGTAATTCAATTGGTATTTTGAAACGTTTCTAACTATATTGTTTTAATTTTTTAAAAATGTAGATAGATGCATGTCGAATTTCTCAAAAGTA
GCGTATCTTTGAGGGATGCAACATGAGTATATAATATTTTTTTAAAGTCCGATTAACATAGTGACTGAAGTATCGTATTTTAGGTTTTTGTTTTCTT
GGAAATTGGAGGGATGGTTTTTGTTGTTGTTGTTGTTTCTAATGACTAATTTGTTTTAGGTCATGATTTTCGCCAAAACGTGTCACACTATTAAGAAT
ATGGTCCTTTAACTTGGCTTCAATTAGTCTTTTTATGCTCCAACTTTCTGTGTCTCAAGTAAACACTTAAATTTGCAGTAGTTTTTCTTATAGTCGGTGT
TTTATGATATAAGTTAAGTTGTAAGAAGCGCGTTTTAGAAGAGTTGGAGGAAACATATAAACTTATTTCAACATATTTCTAACAGGGCATTAGCAAT
CAATGGCCCTGGATTTCATTTGCAACAAGCAAATAGAAGAAGGTAAGCCCATAGAGAAGTTGAAA

cDNA

ATGGTTGATGATGAACATGAAGGCGGCAATCTCTCTCGTGGTTTTGAGTGTAACATTTGCCTAGATCTTGCACATGATCCAGTGGTTACCTACTGTG
GTCACTTGTATTGTTGGCCTTGTATCTACAAATGGATTCACTTACATAGCATCCCTTGTGAAAATCCAGCTCAACGCCACCCACAATGCCCGGTTTGC
AAGGCTCATGTTTCACAAACAACCATGGTTCCACTCTATGGACGCGACCAAGCTACAAAAACATCTCAAGATGATGGCATGGCCATACCAGAAAGA
CCTCAACAATTCCATCATCGGTCTAGTACAAGCACGATTGGTGAAACAGCTCATGCAGAACGCGTTGATGGGAACTCATTTGTTCCATATCCAAACG
CGCATCATCTAGCAGGGACCGCTAGTTTGAGGATGAGAAGGCAACAACTGCAAGCTGATGATAAATCACTTCGCAGAGTGCATTTTTTTCTCTTCTG
TTGTGTGGTGTTGTGTCTTATCTTGTTATGA

Δ(7)-STEROL-C5(6)-DESATURASE (Solyc02g086180)

Promoter

>SL2.50ch02:48923013..48926012
AACTCTCACAAAATTATTAACTGTTATTCGAATTGATGACTCTAGAAATCTCCAAAGACTCAATAAATTTTATCAGCGATAAATCTTAGTATAGCTCTA
TGAATGTAGGTCACTAGTGCAACAACTGGGAGATCCAAAACTCTCATATAACAAATCCTCTATAATCATTTTGCTCTATTCCGGAGAATTTCATTCCA
TTTCCTAAACTCAAAGTATCATAATTACTGCTACTTCATTAGATAGAGAAATTTTGTTTACCTTTATTTATGAATGAAAAACAAGAAGAAAACTAGGT
TCTATAAACGTATGAAATAAGGCGAGAAAGAGATTACTTGAAATTTGGAAACTTGAGCATTGTTGACACCATTTTTATTGGTGAAATTGGGTTGTAG
AGATGGTACTGATGCTATCGCAGCTTCTCTCCTCCCGTCGTCCTCTTCCATTGAAACAGCTCCAGACCCGACCCTGACCCGCTTCGCTCTGTGGCAAT
TTATTCGGGTCTTTGGGTCTACTGACGGGAATGGGCCAGGATACAACATAACACGGGCCTTGTAATTTCATGATCCACTCATTTGTCATACAGTGAA
CATTTTTTGTATAAGTAAATAGATTTTTATTATAATCGATATATCTATGAATGAATAGATTAGGTTAAATTTAAATAGATCAAGATAAATAGAGTTAA
TAAGTTACTGTTAAACAAAATTAACCATAAAAAAAATATCGCATATCTAACGTTGCGTATGACATTTCATGTGAAGATTATTTATTTCTATACAAATTTC
TATTTATATATCGTTTATACAAAAAATATATAATTTTTAATATTTGTTATTTCATAAAATTAACGTATAAATAACTATATTTTTCTTATTTTATTTTTCAG
CATTATTATACTTAAAAATATGAATTTAATTGACATAGAAAAGACAAACAACATGATACGTGGTACTCCCTTACTTTGTCTCCTATTTATATTTATTAA
TTATTAGGACAACATGAGTATTTAATTAAAAGAAAAACCTGCTCCAACTCCCTTCACATGCACACAAAATCCAAAGGGATACATTATATAGGTTTA
TTAAAATGTTTTTTAAGTATTTATTATTATAATTAAGTAAATTATATAAAAATATGTTATTTCTTTAACTATACATATTAATTCTGATTGGAACAAGCTT
GATAATTTAATACTGAGATTTTATACATATAATTAAAAGATGTTATGATAAATTGAACTACTTAATAAACATGTAAAAACCCATATATTTTCAAAATA
TTAAAGTTAAAATATCTAATAAACTTTTTTTAATCATATAGTTAAAATATTCAATTGAAATAGATTGAATTTAAGTGTCTAAATAAATTTAGTAACAAA
CTAAACAAACGAAGGGAAAAAAAAGATAGAAAGGAAAAGATTAAAGAGCATTATTTCTGTTATGCTTGATTTTAATTTAATTCACCACTAAGATTTT
TTATTTTTTAGAAAAAAATTACAAACCATTAGATCACGTTTACTTTTTAGGCTATTGTTAAGTTGTGGAGCTTGATTAATATTTCTGTAACTAAATATA
CTCTGAATTATTAATATATACTTCACCGCCGTGGAAGTTTATTCACGGATATTACCACGAACTATTAATTTCTCTCTCTAGATCTCTCATCTCTCTCTCA
AAAATTCTTATATTTTTCATTTATCAAGTGTGTGTGAATTCGATCCTAACAACTATACGAGTAGAACTACCTAACATTTCAAGTAGCGGCATGAAATG
TCTATTTAGGCTAAAAAACAAAGGACCACGAGGAAAATAACAACGGAGAGAGTGATTTTATAGCCCCTCAGCCAAATAAATAAATAAATAAATAAA
TTCTCTTTGGCCGGCCAGCTAGTTTCAATCTTCGCAAAATAATAGATATTTCTTGATAAGAAGGTTTAGATCCATAGCATAGGTAACATGAGAAATAT
GAAAATTAATATCATATAATATATTTAATTTAATTTTGTAAATTGAATCTACAAAATAATGATGTATGCACTATTTAATATTTTGTGAATGTTGAAATA
TTGTTTCATTAGACTGATATGATAGATTGCTAAACTAAATCTCATAATATATATTACATTAAAGCTCAAACTATAATCTCTTTGATACTATTTCCATAG
AACGAATGTGCATTAACAGAGGGGAAAGGAGAAAGACCACATAATGATTAATCATAGAGATGATAATAAAACCAACAAAAAAAAAAAAAAAAAGA
CAAAGTAGAACAAGCCAAACTTATGAAGGCCATAATAACCATATAATAAAACAGACAAAATTTTAATTTTTGAATTTTTTTAATGTAGTCGAGTTTCC
CCCCTTCAAAATTAATTATTGTGACATCAATATACTGCAACAAGTGGCTGGTTAGTTCAACACAAGTTTAAAGGAATAACACAAACAAACTTTCATCT
AGTTATATCGTTTTAGTAAAATGGACATAAGTTATTGCTTAAATTCCATAACTGTTTATATTTTCTTAAAAGTTTGATACATTTGTTTATGAAAATCAC
TTAATAGCTTAATGTTAGATGTATACATATCACAATTTTAGCTAATATCAAAGTTAATTTTGTGTCTAACTTTTTTATATATATCTACTCGATTTTCAAT
AGTATCTAGATTAATGAAACATGACAAGGTGAATTCAAACTAAGGATTAAAATTAATAAAGCATGATAAGATTATATTAAATATATTATTGTTATCA
GTAATATTTTCTCGATTTATTATTACATAATTATTAATAAATGTATGAAATATTACTCTAAAAAAAAAAGTGATTGAATGATAAAAATAAAATGTGAAA
TCTATCATATTTTTAAAAATATTAAACAAATAAAATCAAACGAACGATCTTATTAGAATCTTTTGCCAGTTTAGCTCACCACACAACACGTGGGAGTA
GTAGTCCAAAATGGAATCAATTTTCATTAAATACAGCCTGCCACACATGCATATA

cDNA

TTGTGTTACCGCACTAGCCCCTCTTTGTTTTCCTCTCTATCTTCACTCCATAGTTCCTCCTTTTCCGGCGGACCATCGCCGCCGGAGTCTGGTAAAACC
CGACCCGAGACGGACATATTTCTCTTTTACCTAGTATACCGTCGGAGATGGAGGATTACTTGAAGCTATTCGTGGAGGAGACATCGTTTTACAACCG
TCTGGTTTTGGGTACATTCTTGCCGGAATCATGGTGGGGACCACTTCCTCATATGCTTCAAGGATGGCTCCGTAACTACATTGGCGGTGTTTTACTTT
ACTTCATCTCCGGTTTCCTCTGGTGCTTCTATATTTATCACTTGAAACGCAATGTCTATATTCCCAAAGATGCCATACCATCAAAGCAAGCAATGCTGT
TGCAAATATCAGTTGCTATGAAAGCTATGCCGTGGTACTGTGCCCTTCCATCACTTTCTGAGTACATGATTGAAAACGGATGGACCAAATGTTTTGC
GAGAATAAGTGATGTTGGATGGCCTACCTACATCATCAATGCGGCTATTTATCTTGTAATAGTGGAGTTTGGAATCTACTGGATGCACAAGTTATTG
CATGACATAAAACCTCTGTATAAATATCTGCATGCTACACATCATATTTACAACAAGCAAAACACACTTTCCCCATTTGCTGGATTGGCATTCCACCC
ATTGGATGGAATACTGCAGGCAGTGCCACACGTTGTAGCTCTTTTCTTGGTGCCAATGCATTTCACTACACACATAGTGCTCATATTCGTGGAAGCC
TTATGGACGGCTAATATTCATGACTGCATACATGGTAAGGTGTGGCCTGTAATGGGTGCCGGTTATCATACCATTCACCATACAACATACCGCCATA
ATTATGGTCATTACACAATATGGATGGACTGGATGTTTGGAACTCTTCGTGATCCCGTTGAAGAGGATGCCAAGAAAATGTAATGTGCATGATTTG
CACTCAAGGATGTTCCTTATTGTTCATTGTCCCAGCAATTTATCTCGTGTTTCCATCGTTGGTGGATGTAAATTTTGATGTAGTAGTTTATGTGAAAAC
AATTTGCTTTTATTGTCTTTTCTTCCCCTATATATTTAATAATAAGCAATAAAATCTGTGGGGAAGAAAATTTCAGTTTCTAACTTTATGCCGGAGTCC
TTTGAGGTAATCATATAACGAAAGCTAAATGTGAATCACATTGTCACCATGTTATGGTGGATA

Supplemental Figure 3. GC-rich boxes found in SGA-related genes.

Three GC-rich boxes in the promoters of target genes are known to be bound by group IXa ERFs (Shoji et al., 2013). A P-box recognized by all ERFs, except clade 1 AtERF1; a CS1-box bound only by clade 2-3 ERFs (e.g. ERF163, ORCA3, AtERF13) and a GCC-box bound by all, except clade 2-1 ERFs (e.g. ERF189, ERF115). For each gene is highlighted the presence of a P-box (green) in the promoter region (~3000 bp) and the start and stop codons (red) in the coding sequence.

Supplemental Table 2. Details of homologs co-expressed with *GAME9* in both tomato and potato presented in Figure 1.

Annotation	Solyc ID (tomato)	Sotub ID (potato)
GAME12 (Aminotransferase-like protein)	Solyc12g006470 [a]	Sotub12g011080 [a]
GAME4 (CYP88D)	Solyc12g006460 [a]	Sotub12g011090 [a]
GAME6 (CYP72A)	Solyc07g043460 [b]	Sotub07g016580 [b]
GAME1 (UDP-galactosyltransferase)	Solyc07g043490 [b]	Sotub07g016600 [b]
GAME11 (2-oxoglutarate-dependent dioxygenase)	Solyc07g043420 [b]	Sotub07g016570 [b]
Cellulose synthase	Solyc07g043390 [b]	Sotub07g016530 [b]
SSR2 (Sterol side chain reductase 2)	Solyc02g069490	Sotub02g015720
Phylloplanin	Solyc01g066680	Sotub01g020750
DNA topoisomerase 2	Solyc01g087500	Sotub01g025950
F-box family protein	Solyc01g104230	Sotub01g041250
Unknown protein	Solyc01g106540	Sotub01g043720
Subtilisin-like serine protease	Solyc02g069630	Sotub02g015850
Ascorbate peroxidase	Solyc02g083630	Sotub02g026940
Male sterility 5 family protein	Solyc02g083640	Sotub02g026950
Laccase	Solyc02g085110	Sotub02g028260
Laccase	Solyc02g085120	Sotub02g028270
Genomic DNA chromosome 5 P1 clone MSI17	Solyc02g086140	Sotub02g029290
Microtubule-associated protein MAP65-1a	Solyc03g007130	Sotub11g029000
Leucine-rich repeat receptor-like protein kinase	Solyc03g083510	Sotub03g020030
E2F transcription factor-like protein	Solyc03g113760	Sotub03g027120
Tubby-like protein 13	Solyc03g117730	Sotub03g031190
Kinesin-like protein	Solyc04g076310	Sotub04g029340
Cyclin A-like protein	Solyc04g078310	Sotub04g031250
Pollen allergen Phl p 11	Solyc05g051870	Sotub05g024500
Glucosyltransferase	Solyc05g053400	Sotub05g026070
Os03g0291800 protein	Solyc06g050920	Sotub06g011200
Cyclin-dependent protein kinase regulator-like protein	Solyc06g065680	Sotub06g024730
Genomic DNA chromosome 5 TAC clone K19B1	Solyc06g070970	Sotub06g026960
Leucine-rich repeat receptor-like protein kinase PEPR2	Solyc06g084420	Sotub06g034920
Dopamine beta-monooxygenase	Solyc07g048050	Sotub07g018160
LRR receptor-like serine/threonine-protein kinase	Solyc08g061560	Sotub08g014090
Pectinesterase	Solyc08g078640	Sotub08g024430
Os04g0690100 protein	Solyc08g079010	Sotub08g024870
Alpha-hydroxynitrile lyase	Solyc09g014970	Sotub09g012700
Cyclin A1	Solyc11g005090	Sotub11g013250
Superoxide dismutase	Solyc11g066390	Sotub11g024220
Genomic DNA chromosome 5 P1 clone MUP24	Solyc11g069420	Sotub11g026580

[a,b]Genes localized in clusters on chromosome 12 and 7, respectively, previously associated with SGA biosynthesis (Itkin et al., 2013).

Supplemental Table 3. Changes in expression of genes involved in the biosynthesis of SGAs and sterol precursors in potato *GAME9*-Ox leaves determined by RNA-Seq analyses.

Process	Gene	Id	Description	Change
Regulation	*GAME9*	Sotub01g029510	Ethylene responsive transcription factor	↑
SGA biosynthesis (Chr 7)	*GAME2*	Sotub07g016550	UDP-rhamnosyltransferase	↑
	GAME11	Sotub07g016570	2-oxoglutarate-dependent dioxygenase	↑
	GAME6	Sotub07g016580	Cytochrome P450 72A188	↑
	GAME1	Sotub07g016600	UDP-galactosyltransferase	↑
SGA biosynthesis (Chr 12)	*GAME12*	Sotub12g011080	Aminotransferase-like protein	↑
	GAME4	Sotub12g011090	Cytochrome P450 88D	↑
Biosynthesis of precursors	*HMGR*[*1]	Sotub02g025630	3-Hydroxy-3-methylglutaryl CoA reductase 1	No change
	SQS[*2]	Sotub01g047720	Squalene synthase	No change
	CAS1	Sotub04g023080	Cycloartenol synthase	↑
	TTS1[*5]	Sotub12g010880	*β*-amyrin synthase	No change
	SMT1[*3]	Sotub01g025880	Sterol C24-methyltransferase type1	No change
	SSR2[*4]	Sotub02g015720	Sterol side chain reductase 2	↑

*Identified in potato (Ginzberg et al., 2012). Accession numbers: [1]L01400, [2]AB022599, [3]DN908821. [4]Sawai et al., 2014. *[5]Identified in tomato (Wang et al., 2011).

Genes are grouped according to their association with SGA biosynthesis or synthesis of precursors. Change is indicated when increase (↑) in gene expression was found as compared to wild-type potato plants.

Supplemental Table 4. Changes in expression of genes involved in the biosynthesis of SGAs and precursors in tomato *GAME9*-RNAi and *GAME9*-Ox leaves as determined by RNA-Seq analyses.

Process	Gene	Id	Description	RNAi	Ox
Regulation	*GAME9*	Solyc01g090340	Ethylene responsive transcription factor	↓	↑
SGA biosynthesis (Chr 7)	*GAME2*	Solyc07g043410	UDP-xylosyltransferase	No change	No change
	GAME11	Solyc07g043420	2-oxoglutarate-dependent dioxygenase	↓	↑
	GAME6	Solyc07g043460	Cytochrome P450 72A188	↓	↑
	GAME17	Solyc07g043480	UDP-glucosyltransferase	↓	↑
	GAME1	Solyc07g043490	UDP-galactosyltransferase	↓	↑
	GAME18	Solyc07g043500	UDP-glucosyltransferase	↓	↑
SGA biosynthesis (Chr 12)	*GAME12*	Solyc12g006470	Aminotransferase-like protein	↓	↑
	GAME4	Solyc12g006460	Cytochrome P450 88D	↓	No change
Biosynthesis of precursors	*HMGR*[*1]	Solyc02g082260	3-Hydroxy-3-methylglutaryl CoA reductase 1	No change	No change
	SQS[*2]	Solyc01g110290	Squalene synthase	No change	No change
	CAS1	Solyc04g070980	Cycloartenol/Lanosterol synthase	No change	No change
	TTS1	Solyc12g006530	*β*-amyrin synthase	No change	No change
	SMT1[*3]	Solyc01g087560	Sterol C24-methyltransferase type1	No change	No change
	SSR2[*4]	Solyc02g069490	Sterol side chain reductase 2	↓	↑

*Identified in potato (Ginzberg et al., 2012). Accession numbers: [1]L01400, [2]AB022599, [3]DN908821. [4]Sawai et al., 2014. *[5]Identified in tomato (Wang et al., 2011).

Genes are grouped according to their association with SGA biosynthesis or synthesis of precursors. Change is indicated when increase (↑) or decrease (↓) in gene expression was found in comparison to wild-type tomato plants.

Supplemental Table 6. Details of promoters tested in tobacco protoplast transfection assays for transactivation by GAME9.

Gene Symbol	Annotation	Solyc ID	Size cloned
GAME1	UDP-galactosyltransferase	Solyc07g043490	~2600 bp
GAME2	UDP-xylosyltransferase	Solyc07g043410	~2300 bp
GAME4	CYP88D	Solyc12g006460	~2600 bp
GAME6	CYP72A	Solyc07g043460	~2600 bp
GAME7	CYP72A	Solyc07g062520	~2700 bp
GAME9	Ethylene responsive transcription factor	Solyc01g090340	~1650 bp
GAME11	2-oxoglutarate-dependent dioxygenase	Solyc07g043420	~2000 bp
GAME12	Aminotransferase-like protein	Solyc12g006470	~1250 bp
GAME17	UDP-glucosyltransferase	Solyc07g043480	~1900 bp
GAME18	UDP-glucosyltransferase	Soly07g043500	~1550 bp
HMGR1	HMG CoA reductase 1	Solyc03g032010	~1500 bp
HMGR2	HMG CoA reductase 2	Solyc03g032020	~1500 bp
SQMO	Squalene monooxygenase	Solyc04g077440	~1200 bp
SQS	Squalene synthase	Solyc01g110290	~1300 bp
SMO1	Methylsterol monooxygenase 2-2-like	Solyc06g005750	~2250 bp
SSR2	Sterol Side Chain Reductase 2	Solyc02g069490	~1600 bp
C-5 SD	Δ(7)-sterol-C5(6)-desaturase	Solyc02g086180	~1600 bp
Δ-14 SR	Δ(14)-sterol reductase	Solyc09g009040	~2100 bp

For each gene, a region upstream of the start site (ATG codon) spanning ~1200-2700 bp was cloned and used for transactivation with GAME9 in tobacco protoplasts.

Supplemental Table 7. Accession numbers for ERF proteins used in the phylogenetic analysis presented in Figure 2.

Name in Figure 2	Organism	Accession number
Sl_GAME9-like1	*Solanum lycopersicum*	Solyc01g090300
Sl_GAME9-like2	*Solanum lycopersicum*	Solyc01g090310
Sl_GAME9-like3	*Solanum lycopersicum*	Solyc01g090320
Sl_GAME9	*Solanum lycopersicum*	Solyc01g090340
Sl_GAME9-like4	*Solanum lycopersicum*	Solyc01g090370
St_GAME9-like1	*Solanum tuberosum*	PGSC0003DMG400026049
St_GAME9-like2	*Solanum tuberosum*	PGSC0003DMG400041045
St_GAME9-like3	*Solanum tuberosum*	PGSC0003DMG400025991
St_GAME9-like4	*Solanum tuberosum*	PGSC0003DMG400046672
St_GAME9	*Solanum tuberosum*	PGSC0003DMG400025989
St_GAME9-like5	*Solanum tuberosum*	PGSC0003DMG400026048
St_GAME9-like6	*Solanum tuberosum*	PGSC0003DMG400026046
St_GAME9-like7	*Solanum tuberosum*	PGSC0003DMG400040573
Nt_ERF189	*Nicotiana tabacum*	ERF189
Nt_ERF115	*Nicotiana tabacum*	ERF115
Nt_ERF179	*Nicotiana tabacum*	ERF179
Nt_ERF168	*Nicotiana tabacum*	ERF168
Nt_ERF221/ORC1	*Nicotiana tabacum*	ERF221
Nt_ERF104	*Nicotiana tabacum*	ERF104
Nt_ERF17	*Nicotiana tabacum*	ERF17
Cr_ORCA1	*Catharanthus roseus*	AJ238739
Cr_ORCA2	*Catharanthus roseus*	AJ238740
Cr_ORCA3	*Catharanthus roseus*	EU072424
Sl_SHN3	*Solanum lycopersicum*	XP_004240977.1
Sl_SHN1	*Solanum lycopersicum*	XP_004235965.1
At_SHN1	*Arabidopsis thaliana*	At1g15360
At_SHN2	*Arabidopsis thaliana*	At5g11190
At_SHN3	*Arabidopsis thaliana*	At5g25390
Sl_ERF1	*Solanum lycopersicum*	AAL75809
Os_SHN1	*Oryza sativa*	BAD15859
At_ERF1	*Arabidopsis thaliana*	At4g17500
At_DREB1A	*Arabidopsis thaliana*	At4g25480
At_DREB1B	*Arabidopsis thaliana*	At4g25490
At_DREB1C	*Arabidopsis thaliana*	At4g25470
At_DREB2A	*Arabidopsis thaliana*	At5g05410
At_DREB2B	*Arabidopsis thaliana*	At3g11020
Sl_AP2a	*Solanum lycopersicum*	ACD62792
Sl_AP2b	*Solanum lycopersicum*	HQ586952
Sl_AP2c	*Solanum lycopersicum*	HQ586951
Sl_AP2d	*Solanum lycopersicum*	HQ586953
Sl_AP2e	*Solanum lycopersicum*	HQ586954
At_ANT	*Arabidopsis thaliana*	Q38914

At_TOE2	*Arabidopsis thaliana*	Q9LVG2
At_ WRI1	*Arabidopsis thaliana*	AY254038
At_AP2	*Arabidopsis thaliana*	At4g369200
Bn_BBM1	*Brassica napus*	AF317904
Bn_BBM2	*Brassica napus*	AF317905
At_PLT1	*Arabidopsis thaliana*	At3g20840

Sequence data can be found in GenBank/EMBL data libraries, Sol Genomics Network or in the tobacco transcription factors (TOBFAC) database, respectively.

Supplemental Table 8. Oligonucleotides used in this study.

Organism	Name	Sequence	Use
Potato	NAC-Fw	ATATAGAGCTGGTGATGACT	qRT-PCR
	NAC-Rv	TCCATGATAGCAGAGACTA	qRT-PCR
	GAME9-Fw	AAGCCGCTTACAAGATTCGG	qRT-PCR (also used in tomato)
	GAME9-Rv	ACGACGCCTAGCCGTTACTC	qRT-PCR (also used in tomato)
	CAS-Fw	AATCATGACGGTCACTGGGCT	qRT-PCR
	CAS-Rv	AATACTGCATTGAGTGCCCCC	qRT-PCR
	SMT1-Fw	GCTTTGCTGTGTATGAGTGG	qRT-PCR
	SMT1-Rv	TTTTGTGTCGATCGAATCTC	qRT-PCR
	SSR2-Fw	CCACCGTTCACCCTAGGAGG	qRT-PCR
	SSR2-Rv	ATACAAGAACGAGAATGGAAGGACA	qRT-PCR
	GAME11-Fw	TGGCGGACCTTCTTTCAAACT	qRT-PCR
	GAME11-Rv	CACAATTTCAACTGGATCCGATG	qRT-PCR
	GAME6-Fw	TTTGGCCGTATGTTTGCCTT	qRT-PCR
	GAME6-Rv	TTCCTTGCTGCCGCAGTT	qRT-PCR
	GAME4-Fw	GGGACTCAAGGCTCGAAAAGTACT	qRT-PCR
	GAME4-Rv	TGTTTGCCCTTGGCATTGAT	qRT-PCR
	GAME12-Fw	GGAATGGCCAAGACTACTAATGGA	qRT-PCR
	GAME12-Rv	GCTCTATCTATAACTAAAGGTCCCATA	qRT-PCR
	GAME1-Fw	TGGGTCCCACAGCTTACGATC	qRT-PCR
	GAME1-Rv	GGCACGCCAAAAGTGATGG	qRT-PCR
	GAME2-Fw	GAAACTGCAATGCCGCATGT	qRT-PCR
	GAME2-Rv	TTGAGGCCATGGAGGGC	qRT-PCR
	GAME9-Ox-Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTA TGAATATTGCAATTGATGATGATGA	Cloning GAME9 for Ox
	GAME9-Ox-Rv	GGGGACCACTTTGTACAAGAAAGCTGGGTT CATTTGTATCAACATTTGTAAATTCACAC	Cloning GAME9 for Ox
	GAME9-RNAi-Fw	aaaaaGCGGCCGCATGAGTATTGTAATTGATG ATGATGAAATC	Cloning GAME9 for RNAi (also used in tomato)
	GAME9-RNAi-Rv	aaaaaGGCGCGCCCACACGCCACAGATGGTTC TT	Cloning GAME9 for RNAi (also used in tomato)
Tomato	TIP41-Fw	AGATGAACTGGCAGATAATGG	qRT-PCR
	TIP41-Rv	CATCAACCCTAAGCCAGAAA	qRT-PCR
	HMGR-Fw	CTGACGCGCTTCCACTCC	qRT-PCR
	HMGR-Rv	GATCTTCTCACGCCACCTTACG	qRT-PCR
	SSR2-Fw	GGCCAAATGTCAAGGGTCACT	qRT-PCR
	SSR2-Rv	ACCCCGAACCCATTGATCA	qRT-PCR
	GAME11-Fw	TGGTCCTGAGAATCCTCCACA	qRT-PCR
	GAME11-Rv	GCTCCAATGAAGCGTGGTACAC	qRT-PCR
	GAME4-Fw	CTTCAATGTGTGGTGATCCAAAA	qRT-PCR
	GAME4-Rv	CCATAATTGTTGGCTTCCCAAA	qRT-PCR
	GAME12-Fw	TATGACTGCCGGTCTCTCCG	qRT-PCR
	GAME12-Rv	GATAGTTCCAATAATGAGGGCAATCA	qRT-PCR
	GAME17-Fw	GCTGCAGGATTCCCTATTCCAC	qRT-PCR
	GAME17-Rv	TACTTAGCATGGTGCTCCAC	qRT-PCR
	GAME9-Ox-Fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTA TGAGTATTGTAATTGATGATGATGAAATC	Cloning GAME9 for Ox
	GAME9-Ox-Rv	GGGGACAAGTTTGTACAAAAAAGCAGGCTA TGAGTATTGTAATTGATGATGATGAAATC	Cloning GAME9 for Ox

Parsed Citations

Aerts, R.J., Gisi, D., De Carolis, E., De Luca, V., and Baumann, T.W. (1994). Methyl jasmonate vapor increases the developmentally controlled synthesis of alkaloids in Catharanthus and Cinchona seedlings. Plant J. 5: 635-643.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Arnqvist, L., Dutta, P.C., Jonsson, L., and Sitbon, F. (2003). Reduction of cholesterol and glycoalkaloid levels in transgenic potato plants by overexpression of a type 1 sterol methyltransferase cDNA. Plant Physiol. 131: 1792-1799.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Baldwin, I.T., Schmelz, E.A., and Ohnmeiss, T.E. (1994). Wound-induced changes in root and shoot jasmonic acid pools correlate with induced nicotine synthesis in Nicotiana sylvestris Spegazzini and Comes. J. Chem. Ecol. 20: 2139-2157.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Buckingham, J., Baggaley, K.H., Roberts, A.D., and Szabo, L.F. eds (2010). Dictionary of Alkaloids 2nd ed. (CRC Press).
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Cárdenas, P.D., Sonawane, P.D., Heinig, U., Bocobza, S.E., Burdman, S., and Aharoni, A. (2014). The bitter side of the nightshades: Genomics drives discovery in Solanaceae steroidal alkaloid metabolism. Phytochemistry http://dx.doi.org/10.1016/j.phytochem.2014.12.010
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Choi, D., Bostock, R.M., Avdiushkot, S., and Hildebrandt, D.F. (1994). Lipid-derived signals that discriminate wound- and pathogen-responsive isoprenoid pathways in plants: Methyl jasmonate and the fungal elicitor arachidonic acid induce different 3-hydroxy-3-methylglutaryl-coenzyme A reductase genes and antimicrobial isoprenoids in Solanum tuberosum L. Proc. Natl. Acad. Sci. U. S. A. 91: 2329-2333.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

De Boer, K., Tilleman, S., Pauwels, L., Vanden Bossche, R., De Sutter, V., Vanderhaeghen, R., Hilson, P., Hamill, J.D., and Goossens, A. (2011). APETALA2/ETHYLENE RESPONSE FACTOR and basic helix-loop-helix tobacco transcription factors cooperatively mediate jasmonate-elicited nicotine biosynthesis. Plant J. 66: 1053-1065.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

De Geyter, N., Gholami, A., Goormachtig, S., and Goossens, A. (2012). Transcriptional machineries in jasmonate-elicited plant secondary metabolism. Trends Plant Sci. 17: 349-359.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

De Sutter, V., Vanderhaeghen, R., Tilleman, S., Lammertyn, F., Vanhoutte, I., Karimi, M., Inzé, D., Goossens, A., and Hilson, P. (2005). Exploration of jasmonate signalling via automated and standardized transient expression assays in tobacco cells. Plant J. 44: 1065-1076.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Diener, A.C., Li, H., Zhou, W., Whoriskey, W.J., Nes, W.D., and Fink, G.R. (2000). STEROL METHYLTRANSFERASE 1 controls the level of cholesterol in plants. Plant Cell 12: 853-870.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Eich, E. (2008). Solanaceae and Convolvulaceae: Secondary Metabolites (Springer: Berlin).

Expósito-Rodríguez, M., Borges, A.A., Borges-Pérez, A., and Pérez, J.A. (2008). Selection of internal control genes for quantitative real-time RT-PCR studies during tomato development process. BMC Plant Biol. 8: 131-142.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H., and Ohme-Takagi, M. (2000). Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. Plant Cell 12: 393-404.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Fujiwara, Y., Takaki, A., Uehara, Y., Ikeda, T., Okawa, M., Yamauchi, K., Ono, M., Yoshimitsu, H., and Nohara, T. (2004). Tomato steroidal alkaloid glycosides, esculeosides A and B, from ripe fruits. Tetrahedron 60: 4915-4920.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ginzberg, I., Thippeswamy, M., Fogelman, E., Demirel, U., Mweetwa, A.M., Tokuhisa, J., and Veilleux, R.E. (2012). Induction of potato steroidal glycoalkaloid biosynthetic pathway by overexpression of cDNA encoding primary metabolism HMG-CoA reductase and squalene synthase. Planta 235: 1341-1353.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hatlestad, G.J., Akhavan, N. a, Sunnadeniya, R.M., Elam, L., Cargile, S., Hembd, A., Gonzalez, A., McGrath, J.M., and Lloyd, A.M. (2015). The beet Y locus encodes an anthocyanin MYB-like protein that activates the betalain red pigment pathway. Nat. Genet. 47: 92-96.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Heftmann, E. (1983). Review of steroids in Solanaceae. Phytochemistry 22: 1843-1860.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hibi, N., Higashiguchi, S., Hashimoto, T., and Yamada, Y. (1994). Gene expression in tobacco low-nicotine mutants. Plant Cell 6: 723-735.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Iijima, Y., Watanabe, B., Sasaki, R., Takenaka, M., Ono, H., Sakurai, N., Umemoto, N., Suzuki, H., Shibata, D., and Aoki, K. (2013). Steroidal glycoalkaloid profiling and structures of glycoalkaloids in wild tomato fruit. Phytochemistry 95: 145-157.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Itkin, M., Heinig, U., Tzfadia, O., Bhide, A.J., Shinde, B., Cardenas, P.D., Bocobza, S.E., Unger, T., Malitsky, S., Finkers, R., Tikunov, Y., Bovy, A., Chikate, Y., Singh, P., Rogachev, I., Beekwilder, J., Giri, A.P., and Aharoni, A. (2013). Biosynthesis of antinutritional alkaloids in Solanaceous crops is mediated by clustered genes. Science 341: 175-179.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Itkin, M., Rogachev, I., Alkan, N., Rosenberg, T., Malitsky, S., Masini, L., Meir, S., Iijima, Y., Aoki, K., de Vos, R., Prusky, D., Burdman, S., Beekwilder, J., and Aharoni, A. (2011). GLYCOALKALOID METABOLISM1 is required for steroidal alkaloid glycosylation and prevention of phytotoxicity in tomato. Plant Cell 23: 4507-4525.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Karimi, M., Inzé, D., and Depicker, A. (2002). GATEWAYTM vectors for Agrobacterium-mediated plant transformation. Trends Plant Sci. 7: 193-195.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Kellner, F., Kim, J., Clavijo, B.J., Hamilton, J.P., Childs, K.L., Vaillancourt, B., Cepela, J., Habermann, M., Steuernagel, B., Clissold, L., McLay, K., Buell, C.R., and O'Connor S.E. (2015). Genome-guided investigation of plant natural product biosynthesis. Plant J. http://dx.doi.org/10.1111/tpj.12827.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., and Higgins, D.G. (2007). Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

McCue, K.F., Allen, P. V, Shepherd, L.V.T., Blake, A., Maccree, M.M., Rockhold, D.R., Novy, R.G., Stewart, D., Davies, H. V, and Belknap, W.R. (2007). Potato glycosterol rhamnosyltransferase, the terminal step in triose side-chain biosynthesis. Phytochemistry 68: 327-334.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

McCue, K.F., Allen, P. V, Shepherd, L.V.T., Blake, A., Whitworth, J., Maccree, M.M., Rockhold, D.R., Stewart, D., Davies, H. V, and Belknap, W.R. (2006). The primary in vivo steroidal alkaloid glucosyltransferase from potato. Phytochemistry 67: 1590-1597.
Pubmed: Author and Title

CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

McCue, K.F., Shepherd, L.V.T., Allen, P. V, Maccree, M.M., Rockhold, D.R., Corsini, D.L., Davies, H. V, and Belknap, W.R. (2005). Metabolic compensation of steroidal glycoalkaloid biosynthesis in transgenic potato tubers: using reverse genetics to confirm the in vivo enzyme function of a steroidal alkaloid galactosyltransferase. Plant Sci. 168: 267-273.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Menke, F.L., Champion, A., Kijne, J.W., and Memelink, J. (1999). A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene Str interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. EMBO J. 18: 4455-4463.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Mennella, G., Lo Scalzo, R., Fibiani, M., D'Alessandro, A., Francese, G., Toppino, L., Acciarri, N., de Almeida, A.E., and Rotino, G.L. (2012). Chemical and bioactive quality traits during fruit ripening in eggplant (S. melongena L.) and allied species. J. Agric. Food Chem. 60: 11821-11831.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Milner, S.E., Brunton, N.P., Jones, P.W., O'Brien, N.M., Collins, S.G., and Maguire, A.R. (2011). Bioactivities of glycoalkaloids and their aglycones from Solanum species. J. Agric. Food Chem. 59: 3454-3484.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Mintz-Oron, S., Mandel, T., Rogachev, I., Feldberg, L., Lotan, O., Yativ, M., Wang, Z., Jetter, R., Venger, I., Adato, A., and Aharoni, A. (2008). Gene expression and metabolism in tomato fruit surface tissues. Plant Physiol. 147: 823-851.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Moco, S., Bino, R.J., Vorst, O., Verhoeven, H.A., de Groot, J., van Beek, T.A., Vervoort, J., and de Vos, C.H.R. (2006). A liquid chromatography-mass spectrometry-based metabolome database for tomato. Plant Physiol. 141: 1205-1218.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Moehs, C.P., Allen, P. V, Friedman, M., and Belknap, W.R. (1997). Cloning and expression of solanidine UDP-glucose glucosyltransferase from potato. Plant J. 11: 227-236.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006). Genome-Wide Analysis of the ERF Gene Family in Arabidopsis and Rice. Plant Physiol. 140: 411-432.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Nes, W.D. (2011). Biosynthesis of cholesterol and other sterols. Chem. Rev. 111: 6423-6451.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ohnmeiss, T.E., McCloud, E.S., Lynds, G., and Baldwin, I.T. (1997). Within plant relationships among wounding, jasmonic acid, and nicotine implications for defence in Nicotiana sylvestris. New Phytol. 137: 441-452.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ohyama, K., Okawa, A., Moriuchi, Y., and Fujimoto, Y. (2013). Biosynthesis of steroidal alkaloids in Solanaceae plants: involvement of an aldehyde intermediate during C-26 amination. Phytochemistry 89: 26-31.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Pekker, I., Alvarez, J.P., and Eshed, Y. (2005). Auxin Response Factors Mediate Arabidopsis Organ Asymmetry via Modulation of KANADI Activity. Plant Cell 17: 2899-2910.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Pollier, J., Moses, T., González-Guzmán, M., De Geyter, N., Lippens, S., Vanden Bossche, R., Marhavý, P., Kremer, A., Morreel, K., Guérin, C.J., Tava, A., Oleszek, W., Thevelein, J.M., Campos, N., Goormachtig, S., and Goossens, A. (2013). The protein quality control system manages plant defence compound synthesis. Nature 504: 148-152.

Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Roddick, J.G. (1996). Steroidal glycoalkaloids: nature and consequences of bioactivity. Adv. Exp. Med. Biol. 404: 277-295.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Saitou, N. and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406-425.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sánchez-Mata, M.C., Yokoyama, W.E., Hong, Y.J., and Prohens, J. (2010). a-solasonine and a-solamargine contents of gboma (Solanum macrocarpon L.) and scarlet (Solanum aethiopicum L.) eggplants. J. Agric. Food Chem. 58: 5502-5508.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sawai, S., Ohyama, K., Yasumoto, S., Seki, H., Sakuma, T., Yamamoto, T., Takebayashi, Y., Kojima, M., Sakakibara, H., Aoki, T., Muranaka, T., Saito, K., and Umemoto, N. (2014). Sterol Side Chain Reductase 2 is a key enzyme in the biosynthesis of cholesterol, the common precursor of toxic steroidal glycoalkaloids in potato. Plant Cell 26: 3763-3774.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Schwahn, K., Perez de Souza, L., Fernie, A.R., and Tohge, T. (2014). Metabolomics-assisted re?nement of the pathways of steroidal glycoalkaloid biosynthesis in the tomato clade. J. Integr. Plant Biol. 56: 864-875.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Shakya, R. and Navarre, D.A. (2008). LC-MS analysis of solanidane glycoalkaloid diversity among tubers of four wild potato species and three cultivars (Solanum tuberosum). J. Agric. Food Chem. 56: 6949-6958.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13: 2498-2504.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sheen, S.J. (1988). Detection of nicotine in foods and plant materials. J. Food Sci. 53: 1572-1573.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Shoji, T., Kajikawa, M., and Hashimoto, T. (2010). Clustered transcription factor genes regulate nicotine biosynthesis in tobacco. Plant Cell 22: 3390-3409.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Shoji, T. and Hashimoto, T. (2011). Tobacco MYC2 regulates jasmonate-inducible nicotine biosynthesis genes directly and by way of the NIC2-locus ERF genes. Plant Cell Physiol. 52: 1117-1130.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Shoji, T., Mishima, M., and Hashimoto, T. (2013). Divergent DNA-binding specificities of a group of ETHYLENE RESPONSE FACTOR transcription factors involved in plant defense. Plant Physiol. 162: 977-990.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sørensen, K.K., Kirk, H.G., Olsson, K., Labouriau, R., and Christiansen, J. (2008). A major QTL and an SSR marker associated with glycoalkaloid content in potato tubers from Solanum tuberosum x S. sparsipilum located on chromosome I. Theor. Appl. Genet. 117: 1-9.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30: 2725-2729.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Todd, A.T., Liu, E., Polvi, S.L., Pammett, R.T., and Page, J.E. (2010). A functional genomics screen identifies diverse transcription factors that regulate alkaloid biosynthesis in Nicotiana benthamiana. Plant J. 62: 589-600.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Vanden Bossche, R., Demedts, B., Vanderhaeghen, R., and Goossens, A. (2013). Transient expression assays in tobacco protoplasts. Methods Mol. Biol. 1011: 227-239.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

van der Fits, L. and Memelink, J. (2000). ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. Science 289: 295-297.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Wang, Z., Guhling, O., Yao, R., Li, F., Yeats, T.H., Rose, J.K., and Jetter, R. (2011). Two oxidosqualene cyclases responsible for biosynthesis of tomato fruit cuticular triterpenoids. Plant Physiol. 155: 540-552.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Xu, X., et al; Potato Genome Sequencing Consortium (2011). Genome sequence and analysis of the tuber crop potato. Nature 475: 189-195.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yamada,Y., Kokabu, Y., Chaki, K., Yoshimoto, T., Ohgaki, M., Yoshida, S., Kato, N., Koyama, T., and Sato, F. (2011). Isoquinoline alkaloid biosynthesis is regulated by a unique bHLH-type transcription factor in Coptis japonica. Plant Cell Physiol. 52: 1131-1141.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yamada, Y., Motomura, Y., and Sato, F. (2015). CjbHLH1 homologs regulate sanguinarine biosynthesis in Eschscholzia californica cells. Plant Cell Physiol. http://dx.doi.org/10.1093/pcp/pcv027.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhang, H., Hedhili, S., Montiel, G., Zhang, Y., Chatel, G., Pré, M., Gantet, P., and Memelink, J. (2011). The basic helix-loop-helix transcription factor CrMYC2 controls the jasmonate-responsive expression of the ORCA genes that regulate alkaloid biosynthesis in Catharanthus roseus. Plant J. 67: 61-71.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Zhong, S., Joung, J.G., Zheng, Y., Chen, Y.R., Liu, B., Shao, Y., Xiang, J.Z., Fei, Z., and Giovannoni, J.J. (2011). High-throughput illumina strand-specific RNA sequencing library preparation. Cold Spring Harb. Protoc. 6: 940-949.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title