(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 November 2019 (21.11.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/219933 A1

(51) International Patent Classification:
C07K 14/415 (2006.01) *C12N 15/82* (2006.01)

(21) International Application Number:
PCT/EP2019/062848

(22) International Filing Date:
17 May 2019 (17.05.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
18172996.3 17 May 2018 (17.05.2018) EP

(71) Applicants: **VIB VZW** [BE/BE]; Rijvisschestraat 120, 9052 Gent (BE). **UNIVERSITEIT GENT** [BE/BE]; Sint-Pietersnieuwstraat 25, 9000 Gent (BE).

(72) Inventors: **GOOSSENS, Alain**; Nijverheidstraat 92, 9160 Lokeren (BE). **POLLIER, Jacob**; Hertooiebos 12A, 9052 Gent (BE). **COLINAS MARTINEZ, Maite**; Forelstraat 24, 9000 Gent (BE). **SCHWEIZER, Fabian**; Chemin du Sésau 1, 1084 Carrouge (CH).

(74) Common Representative: **VIB VZW**; Rijvisschestraat 120, 9052 Gent (BE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title:** AN ENGINEERED COMBINATORIAL MODULE OF TRANSCRIPTION FACTORS TO BOOST PRODUCTION OF MONOTERPENOID INDOLE ALKALOIDS

(57) **Abstract:** The present invention relates to the field of plant secondary metabolites, more particularly to methods of increasing the production of specific plant secondary metabolites. Even more particularly the present invention relates to the identification of a combination of transcription factors that enhances the production of monoterpenoid indole alkaloids.

WO 2019/219933 A1

AN ENGINEERED COMBINATORIAL MODULE OF TRANSCRIPTION FACTORS TO BOOST PRODUCTION OF MONOTERPENOID INDOLE ALKALOIDS

Field of the invention

The present invention relates to the field of plant secondary metabolites with agricultural, pharmacological or other industrial properties, more particularly to methods of increasing the activity of specific plant secondary metabolites. Even more particularly the present invention relates to the identification of a combination of transcription factors that enhances the production of monoterpenoid indole alkaloids.

Background

Plants synthesize an overwhelming variety of specialized metabolites such as monoterpenoid indole alkaloids (MIAs) with an enormous range of biological activities relevant for the pharmaceutical (e.g. antitumor), chemical industry (e.g. additives to foods and cosmetics) as well as for the agricultural industry (e.g. antifungal, antibacterial, molluscicidal, insecticidal, and anti-feeding activities) (Suzuki et al. 2002 Plant Journal 32: 1033-1048; Huhman et al. 2005 J Agric Food Chem 53: 1914-1920; Augustin et al. 2011 Phytochemistry 72: 435-457; Osbourn et al. 2011 Natural Product Reports 28: 1261-1268).

Like for many specialized metabolites, MIA biosynthesis is regulated by the plant hormone jasmonate (JA) and serves predominantly a defensive role against insects and pathogens (De Geyter et al., 2012; Dugé de Bernonville et al., 2017; Goossens et al., 2017; Roepke et al., 2010; Wasternack and Hause, 2013; Zhang et al., 2011; Zhou and Memelink, 2016). In general, in the absence of JA, in particular its bioactive form JA-isoleucine (JA-Ile), the activity of positive regulators of the JA response, such as the transcription factor (TF) MYC2, is physically blocked by the interaction with JASMONATE-ZIM DOMAIN (JAZ) proteins (Chini et al., 2016; Pauwels and Goossens, 2011; Wasternack and Strnad, 2018). Insect attack, wounding or JA treatment is followed by a burst of JA-Ile that promotes the interaction between the JAZ repressor proteins and the F-box CORONATINE INSENSITIVE 1 (COI1) protein, leading to ubiquitination and consequent proteasomal degradation of the JAZ proteins. Subsequently, TFs such as MYC2 are de-repressed, leading to transcriptional activation of their target genes (Chini et al., 2016; Goossens et al., 2017; Wasternack and Strnad, 2018; Zhou and Memelink, 2016). Recently, it has been shown that specific amino acid changes in the JAZ Interaction Domain (JID) of Arabidopsis thaliana MYC2 (AtMYC2) can prevent the interaction with the JAZ proteins, resulting in a constitutively active (hereafter called 'de-repressed') form, such as AtMYC2D105N (Frerigmann et al., 2014; Gasperini et al., 2015; Goossens et al., 2015; Smolen et al., 2002; Zhang et al., 2015).

Besides MYC2, JA-mediated elicitation of MIA biosynthesis in *C. roseus* organs involves several other TFs. Expression of the IPAP-specific MIA biosynthesis genes is controlled by the BASIC HELIX-LOOP-HELIX

IRIDOID SYNTHESIS (BIS) 1 and BIS2 TFs (Fig. 1A) (Van Moerkercke et al., 2016; Van Moerkercke et al., 2015). The expression of the epidermis-specific genes steering the following steps leading to strictosidine is controlled by the clade of the OCTADECANOID DERIVATIVE RESPONSIVE CATHARANTHUS APETALA2-DOMAIN (ORCA) 2, ORCA3, ORCA4, and ORCA5 TFs (Menke et al., 1999; Paul et al., 2017; van der Fits and Memelink, 2000; Zhou and Memelink, 2016). The expression of the BIS TFs as well as the ORCA TFs can be stimulated by JAs (Paul et al., 2017; Van Moerkercke et al., 2016; Van Moerkercke et al., 2015; Zhang et al., 2011). Overexpression of CrMYC2 in cell suspension and hairy root cultures has been shown to induce the expression of ORCA genes (Paul et al., 2017; Zhang et al., 2011) and CrMYC2 and ORCA3 have both been shown to be capable of transactivating the TDC promoter in tobacco protoplasts (Paul et al., 2017). However, the mechanism by which the whole MIA pathway and/or the above TFs are controlled by the primary MYC2 JA signaling module is not well understood yet, because evidence for both indirect and direct actions exist.

The medicinal plant *Catharanthus roseus* (Madagascar periwinkle) produces over 130 monoterpenoid indole alkaloids (MIAs), including the anticancer compounds vinblastine and vincristine (Fig. 1A) (van der Heijden et al., 2004). Today, *C. roseus* still remains the only source of these compounds, as their total chemical synthesis is economically not viable (Eastman, 2011; van der Heijden et al., 2004). Developing alternatives such as biotechnology-based production methods in hairy root, suspension cell or yeast cultures has been a longstanding research subject (Facchini and De Luca, 2008; O'Connor, 2015; van der Heijden et al., 2004). Some precursors of vinblastine have been produced successfully in heterologous systems, including the MIA precursor strictosidine in yeast and tobacco (Brown et al., 2015; Miettinen et al., 2014) and vindoline in yeast after feeding with the precursor tabersonine (Qu et al., 2015). However, because of the still incomplete knowledge of the entire MIA pathway, yields of current approaches are suboptimal and thus it would be advantageous to find alternative strategies to boost the production of MIAs, more particular the anticancer compounds vinblastine and vincristine.

Summary

Applicant made use of a flower petal transformation method in the medicinal plant *Catharanthus roseus* to shed light on the complex regulatory mechanisms steering the jasmonate-modulated biosynthesis of monoterpenoid indole alkaloids (MIAs), to which the anti-cancer compounds vinblastine and vincristine belong. By combinatorial overexpression of the transcriptional activators BIS1, ORCA3 and MYC2a, an unprecedented insight into the modular transcriptional control of MIA biosynthesis is provided. Furthermore, it is shown that the expression of an engineered de-repressed MYC2a triggers a tremendous reprogramming of the MIA pathway, finally leading to massively increased accumulation of at least 23 MIAs.

Therefore, a first object of the application is a *Catharanthus roseus* MYC2 protein with a D to N mutation at a position relative to position 126 of SEQ ID No. 1. In one embodiment, said MYC2 protein further comprises a JAZ interacting domain comprising SEQ ID No 7 and/or SEQ ID No. 8. In a further embodiment, said MYC2 protein has at least 90% homology to SEQ ID No. 1. In a more particular embodiment, a mutant *Catharanthus roseus* MYC2a protein as depicted in SEQ ID No. 2 is provided. It is also an object of the application to provide a nucleic acid molecule encoding one of the above MYC2 or MYC2a proteins.

A third object of the application is a plant cell expressing the above described nucleic acid molecule or comprising one of the MYC2 or MYC2a proteins described. In a more particular embodiment, said plant cell further expresses the transcriptional activators BIS1 and ORCA3. In an even more particular embodiment, the expression level of BIS1 and ORCA3 in said plant cell is at least 10% higher than the expression level of BIS1 and ORCA3 in a control plant cell.

A fourth object of the application is a regulatory module comprising one of the MYC2 or MYC2a proteins described above, further comprising the transcription factors BIS1 and ORCA3, wherein said BIS1 and ORCA3 transcription factors are transcriptionally controlled by promoters different to the promoters to which they are naturally linked.

A fifth object of the application is the use of one of the described MYC2 or MYC2a proteins or the use of the described nucleic acid molecule or the use of the described regulatory module to enhance the production of at least one monoterpenoid indole alkaloid in a plant cell with at least 10% compared to a control plant cell. In a particular embodiment, said at least one monoterpenoid indolalkaloid is a 16-hydroxytabersonine-derived MIA and/or 19-hydroxytabersonine-derived MIA. In an even more particular embodiment, said at least one MIA is selected from the list consisting of strictosidine, catharanthine, 16-hydroxytabersonine, 16-methoxylhörhammericine, hörhammericine, serpentine, vindoline, vinblastine, akuammicine, isositsirikine, 16-methoxytabersonine, 3-hydroxy-16-methoxy-2,3-dihydrotabersonine, 19-hydroxytabersonine, 16-hydroxyhörhammericine, minovincinine, vincadifformine, 16-hydroxyvincadifformine, 16-hydroxy-19-O-acetyltabersonine, perivine and O-acetylstemmadine.

A sixth object of the application is a method of producing monoterpenoid indole alkaloids, said method comprising expressing one of the MYC2 or MYC2a proteins described in this specification in a plant cell; optionally further expressing the BIS1 and/or ORCA3 transcription factors; and selecting plant cells with increased production of at least one monoterpenoid indole alkaloid compared to control plant cells not expressing one of said MYC2 of MYC2a proteins.

Brief description of the drawings

Figure 1. MIA biosynthesis in *C. roseus* is regulated by CrMYC2a, BIS1 and ORCA3. (A) Pathway leading to the production of shoot-specific MIAs in *C. roseus*. (B) Pathway leading to the root-specific MIAs. Genes regulated by CrMYC2a, BIS1 and ORCA are boxed in orange, blue and green, respectively. Genes not regulated by any of these TFS are boxed in yellow. 7DLGT, 7-deoxyloganetic acid glucosyl transferase; 7DLH, 7-deoxyloganic acid hydroxylase; 8HGO, 8-hydroxygeraniol oxidoreductase; BIS1/2, bHLH iridoid synthesis 1 and 2; D4H, desacetoxy vindoline 4-hydroxylase; DAT, deacetylvindoline O-acetyltransferase; G8O, geraniol-8-oxidase; GES, geraniol synthase; GO, geissoschizine oxidase; GS1/2, geissoschizine synthase 1 and 2; HL1/2, α/β hydrolase 1 and 2; HYS, heteroyohimbine synthase; IO, iridoid oxidase; IS, iridoid synthase; LAMT, loganic acid O-methyltransferase; MAT, minovincinine-19-hydroxy-O-acetyltransferase; NMT, N methyltransferase; NPF2.4/5/6/9, nitrate/peptide family; ORCA3, octadecanoid-derivative responsive Catharanthus AP2-domain protein 3; PRX1, peroxidase 1; SAT, stemmadenine-O-acetyltransferase; SGD, strictosidine B-glucosidase; SLS, secologanin synthase; STR, strictosidine synthase; T16H1/2, tabersonine 16-hydroxylase 1 and 2; T16OMT, tabersonine 16-O-methyltransferase; T19H, tabersonine 19-hydroxylase; T3O, tabersonine 3-oxygenase; T3R, tabersonine 3-reductase; TAT, tabersonine derivative 19-O-acetyltransferase; TDC, tryptophan decarboxylase; THAS, tetrahydroalstonine synthase.

Figure 2. Y2H assays to assess interaction between *C. roseus* JAZ and bHLH proteins. A: BIS1 and BIS2 do not interact with JAZ proteins in a Y2H assay. The full-length open reading frame sequences of BIS1 and BIS2 were cloned in the activation domain (AD) vectors whereas JAZ1, 2, 3, 4, 7 and 8 were cloned in DNA-binding domain (BD) vectors, as indicated. Transformed yeasts were spotted in 10-fold and 100-fold dilutions on control medium (-2) or selective medium (-3). Empty vectors were used as the negative control, whereas a clone with the BIS2-ACT domain was used as the positive control. B: Y2H growth control. Yeast cultures used for the Y2H assay shown in Fig. 3B to detect MYC2-JAZ interaction were additionally grown in 1:10 and 1:100 dilutions on control medium lacking Leu and Trp (-2).

Figure 3. De-repressed CrMYC2aD126N transactivates the promoters of BIS1, ORCA3 and MIA biosynthesis genes. (A) De-repressed CrMYC2aD126N and AtMYC2D105N can transactivate pBIS1, pORCA3 and pSTR in transient expression assays in N. tabacum protoplasts. The y-axis shows fold change in normalized fLUC activity relative to the control transfection with GUS. The error bars designate SE of the mean (n = 8). Statistical significance was determined by the Student's t test (*P < 0.05, **P < 0.01, ***P < 0.001). (B) Y2H assay showing that de-repressed CrMYC2aD126N and AtMYC2D105N lose their capacity to interact with some but not all JAZ proteins, as compared to the wild-type versions. The FL-

ORF sequences of CrMYC2a, CrMYCaD126N, AtMYC2 and AtMYC2D105N were cloned in the activation domain (AD) vectors whereas those of the JAZ genes were cloned in DNA-binding domain (BD) vectors, as indicated. Transformed yeasts were spotted in 10-fold and 100-fold dilutions on selective medium (this figure) and control medium (see Fig. 2B). Empty vectors were used as the negative control. (C) CrMYC2aD126N, but not CrMYC2a, transactivates the promoter of MIA pathway genes in N. tabacum protoplasts. The y-axis shows fold change in normalized fLUC activity relative to the control transfection with GUS. The error bars designate SE of the mean (n = 4). Statistical significance was determined by the Student's t-test (*P < 0.05, **P < 0.01, ***P < 0.001). (D-G) qPCR data show that only the derepressed version of CrMYC2a is able to activate expression of BIS1 (D), BIS2 (E), ORCA2 (F) and ORCA3 (G).

Figure 4. Phylogenetic tree of *A. thaliana* and *C. roseus* bHLH proteins showing the relationship between clade III, clade IV and previously published MYC proteins. Arabidopsis bHLH proteins were used as bait to screen for clade III orthologs within the *C. roseus* transcriptome (http://bioinformatics.psb.ugent.be/orcae/overview/SmartCell). A neighbor-joining tree was constructed in MEGA6 based on an alignment of the full-length proteins (Tamura et al., 2013). Bootstrap analyses were performed with 1,000 bootstrap replicates. The tree was rooted on to the branch containing clade XII, IX, VIIb bHLH TFs.

Figure 5. Amino acid alignment showing conservation of aspartate in the JID domain. ClustalW multiple alignment of the JAZ interacting domain (JID) within *C. roseus* and *A. thaliana* MYC2 homologs. Red box indicates the mutated aspartate (D) in the de-repressed MYC2 homologs.

Figure 6. Expression of MIA pathway genes and metabolite profiling of *C. roseus* flower petals transiently overexpressing GUS, CrMYC2a, CrMYC2aD126N, AtMYC2 or AtMYCD105N under the CaMV35S promoter. (A) Expression of transcription factors, MIA pathway genes and transporters was measured by qPCR and represented in fold change compared to control samples (infiltrated with pCaMV35S::GUS constructs). (B) High abundant MIAs were measured on LC-ESI-FT-ICR-MS and expressed in average total ion current (TIC). (C) PCA plot and PC loading plot derived from the LC-ESI-FT-ICR-MS metabolome data of flower petals transiently overexpressing GUS, CrMYC2a, CrMYC2aD126N, AtMYC2, or AtMYC2D105N. The color code for the overexpression cassettes is depicted at the bottom of the figure. The error bars designate SE of the mean (n=4). Different letters within each expression series indicate statistically significant differences at P < 0.01 (Tukey's highly significant difference test).

Figure 7. Profiling of MIA pathway gene expression and metabolite accumulation in *C. roseus* flower petals transiently overexpressing the combination of CrMYC2aD126N, BIS1 and ORCA3 under control of the CaMV35S promoter. (A) Expression of MIA biosynthesis genes and transporters was measured by qPCR and represented in fold change compared to control samples (infiltrated with pCaMV35S::GUS). Error bars designate SE of the mean (n = 4). Different letters within each sample indicate statistically significant differences at $P < 0.01$ (Tukey's highly significant difference test). (B) LC-ESI-FT-ICR-MS chromatograms of extracts from flowers transiently overexpressing the GUS control (black) and BIS1/MYC2aD126N/ORCA3 (red) are depicted. (C) Levels of differentially accumulating high abundant MIAs are expressed in peak area of total ion current (TIC). Error bars designate SE of the mean (n = 4). Different letters within each expression series indicate statistically significant differences at $P < 0.01$ (Tukey's highly significant difference test). (D) PCA plot and PC loading plot derived from LC-ESI-FT-ICR-MS metabolome data of flower petals transiently overexpressing GUS, MYC2aD126N, ORCA3, BIS1 or the combination of MYC2aD126N, ORCA3 and BIS1.

Figure 8. Analysis of TF amplification loops in the regulation of *C. roseus* MIA biosynthesis. Expression profile of TFs in *C. roseus* flower petals transiently overexpressing the combination of CrMYC2aD126N, BIS1 and ORCA3. Expression of MIA regulators was measured by qPCR and represented in fold change compared to control samples (infiltrated with pCaMV35S::GUS). Error bars designate SE of the mean (n = 4). Different letters within each sample indicate statistically significant differences at $P < 0.01$ (Tukey's highly significant difference test).

Figure 9. A simplified model illustrating the complex transcriptional regulation of the iridoid and MIA pathway. In the absence of stimuli, JAZ repressors bind the JID domain of CrMYC2a and thereby inhibit its transcriptional activity. Upon insect feeding, wounding or JA elicitation, JAZ proteins are degraded, leading to the de-repression of CrMYC2a and consequently direct induction of the whole MIA pathway. Concomitantly, active CrMYC2a induces expression of BIS and ORCA TFs, inserting a second amplification layer in the induction of the MIA pathway. Additionally, BIS1 induces an amplification loop via BIS2, whereas ORCA3 seems to be controlled by an auto-amplification loop and potential cross-amplification loops involving other ORCA3 homologs. In parallel, we postulate that active CrMYC2a induces transcription of JAZ repressors, as in other plant species, resulting in de novo repression of CrMYC2a activity, followed by subsequent downregulation of BIS TFs, ORCA TFs and MIA pathway genes. In the case of the constitutive de-repression due to the D126N amino acid change in CrMYC2a, the negative feedback loop of JAZ is disrupted, resulting in high transcript levels of BIS1, BIS2, ORCA3 and MIA genes. The whole regulatory mechanism is therefore in a free running state leading to a constant high

expression of the MIA pathway genes, resulting in higher accumulation of MIAs. GPP, geranyl diphosphate.

Detailed description of the invention

The present invention will be described with respect to particular embodiments and with reference to certain drawings but the invention is not limited thereto but only by the claims. Any reference signs in the claims shall not be construed as limiting the scope. The drawings described are only schematic and are non-limiting. In the drawings, the size of some of the elements may be exaggerated and not drawn on scale for illustrative purposes. Where the term "comprising" is used in the present description and claims, it does not exclude other elements or steps. Where an indefinite or definite article is used when referring to a singular noun e.g. "a" or "an", "the", this includes a plural of that noun unless something else is specifically stated. Furthermore, the terms first, second, third and the like in the description and in the claims, are used for distinguishing between similar elements and not necessarily for describing a sequential or chronological order. It is to be understood that the terms so used are interchangeable under appropriate circumstances and that the embodiments of the invention described herein are capable of operation in other sequences than described or illustrated herein.

Unless otherwise defined herein, scientific and technical terms and phrases used in connection with the present invention shall have the meanings that are commonly understood by those of ordinary skill in the art. Generally, nomenclatures used in connection with, and techniques of molecular and cellular biology, genetics and protein and nucleic acid chemistry and hybridization described herein are those well-known and commonly used in the art. The methods and techniques of the present invention are generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification unless otherwise indicated. See, for example, Sambrook et al. Molecular Cloning: A Laboratory Manual, 2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989); Ausubel et al., Current Protocols in Molecular Biology, Greene Publishing Associates (1992, and Supplements to 2002); Leach, Molecular Modelling: Principles and Applications, 2d ed., Prentice Hall, New Jersey (2001).

To fend off microbial pathogens and herbivores, plants have evolved a wide range of defense strategies such as physical barriers, or the production of anti-digestive proteins or bioactive specialized metabolites. Accumulation of the latter compounds is often regulated by transcriptional activation of the biosynthesis pathway genes by the phytohormone jasmonate-isoleucine. Applicant has developed a transient expression system in *Catharanthus roseus* flower petals, thereby creating a rapid and non-labor

intensive screening platform for pathway mapping, gene discovery and assessment of engineering tools (Van Moerkercke et al., 2016). Here, this platform has been employed to further unravel the organization of the regulatory transcription factor circuit that drives MIA biosynthesis in *C. roseus*. Subsequently, a combinatorial module has been designed of 'wild-type' and engineered, i.e. gain-of-function mutant, versions of three *C. roseus* transcription factors (TFs) that together boost MIA biosynthesis in a coordinated manner across cell types. Some of these MIAs have highly interesting activities on their own or are base molecules (e.g. strictosidine) for numerous pharmaceutically valuable metabolites including quinine, camptothecin, ajmalicine, serpentine, vinblastine and vincristine.

MIA biosynthesis in *C. roseus* is organ and cell-type specific and compartmentalized at the cellular and subcellular level (Courdavault et al., 2014; Dugé de Bernonville et al., 2015). At the organ level, bisindole alkaloids such as vinblastine and vincristine, as well as vindoline are produced exclusively in the areal parts of the plant, whereas catharanthine accumulates in all organs (Pan et al., 2016; van der Heijden et al., 2004). Other abundant MIAs such as ajmalicine, serpentine and the tabersonine-derived MIAs hörhammericine, lochnericine and 19-hydroxytabersonine are mostly present in roots (Laflamme et al., 2001; Pan et al., 2016; Rodriguez et al., 2003; van der Heijden et al., 2004). At the cellular level, the first seven biosynthesis steps from geranyl diphosphate (GPP) to the iridoid loganic acid take place in internal phloem associated parenchyma (IPAP) cells and involve GERANIOL SYNTHASE (GES), 8-HYDROXYGERANIOL OXIDOREDUCTASE (8HGO), GERANIOL 8-OXIDASE (G8O), IRIDOID SYNTHASE (IS), IRIDOID OXIDASE (IO), 7-DEOXYLOGANETIC ACID GLUCOSYLTRANSFERASE (7DLGT) and 7-DEOXYLOGANIC ACID HYDROXYLASE (7DLH) (Asada et al., 2013; Geu-Flores et al., 2012; Miettinen et al., 2014; Salim et al., 2014; Simkin et al., 2013). The next four enzymatic steps involving LOGANIC ACID O-METHYLTRANSFERASE (LAMT), SECOLOGANIN SYNTHASE (SLS), TRYPTOPHAN DECARBOXYLASE (TDC) and STRICTOSIDINE SYNTHASE (STR) take place in epidermal cells, resulting in the formation of strictosidine (Courdavault et al., 2014; St-Pierre et al., 1999) (Fig. 1A). Transport of the iridoid intermediate(s) from the IPAP to the epidermal cells involves NITRATE/PEPTIDE FAMILY (NPF) transporters (NPF2.4, NPF2.5, NPF2.6) (Larsen et al., 2017). Likewise, strictosidine is exported from the epidermal vacuole by NPF2.9 (Payne et al., 2017), transported to the nucleus via a yet unknown mechanism and subsequently deglucosylated by STRICTOSIDINE B-GLUCOSIDASE (SGD), resulting in a highly reactive aglycone that can be converted to (i) the corynanthe-type ajmalicine or tetrahydroalstonine by HETEROYOHIMBINE SYNTHASE (HYS) or TETRAHYDROALSTONINE SYNTHASE (THAS), respectively, or (ii) vindoline or catharanthine by the action of multiple enzymes (Courdavault et al., 2014; Stavrinides et al., 2016) (Fig. 1A). The latter involve GEISSOSCHIZINE SYNTHASE 1 and 2 (GS1 and 2) that, in conjunction with GEISSOSCHIZINE OXIDASE (GO), produce preakuammicine, a precursor of both vindoline and catharanthine (Tatsis et al., 2017). The pathway enzymes catalyzing the conversion

of the intermediate preakuammicine to catharanthine or tabersonine remain elusive. In contrast, conversion of tabersonine to vindoline has been resolved and involves the action of seven enzymes spread over three cell types, i.e. the epidermis, the laticifers and the idioblasts (Besseau et al., 2013; Guirimand et al., 2011; Levac et al., 2008; Liscombe et al., 2010; Qu et al., 2015; St-Pierre et al., 1999). The monomeric MIAs vindoline and catharanthine can dimerize to form 3',4'-anhydrovinblastine, the direct precursor of vinblastine and vincristine (van der Heijden et al., 2004).

In a first aspect, a *Catharanthus roseus* MYC2 protein is provided. More particularly said MYC2 protein has a D to N mutation at a position relative to position 126 of SEQ ID No. 1. Even more particularly, said MYC2 protein has a D to N mutation at position 126 of SEQ ID No. 1.

As used herein, the terms "polypeptide", "protein, "peptide" are used interchangeably and refer to a polymeric form of amino acids of any length, which can include coded and non-coded amino acids, chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones.

A "position relative to position 126" as used herein refers to the precise location of the D (aspartic acid) in the original sequence SEQ ID No. 1. As the D to N mutation at position 126 of SEQ ID No. 1 is crucial to the invention because it makes the activity of the mutated MYC2 independent of JA signaling (i.e. loss of interaction with and loss of repression by JAZ proteins) and thus turns the mutant MYC2 protein into a derepressed form, the upstream amino acid sequence could be modified (e.g. by adding one or more amino acids to the N-terminal part of said MYC2) without losing the characteristics of the depressed MYC2 protein. However, in that case it could be that position 126 does not refer anymore to the crucial D. To make it clear that the D at position 126 of SEQ ID No. 1 should be changed into N, Applicant recites the "relative position", meaning that the precise location of the D could be moved. In case the crucial D to N mutation maintains its position at 126 of said MYC2 protein of the invention, the position relative to position 126 of SEQ ID No. 1 is identical as the position 126 of SEQ ID No. 1.

In a more particular embodiment, said MYC2 protein further comprises a JAZ interacting domain comprising SEQ ID No 7 (VLGWGNGYYKGEEDK) and/or SEQ ID No. 8 (WTYAIFWQXS, wherein X is S or P). In a further particular embodiment, said MYC2 protein comprises a JAZ interacting domain consisting of SEQ ID No 7 (VLGWGNGYYKGEEDK) and/or SEQ ID No. 8 (WTYAIFWQXS, wherein X is S or P). In a more particular embodiment, said MYC2 protein comprises a JAZ interacting domain comprising WTYAIFWQXSXXXFXGXS (SEQ ID No. 9), wherein X can be every amino acid or a gap. In an even more particular embodiment, said MYC2 protein comprises a JAZ interacting domain comprising SEQ ID No 7 (VLGWGNGYYKGEEDK) and WTYAIFWQXSXXXFXGXS (SEQ ID No. 9), wherein X can be every amino acid

or a gap. In a most particular embodiment said MYC2 protein comprises a JAZ interacting domain comprising or consisting of SEQ ID No. 10 (WTYAIFWQSSVVEFAGPSVLGWGNGYYKGEEDK).

The "JAZ interacting domain" of "JID" as used herein is well-known by the skilled person as a conserved amino acid domain which is essential for interaction with JAZ proteins (Fernández-Calvo et al., 2011 Plant Cell 23: 701-715).

In an even more particular embodiment, said MYC2 protein has at least 90%; at least 91%; at least 92%; at least 93%; at least 94%; at least 95%; at least 96%; at least 97%; at least 98%; at least 99% or at least 99.5% homology to SEQ ID No. 1.

Homologs of a protein encompass peptides, oligopeptides and polypeptides having amino acid substitutions, deletions and/or insertions, preferably by a conservative change, relative to the unmodified protein in question and having similar biological and functional activity as the unmodified protein from which they are derived; or in other words, without significant loss of function or activity. Orthologs and paralogs, which are well-known terms by the skilled person, define subcategories of homologs and encompass evolutionary concepts used to describe the ancestral relationships of genes. Paralogs are genes within the same species that have originated through duplication of an ancestral gene; orthologs are genes from different organisms that have originated through speciation, and are also derived from a common ancestral gene. Several different methods are known by those of skill in the art for identifying and defining these functionally homologous sequences. General methods for identifying orthologues and paralogues include phylogenetic methods, sequence similarity and hybridization methods. Percentage similarity and identity can be determined electronically. Examples of useful algorithms are PILEUP (Higgins & Sharp, CABIOS 5:151 (1989), BLAST and BLAST 2.0 (Altschul et al. J. Mol. Biol. 215: 403 (1990). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). The "homology" or "sequence identity" of two related nucleotide or amino acid sequences, expressed as a percentage, refers to the number of positions in the two optimally aligned sequences which have identical residues (x100) divided by the number of positions compared. A gap, i.e., a position in an alignment where a residue is present in one sequence but not in the other is regarded as a position with non-identical residues. The alignment of the two sequences is performed by the Needleman and Wunsch algorithm (Needleman and Wunsch (1970) J Mol Biol. 48: 443-453). The computer-assisted sequence alignment above, can be conveniently performed using standard software program such as GAP which is part of the Wisconsin Package Version 10.1 (Genetics Computer Group, Madison, Wisconsin, USA) using the default scoring matrix with a gap creation penalty of 50 and a gap extension penalty of 3. Sequences are indicated as "essentially similar" when such sequences have a sequence identity of at least about 75%,

particularly at least about 80 %, more particularly at least about 85%, quite particularly about 90%, especially about 95%, more especially about 100%, quite especially are identical.

In most particular embodiments, said MYC2 protein from the invention is a MYC2a protein, more particularly a mutant MYC2a protein, even more particularly a mutant *Catharanthus roseus* MYC2a protein, most particularly the mutant *Catharanthus roseus* MYC2a protein as depicted in SEQ ID No. 2.

In particular embodiments of the first aspect, the above described MYC2 or MYC2a proteins are recombinantly produced proteins. A "recombinantly produced protein", as used herein, refers to a protein produced by recombinant methods, e.g. recombinant DNA technology. For the purposes of the invention, "recombinant", "transgene" or "transgenic" means with regard to, for example, a protein, a nucleic acid sequence, a chimeric gene construct or a vector comprising said nucleic acid sequence or a cell transformed with the said nucleic acid sequences, chimeric gene constructs or vectors, all those constructions brought about by recombinant methods in which either (a) the nucleic acid sequences encoding proteins useful in the methods of the invention, or (b) genetic control sequence(s) which is operably linked with the nucleic acid sequence according to the invention, for example a promoter, or (c) a) and b) are not located in their natural genetic environment or have been modified by recombinant methods, it being possible for the modification to take the form of, for example, a substitution, addition, deletion, inversion or insertion of one or more nucleotide residues. The natural genetic environment is understood as meaning the natural genomic or chromosomal locus in the original plant or the presence in a genomic library. In the case of a genomic library, the natural genetic environment of the nucleic acid sequence is preferably retained, at least in part. The environment flanks the nucleic acid sequence at least on one side and has a sequence length of at least 50 bp, preferably at least 500 bp, especially preferably at least 1000 bp, most preferably at least 5000 bp. A naturally occurring expression cassette – for example the naturally occurring combination of the natural promoter of the nucleic acid sequences with the corresponding nucleic acid sequence encoding one of the MYC2 or MYC2 proteins described in this specification – becomes a transgenic expression cassette when this expression cassette is modified by non-natural, synthetic ("artificial") methods such as, for example, mutagenic treatment. Suitable methods are described, for example, in US 5,565,350 or WO 00/15815.

In one embodiment, said recombinantly produced protein is a man-made or non-naturally occurring fusion protein.

In a second aspect, a nucleic acid molecule is provided encoding any of the MYC2 or MYC2a proteins described in the first aspect of this specification.

As used herein, the terms "nucleic acid", "polynucleotide", “polynucleic acid” are used interchangeably and refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof. Polynucleotides may have any three-dimensional structure, and may perform any function, known or unknown. Non-limiting examples of polynucleotides include a gene, a gene fragment, exons, introns, messenger RNA (mRNA), transfer RNA, ribosomal RNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, control regions, isolated RNA of any sequence, nucleic acid probes, and primers. The polynucleotide molecule may be linear or circular. The polynucleotide may comprise a promoter, an intron, an enhancer region, a polyadenylation site, a translation initiation site, 5’ or 3’ untranslated regions, a reporter gene, a selectable marker or the like. The polynucleotide may comprise single stranded or double stranded DNA or RNA. The polynucleotide may comprise modified bases or a modified backbone. A nucleic acid that is up to about 100 nucleotides in length, is often also referred to as an oligonucleotide.

To express one of the above disclosed nucleic acid molecules, said nucleic acid molecule must be linked operably to or comprise a suitable promoter which expresses said nucleic acid molecule at the right point in time and with the required spatial expression pattern. The term "operably linked" as used herein refers to a linkage in which the promoter or regulatory sequence is contiguous with the gene of interest to control the gene of interest (i.e. initiate the transcription of the gene of interest), as well as a promoter that act in trans or at a distance to control the gene of interest. For example, a DNA sequence is operably linked to a promoter when it is ligated to the promoter downstream with respect to the transcription initiation site of the promoter and allows transcription elongation to proceed through the DNA sequence. Linkage of DNA sequences to regulatory sequences is typically accomplished by ligation at suitable restriction sites or adapters or linkers inserted instead of using restriction endonucleases known to one of skill in the art.

A promoter that enables the initiation of gene transcription in a host cell is referred to as being “active”. To identify a promoter which is active in a host cell, the promoter can be operably linked to a reporter gene after which the expression level and pattern of the reporter gene can be assayed. Suitable well-known reporter genes include for example beta-glucuronidase, beta-galactosidase or any fluorescent or luminescent protein. The term "promoter activity" refers to the extent of transcription of a polynucleotide sequence, homologue, variant or fragment thereof that is operably linked to the promoter whose promoter activity is being measured. The promoter activity is assayed by measuring the enzymatic activity of the beta-glucuronidase or beta-galactosidase. Alternatively, promoter strength may also be assayed by quantifying mRNA levels or by comparing mRNA levels of the nucleic acid, with mRNA levels of housekeeping genes such as 18S rRNA, using methods known in the art, such as Northern

blotting with densitometric analysis of autoradiograms, quantitative real-time PCR or RT-PCR (Heid et al., 1996 Genome Methods 6: 986-994).

The promoter linked to the nucleic acid molecule of interest can be the endogenous promoter of said nucleic acid molecule, i.e. the promoter that regulates expression of said nucleic acid molecule as found in nature. Alternatively, the promoter can be an exogenous promoter, i.e. a regulatory nucleic acid sequence which differs from the sequence to which said nucleic acid molecule is associated in nature. Therefore also a chimeric gene is provided in this application, wherein said chimeric gene is a recombinant nucleic acid sequence comprising one of the nucleic acid molecules disclosed in this specification and a promoter or regulatory nucleic acid sequence, wherein said promoter or regulatory nucleic acid sequence of the chimeric gene is not operably linked to the said nucleic acid sequence as found in nature. This is equivalent as saying that said promoter or regulatory nucleic acid sequence to which the nucleic acid molecule is operably linked differs from the promoter or regulatory nucleic acid sequence operably linked or associated with said nucleic acid molecule in the natural environment.

A non-limiting example of an exogenous promoter for expression of a gene of interest in plants is the 35S promoter. The "35S promoter" or the "cauliflower mosaic virus (CaMV) 35S promoter" is a constitutive or constant active promoter that directs high-level expression in a wide range of cells under a wide range of conditions and in most plant tissues including monocots. Examples of other constitutive plant promoters useful for expressing heterologous, modified or non-modified polypeptides in plant cells include, but are not limited to, the nopaline synthase promoter and the octopine synthase promoter.

For expression of the nucleic acid molecule encoding one of the MYC2 or MYC2a protein disclosed in this specification in yeast, the "AOX promoter" or "AOX1 promoter" or "inducible AOX1 promoter" that originates from *Pichia pastoris* can be used. The alcohol oxidase AOX1 and AOX2 promoters are strongly inducible by methanol and are repressed by e.g. glucose. These genes allow *Pichia* to use methanol as a carbon and energy source. Usually the gene for the desired protein is introduced under the control of the AOX1 promoter, which means that protein production can be induced by the addition of methanol.

In a particular embodiment, said chimeric gene comprising a nucleic acid molecule encoding for one of the MYC2 or MYC2a proteins described in this specification operably linked to an endogenous or exogenous promoter, further comprises a 3' polyadenylation and/or transcript termination region. The term "3' polyadenylation and transcript termination region" refers a 3' end region involved in transcription termination or polyadenylation and encompasses a control sequence which is a DNA sequence at the end of a transcriptional unit which signals 3' processing or polyadenylation of a primary transcript and is involved in termination of transcription. The control sequence for transcription termination or terminator can be derived from a natural gene or from a variety of genes. For expression in plants, the 3' end sequence to be added may be derived from, for example, the nopaline synthase or

octopine synthase genes, or alternatively from another plant gene, or less preferably from any other eukaryotic or viral gene. For expression in yeast the terminator to be added may be derived from, for example, the TEF or CYC1 genes or alternatively from another yeast gene or less preferably from any other eukaryotic or viral gene.

A vector comprising a nucleic acid molecule encoding one of the MYC2 or MYC2a proteins disclosed in this specification or a chimeric gene comprising said nucleic acid molecule also forms part of the present invention. Therefore, in a particular embodiment, a vector is provided comprising a nucleic acid molecule encoding any of the MYC2 or MYC2a proteins described in the first aspect of this specification or comprising a chimeric gene comprising said nucleic acid molecule.

The term "vector" as used herein is intended to refer to a nucleic acid molecule capable of transporting another nucleic acid molecule to which it has been linked. The vector may be of any suitable type including, but not limited to, a phage, virus, plasmid, phagemid, cosmid, bacmid or even an artificial chromosome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., vectors having an origin of replication which functions in the host cell). Other vectors can be integrated into the genome of a host cell upon introduction into the host cell, and are thereby replicated along with the host genome. In addition to the replication system, there will frequently be (but not necessarily) at least one marker present, which may be useful in one or more hosts, or different markers for individual hosts. The markers may a) code for protection against a biocide, such as antibiotics, toxins, heavy metals, certain sugars or the like; b) provide complementation, by imparting prototrophy to an auxotrophic host: or c) provide a visible phenotype through the production of a novel compound in the plant. Exemplary genes which may be employed include neomycin phosphotransferase (NPTII), hygromycin phosphotransferase (HPT), chloramphenicol acetyltransferase (CAT), nitrilase, and the gentamicin resistance gene. For plant host selection, non-limiting examples of suitable markers are β-glucuronidase, providing indigo production, luciferase, providing visible light production, Green Fluorescent Protein and variants thereof, NPTII, providing kanamycin resistance or G418 resistance, HPT, providing hygromycin resistance, and the mutated aroA gene, providing glyphosate resistance.

Additionally, certain preferred vectors are capable of directing the expression of certain genes of interest. Such vectors are referred to herein as "recombinant expression vectors" (or simply, "expression vectors"). Suitable vectors have regulatory sequences, such as promoters, enhancers, terminator sequences, and the like as desired and according to a particular host organism (e.g. plant cell). Typically, a recombinant vector according to the present invention comprises at least one “chimeric gene” or “expression cassette”. Expression cassettes are generally DNA constructs preferably including (5’ to 3’ in the direction of transcription): a promoter region, a polynucleotide sequence, homologue, variant or

fragment thereof of the present invention operably linked with the transcription initiation region, and a termination sequence including a stop signal for RNA polymerase and a polyadenylation signal. It is understood that all of these regions should be capable of operating in biological cells, such as plant cells, to be transformed. The promoter region comprising the transcription initiation region, which preferably includes the RNA polymerase binding site, and the polyadenylation signal may be native to the biological cell to be transformed or may be derived from an alternative source, where the region is functional in the biological cell.

Additionally, also host cells comprising one of the MYC2 or MYC2a proteins disclosed in this specification or comprising a nucleic acid molecule encoding one of said MYC2 or MYC2a proteins or comprising a chimeric gene comprising said nucleic acid molecule or comprising a vector comprising said nucleic acid molecule or said chimeric gene form part of the present invention. Therefore, in other particular embodiments, host cells are provided comprising any of the MYC2 or MYC2a proteins described in the first aspect of this specification of comprising a nucleic acid sequence encoding said MYC2 or MYC2a protein.

In particular embodiments, said host cell is a eukaryotic cell. Eukaryotic cells provided in this application can be of any unicellular or multicellular eukaryotic organism, but in particular embodiments yeast, plant, and algal cells are envisaged, in most particular embodiments, plant cells are envisaged. The nature of the cells used will typically depend on the ease and cost of producing monoterpenoid indole alkaloids (MIAs). It is clear for the skilled person that when using non-plant host cells the necessary molecular framework to produce plant-derived MIAs should be introduced or engineered in said non-plant host cells. Therefore, in a preferred embodiment, said host cell is a plant cell. The term “plant” as used herein refers to vascular plants (e.g. gymnosperms and angiosperms). A “plant cell” is understood, according to the invention, as being any cell which is derived from or found in a plant and which is able to form or is part of undifferentiated tissues, such as calli or cell cultures, differentiated tissues such as embryos, parts of plants, plants or seeds.

Given the conservation of the MIA pathway organization, there is no reason to assume that the MYC2 or MYC2a proteins described in this specification or the triple transcription factor cassette (see later) would not work well in other MIA-producing plants, for example to stimulate the production of valuable compounds such as reserpine in *Rauwolfia serpentina*, camptothecin in *Camptotheca acuminata* or quinine in *Cinchona officinalis* (Aerts et al., 1994; Deepthi and Satheeshkumar, 2017; Gundlach et al., 1992; Song and Byun, 1998). Plants that are particularly suitable for the production of MIAs are plants from the genus selected from the group consisting of *Catharanthus*, *Medicago*, *Rauwolfia*, *Camptothecal*, *Cinchona*, *Arabidopsis* and *Nicotiana*.

Plants or plant cells that are also relevant to this application include, but are not limited to, plants or plant cells of agronomically important crops which are or are not intended for animal or human nutrition, such as maize or corn, wheat, barley, oat, *Brassica* spp. plants such as *Brassica napus* or *Brassica juncea*, soybean, bean, alfalfa, pea, rice, sugarcane, beetroot, tobacco, sunflower, cotton, vegetable plants such as cucumber, leek, carrot, tomato, lettuce, peppers, melon, watermelon, diverse herbs such as oregano, basilicum and mint. It may also be applied to plants that produce valuable compounds, e.g. useful as for instance pharmaceuticals, as ajmalicine, vinblastine, vincristine, ajmaline, reserpine, rescinnamine, camptothecine, ellipticine, quinine, and quinidine, taxol, morphine, scopolamine, atropine, cocaine, sanguinarine, codeine, genistein, daidzein, digoxin, calystegins or as food additives such as anthocyanins, vanillin; including but not limited to the classes of compounds mentioned above. Examples of such plants include, but not limited to, *Papaver spp.*, *Rauwolfia spp.*, *Taxus spp.*, *Cinchona spp.*, *Eschscholtzia californica*, *Camptotheca acuminata*, *Hyoscyamus spp.*, *Berberis spp.*, *Coptis spp.*, *Datura spp.*, *Atropa spp.*, *Thalictrum spp.*, *Peganum spp*. Preferred members of the genus *Taxus* comprise *Taxus brevifolia*, *Taxus baccata*, *Taxus cuspidata*, *Taxus canadensis* and *Taxus floridana*.

According to alternative embodiments, the eukaryotic cells as used to produce the MIAs are not plant cells but for example algal or yeast cells. According to particular embodiments, the yeast cell is a yeast cell of a *Saccharomyces* species (e.g. *Saccharomyces cerevisiae*), a *Hansenula* species (e.g. *Hansenula polymorpha*), a *Yarrowia* species (e.g. *Yarrowia lipolytica*), a *Kluyveromyces* species (e.g. *Kluyveromyces lactis*), a *Pichia* species (e.g. *Pichia pastoris*) or a *Candida* species (e.g. *Candida utilis*). According to further particular embodiments, the algal cells are derived from algae of the genus selected from the group comprising *Dunaliella, Chlorella,* or *Chlamydomonas.* In other alternative embodiments, said host cell is a microbial cell. The microbial cells particularly envisioned in current application are bacterial cells and yeast cells.

In another alternative embodiment, the prokaryotic or eukaryotic host cells as used herein themselves do not naturally produce the MIAs of the application, but may do so after genetic engineering. Thus, preferably, host cells artificially producing MIAs of the application refers to cells that, while not naturally having the ability to synthesize the MIAs of the application, have acquired such ability by means of genetic modification processes including transgenesis. Therefore, in another particular embodiment, said host cell is a recombinant host cell. The term "recombinant host cell" ("expression host cell", "expression host system", "expression system" or simply "host cell"), as used herein, is intended to refer to a cell into which a recombinant vector has been introduced. It should be understood that such terms are intended to refer not only to the particular subject cell but to the progeny of such a cell. Because

certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term “host cell” as used herein. A recombinant host cell may be an isolated cell or cell line grown in culture or may be a cell which resides in a living tissue or organism. Host cells can be of bacterial, fungal, plant or mammalian origin.

Thus, in case the host cell is a plant cell then the plant cell as used may be a genetically engineered plant cell, which is a plant cell derived from a recombinant or genetically engineered plant. A "recombinant plant" or a “genetically engineered plant”, as used herein, refers to a plant comprising a recombinant polynucleotide and/or a recombinant polypeptide resulting in the expression of one of the MYC2 or MYC2a proteins as disclosed in the first aspect of this specification. In a particular embodiment, a recombinant plant cell is provided comprising one of the MYC2 or MYC2a proteins described in the first aspect of this specification or expressing a nucleic acid molecule encoding one of said MYC2 or MYC2a proteins.

A recombinant plant refers to a whole plant as well as to a plant part, such as seed, fruit, leaf, or root, plant tissue, plant cells or any other plant material, and progeny thereof. A recombinant plant can be obtained by transforming a plant cell with an expression cassette and regenerating such plant cell into a transgenic, cisgenic or intragenic plant. Such plants can be propagated vegetatively or reproductively. The transforming step may be carried out by any suitable means, including by *Agrobacterium*-mediated transformation and non-*Agrobacterium*-mediated transformation, as discussed further below. Plants can be regenerated from the transformed cell (or cells) by techniques known to those skilled in the art. Where chimeric plants are produced by the process, plants in which all cells are transformed may be regenerated from chimeric plants having transformed germ cells, as is known in the art. Methods that can be used to transform plant cells or tissue with expression vectors include both *Agrobacterium* and non-*Agrobacterium* vectors. *Agrobacterium*-mediated gene transfer exploits the natural ability of *Agrobacterium tumefaciens* to transfer DNA into plant chromosomes and is described in detail in Gheysen, G., Angenon, G. and Van Montagu, M. 1998. *Agrobacterium*-mediated plant transformation: a scientifically intriguing story with significant applications. In K. Lindsey (Ed.), Transgenic Plant Research. Harwood Academic Publishers, Amsterdam, pp. 1-33 and in Stafford, H.A. (2000) Botanical Review 66: 99-118. A second group of transformation methods is the non-*Agrobacterium* mediated transformation and these methods are known as direct gene transfer methods. An overview is brought by Barcelo, P. and Lazzeri, P.A. (1998) Direct gene transfer: chemical, electrical and physical methods. In K. Lindsey (Ed.), Transgenic Plant Research, Harwood Academic Publishers, Amsterdam, pp.35-55. Methods include particle gun delivery, microinjection, electroporation of intact cells, polyethyleneglycol-mediated protoplast transformation, electroporation of protoplasts, liposome-mediated transformation, silicon-

whiskers mediated transformation etc. A suitable control plant would include a genetically unaltered or non-transgenic plant of the parental line (wild type) used to generate a transgenic plant herein.

Particularly useful for current application are genetically transformed hairy root cultures that can be obtained by transformation with virulent strains of *Agrobacterium rhizogenes*, and can produce high levels of MIAs of current application. Protocols used for establishing of hairy root cultures vary, as well as the susceptibility of plant species to infection by *Agrobacterium* (Toivounen et al. 1993; Vanhala et al. 1995). It is possible by systematic clone selection e.g. via protoplasts, to find high yielding, stable, and from single cell derived-hairy root clones. This is possible because the hairy root cultures possess a great somaclonal variation. Another possibility of transformation is the use of viral vectors (Turpen 1999). Any plant tissue or plant cell capable of subsequent clonal propagation, whether by organogenesis or embryogenesis, may be transformed with an expression vector of interest. The term 'organogenesis' means a process by which shoots and roots are developed sequentially from meristematic centers; the term 'embryogenesis' means a process by which shoots and roots develop together in a concerted fashion (not sequentially), whether from somatic cells or gametes. The particular tissue chosen will vary depending on the clonal propagation systems available for, and best suited to, the particular species being transformed. Exemplary tissue targets include protoplasts, leaf disks, pollen, embryos, cotyledons, hypocotyls, megagametophytes, callus tissue, existing meristematic tissue (e.g. apical meristems, axillary buds, and root meristems), and induced meristem tissue (e.g., cotyledon meristem and hypocotyls meristem). Also useful for the industrial production of MIAs are whole transgenic plants. MIAs can then be extracted or isolated from the whole plant or from easily accessible tissue, e.g. leaves. Also particularly useful for current application are cell suspension cultures of plant cells expressing one of the MYC2 or MYC2a proteins disclosed in the first aspect of this specification or expressing the triple transcription factor cassette (see later) of current application. Suitable cell culture media for eukaryotic cells, in particular plant cells and microbial cells, are known in the art. For plant cells, a number of suitable culture media for callus induction and subsequent growth on aqueous or solidified media are known. Exemplary media include standard growth media, many of which are commercially available (e.g., Sigma Chemical Co., St. Louis, Mo.). Examples include Schenk-Hildebrandt (SH) medium, Linsmaier-Skoog (LS) medium, Murashige and Skoog (MS) medium, Gamborg's B5 medium, Nitsch & Nitsch medium, White's medium, and other variations and supplements well known to those of skill in the art (see, e.g., Plant Cell Culture, Dixon, ed. IRL Press, Ltd. Oxford (1985) and George et al., Plant Culture Media, Vol 1, Formulations and Uses Exegetics Ltd. Wilts, UK, (1987)). (see, e.g., Plant Cell Culture, Dixon, ed. IRL Press, Ltd. Oxford (1985) and George et al., Plant Culture Media, Vol 1, Formulations and Uses Exegetics Ltd. Wilts, UK, (1987)).

For yeast cells, exemplary media include standard growth media, many of which are commercially available (e.g., Clontech, Sigma Chemical Co., St. Louis, Mo.). Examples include Yeast Extract Peptone Dextrose (YPD or YPED) medium, Yeast Extract Peptone Glycerol (YPG or YPEG) medium, Hartwell's complete (HC) medium, Synthetic complete (SC) medium, Yeast Nitrogen Base (YNB), and other variations and supplements well known to those of skill in the art (see, Yeast Protocol Handbook, Clontech). The incubation conditions (temperature, photoperiod, shaking, auxin/cytokine hormone ratio, promoter inducing conditions, promoter repressing conditions, etc.) will depend, among other factors, on the cells to be incubated and are standard techniques in the art.

In a third aspect, a plant cell is provided comprising one of the MYC2 or MYC2a proteins disclosed in the first aspect of this specification or expressing a nucleic acid molecule encoding one of said MYC2 or MYC2a proteins, wherein said plant cell further expresses the transcriptional activators BIS1 and/or ORCA3. In a particular embodiment, said transcriptional activators BIS1 and/or ORCA3 are transcriptionally controlled by promoters different to the promoters to which they are naturally associated or linked. In a particular embodiment, said transcriptional activators BIS1 and/or ORCA3 are transcriptionally controlled by exogenous promoters.

In one embodiment, said transcriptional activator BIS1 has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% homology to SEQ ID No. 18. In another embodiment, said transcriptional activator ORCA3 has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 93% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 95% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 97% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 99% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 is defined by SEQ ID No. 22.

In yet another embodiment, said transcriptional activator BIS1 has at least 93% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 95% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 97% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 99% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 is defined by SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22.

In yet another embodiment, said transcriptional activator BIS1 has at least 93% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 93% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 95% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 95% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 97% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 97% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 99% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 99% homology to SEQ ID No. 22.

In yet another embodiment, said transcriptional activator BIS1 is defined by SEQ ID No. 18 and said transcriptional activator ORCA3 is defined by SEQ ID No.22.

In a particular embodiment, the plant cell of the third aspect of this specification is a recombinant or genetically engineered plant cell. According to a particularly embodiment, the plant cell of the third aspect of this specification comprises or expresses an exogenous or wholly exogenous MYC2, BIS1 and ORCA3. The term “exogenous” or “heterologous” as used herein is any material originated outside of an organism, tissue, or cell, but that is present (and typically can become active) in that organism, tissue, or cell. Analogously, “endogenous”, refers to substances (e.g. genes) originating from within an organism, tissue, or cell. The term “wholly exogenous” as used herein refers to a wholly exogenous protein and means that the whole protein (especially relevant in case of fusion proteins) is exogenous. For example, a plant protein endogenous to a specific plant species will become exogenous after fusion of said endogenous protein with an exogenous protein, for example but without the purpose of limiting the green fluorescent protein (GFP). The term “wholly exogenous” is used when every functional fragment of a fusion protein is exogenous to a specific host.

In particular embodiments, a plant cell is provided comprising one of the MYC2 or MYC2a proteins disclosed in the first aspect of this specification or expressing a nucleic acid molecule encoding one of said MYC2 or MYC2a proteins, wherein said plant cell further expresses an orthologue of the *Catharanthus roseus* transcriptional activator BIS1 and/or an orthologue of the *Catharanthus roseus* transcriptional activator ORCA3. In a particular embodiment, said orthologues of the transcriptional activators BIS1 and/or ORCA3 are transcriptionally controlled by promoters different to the promoters to which they are naturally associated or linked. In a particular embodiment, said orthologues of the transcriptional activators BIS1 and/or ORCA3 are transcriptionally controlled by exogenous promoters.

Orthologues of BIS1 that are particularly envisaged are the *Medicago truncatula* TRITERPENE SAPONIN BIOSYNTHESIS ACTIVATING REGULATOR1 (TSAR1) and TSAR2 proteins (see also Mertens et al 2016 Plant Physiol 170:194-210). In particular embodiments, said orthologue of the *Catharanthus roseus* transcriptional activator BIS1 is MtTSAR1 defined by SEQ ID No. 29 or MtTSAR2 defined by SEQ ID No. 30.

Orthologues of ORCA3 can be identified by the presence of one or more conserved domains. In one embodiment, said orthologue of the *Catharanthus roseus* transcriptional activator ORCA3 comprises one AP2/ERF domain as defined by SEQ ID No. 24. In other embodiments, said ORCA3 orthologue further comprises a C-terminal domain as defined by SEQ ID No. 25 and/or a nuclear localization sequence as defined by SEQ ID No. 26 and/or an acidic domain as defined by SEQ ID No. 27 and/or a serine-rich domain as defined by SEQ ID No. 28.

In another particular embodiment, the plant cell of the third aspect of this specification is provided wherein BIS1 or an orthologue thereof and ORCA3 or an orthologue thereof are overexpressed. Overexpression of a gene of interest can be achieved by adding additional copies of said gene in the host cell of interest. The additional copies of the gene of interest (e.g. BIS1 and/or ORCA3) can be transcriptionally regulated by their endogenous promoter but more preferably by a non-endogenous promoter. The result of said overexpression is that the expression level of a gene of interest exceeds the expression level of said gene in a control host cell. Therefore, this application also provides a plant cell comprising one of the MYC2 or MYC2a proteins disclosed in the first aspect of this specification or expressing a nucleic acid molecule encoding one of said MYC2 or MYC2a proteins, wherein said plant cell further expresses the transcriptional activators BIS1 or an orthologue thereof and/or ORCA3 or an orthologue thereof, wherein the expression levels of BIS1 and/or ORCA3 or said orthologues thereof are at least 10% higher, at least 20% higher, at least 30% higher, at least 40% higher, at least 50% higher, at least 60% higher, at least 70% higher, at least 80% higher, at least 90% higher, at least 100% higher, at least 2-fold higher, at least 3-fold higher, at least 5-fold higher, at least 10-fold higher, between 30% and

150% higher, between 50% and 200% higher or between 2-fold and 8-fold higher than the corresponding expression level of BIS1 and/or ORCA3 of said orthologues thereof in a control plant cell.

In alternative embodiments of the third aspect, a plant cell is provided comprising one of the MYC2 or MYC2a proteins disclosed in the first aspect of this specification or expressing a nucleic acid molecule encoding one of said MYC2 or MYC2a proteins, wherein said plant cell further expresses the transcriptional activators BIS2 and ORCA3 or BIS2 and ORCA2 or BIS1 and ORCA2. In a particular embodiment, said transcriptional activators BIS1 and/or BIS2 and/or ORCA2 and/or ORCA3 are transcriptionally controlled by promoters different to the promoters to which they are naturally associated or linked. In a particular embodiment, said transcriptional activators BIS1 and/or BIS2 and/or ORCA2 and/or ORCA3 are transcriptionally controlled by exogenous promoters. BIS2 is a homologue of BIS1 in *C. roseus* and defined by SEQ ID No. 19, while ORCA2 is a homologue of ORCA3 in *C. roseus* and defined by SEQ ID No. 31.

In a fourth aspect, a regulatory module is provided comprising one of the MYC2 or MYC2a proteins disclosed in the first aspect of this specification and further comprising the transcription factors BIS1 and/or ORCA3, wherein said BIS1 and ORCA3 are transcriptionally controlled by promoters different to the promoters to which BIS1 and ORCA3 are naturally linked or associated.

A “regulatory module” as used herein refers to a combination of at least two transcription factors controlling the expression of one or more genes within one molecular pathway, for example the pathway that leads to the production of monoterpenoid indole alkaloids. The at least two transcription factors can be introduced into a host cell using one or more vectors. Therefore, in one embodiment, a vector is provided comprising a nucleic acid molecule encoding one of the MYC2 or MYC2a proteins described in the first aspect of this specification and a nucleic acid molecule encoding BIS1. In another embodiment, a vector is provided comprising a nucleic acid molecule encoding one of the MYC2 or MYC2a proteins described in the first aspect of this specification and a nucleic acid molecule encoding ORCA3. In yet another embodiment, a vector is provided comprising a nucleic acid molecule encoding one of the MYC2 or MYC2a proteins described in the first aspect of this specification, a nucleic acid molecule encoding BIS1 and a nucleic acid molecule encoding ORCA3.

In a particular embodiment, said regulatory module is a triple transcription factor cassette, consisting of one of the MYC2 or MYC2a proteins disclosed in the first aspect of this specification and the transcription factors BIS1 and ORCA3. Said triple transcription factor cassette can consist of one vector, two vector or three vectors.

In one embodiment, said transcriptional activator BIS1 has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% homology to SEQ ID No. 18. In another embodiment, said transcriptional activator ORCA3 has at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99% or 100% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 93% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 95% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 97% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 99% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 90% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 is defined by SEQ ID No. 22.

In yet another embodiment, said transcriptional activator BIS1 has at least 93% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 95% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 97% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 99% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 is defined by SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 90% homology to SEQ ID No. 22.

In yet another embodiment, said transcriptional activator BIS1 has at least 93% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 93% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 95% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 95% homology to SEQ ID No. 22. In yet another embodiment, said transcriptional activator BIS1 has at least 97% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 97% homology to SEQ ID No. 22. In yet another embodiment, said

transcriptional activator BIS1 has at least 99% homology to SEQ ID No. 18 and said transcriptional activator ORCA3 has at least 99% homology to SEQ ID No. 22.

In yet another embodiment, said transcriptional activator BIS1 is defined by SEQ ID No. 18 and said transcriptional activator ORCA3 is defined by SEQ ID No. 22.

In a fifth aspect, the use of one of the MYC2 or MYC2a proteins described in the first aspect of this specification or of a nucleic acid molecule encoding one of said MYC2 or MYC2a proteins or of the above described regulatory module is provided to produce at least one monoterpenoid indole alkaloid in a plant cell. Also the use of one of the plant cells described in the third aspect of this specification is provided for the production of at least one MIA. In one embodiment, the use of one of the MYC2 or MYC2a proteins described in the first aspect of this specification or of a nucleic acid molecule encoding one of said MYC2 or MYC2a proteins or of the above described regulatory module is provided to enhance the production of at least one monoterpenoid indole alkaloid in a plant cell with at least 10% compared to a control plant cell. In more particular embodiments, said use is provided to enhance the production of at least one monoterpenoid indole alkaloids in a plant cell with at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 2-fold, at least 3-fold, at least 5-fold, at least 7-fold, at least 10-fold, at least 50-fold, at least 100-fold, between 2-fold and 10-fold or between 5-fold and 100-fold or between 50-fold and 1000-fold compared to a control plant cell. For this aspect of the application, said control plant cell is a plant cell without one of said MYC2 or MYC2a proteins or without the nucleic acid molecule encoding one of said MYC2 or MYC2a proteins or without the above described regulatory module.

In other embodiments, the use of one of the MYC2 or MYC2a proteins described in the first aspect of this specification or of a nucleic acid molecule encoding one of said MYC2 or MYC2a proteins is provided in combination with the expression of the *C. roseus* BIS1 or an orthologue thereof and/or the *C. roseus* ORCA3 or an orthologue thereof to produce at least one MIA in a host cell or the increase the production of at least one MIA in a plant cell.

In alternative embodiments, the use of one of the MYC2 or MYC2a proteins described in the first aspect of this specification or of a nucleic acid molecule encoding one of said MYC2 or MYC2a proteins is provided in combination with the expression of the *C. roseus* BIS2 and ORCA3 or of the *C. roseus* BIS2 and ORCA2 or of the *C. roseus* BIS1 and ORCA2 to produce at least one MIA in a host cell or the increase the production of at least one MIA in a plant cell. Said increase in production is defined as at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 2-fold, at least 3-fold, at least 5-fold, at least 7-fold, at least 10-fold, at least 50-fold, at least 100-fold, between 2-fold and 10-fold or between 5-fold and 100-fold or between 50-fold and 1000-

fold higher compared to a control plant cell. In particular embodiments, BIS1, BIS2, ORCA2, ORCA3 and their orthologues are defined as described in the third aspect of this specification.

In a particular embodiment, said use of the fifth aspect is provided, wherein the at least one monoterpenoid indole alkaloid is a 16-hydroxytabersonine-derived MIA and/or 19-hydroxytabersonine-derived MIA.

In another particular embodiment, said use is provided wherein the at least one monoterpenoid indole alkaloids is selected from the list consisting of strictosidine, catharanthine, 16-hydroxytabersonine, 16-methoxylhörhammericine, hörhammericine, serpentine, vindoline, vinblastine, akuammicine, isositsirikine, 16-methoxytabersonine, 3-hydroxy-16-methoxy-2,3-dihydrotabersonine, 19-hydroxytabersonine, 16-hydroxyhörhammericine, minovincinine, vincadifformine, 16-hydroxyvincadifformine, 16-hydroxy-19-O-acetyltabersonine, perivine and O-acetylstemmadine.

In a sixth aspect, a method is provided of producing at least one monoterpenoid indole alkaloid in a host cell, comprising:

- expressing one of the MYC2 or MYC2a proteins described in the first aspect of this specification in a host cell;
- optionally further expressing the *C. roseus* BIS1 transcription factor or an orthologue thereof and/or the *C. roseus* ORCA3 transcription factor or an orthologue thereof;
- selecting host cells that produce at least one monoterpenoid indole alkaloid.

In particular embodiments, a method is provided of increasing the production of at least one monoterpenoid indole alkaloid in a host cell, comprising:

- expressing one of the MYC2 or MYC2a proteins described in the first aspect of this specification in a host cell;
- optionally further expressing the *C. roseus* BIS1 transcription factor or an orthologue thereof and/or the *C. roseus* ORCA3 transcription factor or an orthologue thereof;
- selecting host cells with increased production of at least one monoterpenoid indole alkaloid compared to control cells not expressing one of said MYC2 of MYC2a proteins.

In alternative embodiments, methods are provided of producing at least one monoterpenoid indole alkaloid in a host cell, said method comprises expressing one of the MYC2 or MYC2a proteins described in the first aspect of this specification in a host cell, expressing the *C. roseus* BIS1 transcription factor or an orthologue thereof and/or the *C. roseus* BIS2 transcription factor or an orthologue thereof and/or the *C. roseus* ORCA2 transcription factor or an orthologue thereof and/or the *C. roseus* ORCA3 transcription factor or an orthologue thereof, and selecting host cells that produce at least one monoterpenoid indole alkaloid. In other alternative embodiments, method are provided of increasing the production of at least

one monoterpenoid indole alkaloid in a host cell, said method comprises expressing one of the MYC2 or MYC2a proteins described in the first aspect of this specification in a host cell, expressing the *C. roseus* BIS1 transcription factor or an orthologue thereof and/or the *C. roseus* BIS2 transcription factor or an orthologue thereof and/or the *C. roseus* ORCA2 transcription factor or an orthologue thereof and/or the *C. roseus* ORCA3 transcription factor or an orthologue thereof, and selecting host cells with increased production of at least one monoterpenoid indole alkaloid compared to control cells not expressing one of said MYC2 of MYC2a proteins.

In particular embodiments, said host cell is a eukaryotic cell, more particularly a plant or yeast cell, even more particularly a plant cell.

In more particular embodiments, increased production of at least one MIA means that the production level of said at least one MIA in the selected host cells is at least 10% higher than that of control cells. In even more particular embodiments, said production level is at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 2-fold, at least 5-fold, at least 10-fold, at least 50-fold, at least 100-fold, or between 2-fold and 10-fold, between 10-fold and 100-fold or between 50-fold and 1000-fold higher than that of control cells.

In particular embodiments, BIS1, BIS2, ORCA2, ORCA3 and their orthologues are defined as described in the third aspect of this specification.

The following examples are intended to promote a further understanding of the present invention. While the present invention is described herein with reference to illustrated embodiments, it should be understood that the invention is not limited hereto. Those having ordinary skill in the art and access to the teachings herein will recognize additional modifications and embodiments within the scope thereof. Therefore, the present invention is limited only by the claims attached herein.

EXAMPLES

Example 1. A protoplast-based screen for transcriptional regulator(s) of BIS1 yields no hits from a collection of 'wild-type' *C. roseus* TFs

Positive amplification loops are a hallmark of the JA-signaling pathway (Memelink, 2009; Pauwels et al., 2009). As such, the expression of the BIS as well as the ORCA and most MIA pathway genes can be induced by JA treatment (data not shown) (Giddings et al., 2011; Liscombe et al., 2010; Miettinen et al., 2014; Van Moerkercke et al., 2016; Van Moerkercke et al., 2015). However, thus far we could not establish an auto-amplification loop for the BIS genes (Van Moerkercke et al., 2016). Furthermore, a direct physical interaction between the BIS TFs and the JAZ repressors is unlikely, because the BIS TFs lack a typical JID domain. Nevertheless, we mined the *C. roseus* transcriptomes for *CrJAZ* genes, cloned

several of them and carried out a yeast two-hybrid (Y2H) assay to test for potential interaction between BIS TFs and JAZ. As recently also suggested (Patra et al., 2018), BIS TFs failed to interact with all tested JAZ proteins (Fig. 2A). This suggests that the JA-elicitation of BIS expression is mediated by another TF that is repressed by JAZ proteins. Such a regulatory loop has been established for the JA-induction of ORCA genes, and as a consequence the downstream epidermis-specific MIA biosynthesis genes, with MYC2 being the responsible JAZ-interacting regulatory TF (Fig. 1A) (Zhang et al., 2011; Zhou and Memelink, 2016). To investigate whether the same or a similar cascade is driving the expression of BIS genes and consequently that of the IPAP-specific iridoid pathway, we tested the transactivation of the BIS1 promoter (pBIS1) by JA-modulated, potentially JAZ-interacting *C. roseus* TFs in our well-established tobacco protoplast-based expression system that we previously used to discover the BIS TFs (Van Moerkercke et al., 2016; Van Moerkercke et al., 2015; Vanden Bossche et al., 2013). CrMYC2 (i.e. CrMYC2a, see later), as well as the Arabidopsis homolog AtMYC2, are both able to transactivate the ORCA3 promoter (pORCA3) but both MYC2 orthologs failed to transactivate pBIS1 (Fig. 3A), suggesting that CrMYC2 may not be directly regulating BIS1 expression.

The past decade, numerous JAZ-interacting TFs have been identified, mainly in Arabidopsis, including bHLH TFs from the same or other bHLH clades than MYC2, namely clades IIIb (the 'ICE' TFs), IIId (the 'JAM' TFs), IIIe (the 'MYC2' TFs) and IIIf (the 'GL' TFs) (Chini et al., 2016; De Geyter et al., 2012; Goossens et al., 2016; Heim et al., 2003; Kazan and Manners, 2012; Pauwels and Goossens, 2011). We therefore screened the available *C. roseus* transcriptomes for members of these bHLH clades using the Arabidopsis proteins as a bait and cloned them for expression in the protoplast assay. Besides homologs of well-described TFs such as GLABROUS 3 (GL3), JASMONATE ASSOCIATED MYC2-LIKE 1 (JAM1), and INDUCER OF CBP EXPRESSION 1 (ICE1), we also identified two additional *C. roseus* MYC2 homologs that we named CrMYC2b and CrMYC2c (Fig. 4). The 'original' CrMYC2 was therefore renamed to CrMYC2a here. This nomenclature was chosen to avoid any confusion with previously described bHLH TFs from *C. roseus*, which are phylogenetically not closely related to CrMYC2a, despite being called CrMYC1, CrMYC3, CrMYC4 and CrMYC5 (Chatel et al., 2003; Zhang et al., 2011) (Fig. 4). Expression of CrMYC2a, CrMYC5, CrJAM2 and CrJAM3 but not of CrMYC1, CrMYC3, CrMYC4, CrICE1a, CrICE1b and CrGL3 was JA-inducible in plants (data not shown). However, none of the clade III bHLH TFs could transactivate pBIS1 in the protoplast assay (data not shown).

Besides the clade III bHLH TF BIS, and the ORCA TF-encoding genes, mining of the available *C. roseus* transcriptome data revealed numerous other JA-inducible TFs of the WRKY, MYB and NAC families, several of which were included in our screen. However, again none of the tested TFs could transactivate the pBIS1 fragment in the protoplast assay (data not shown).

Example 2. Contrary to wild-type MYC2 versions and de-repressed mutant CrMYC2b and CrMYC2c, de-repressed mutant CrMYC2a transactivates pBIS1 and pORCA3 in tobacco protoplasts

Because protoplasting is known to trigger JA signaling (Birnbaum et al., 2003), we postulated that endogenous tobacco JAZ proteins may repress the transcriptional activity of bHLH TFs containing JID domains such as CrMYC2a-c, possibly leading to false negatives. To explore this hypothesis, we tested whether de-repressed CrMYC2 mutants could transactivate pBIS1. In Arabidopsis, in contrast to wild-type AtMYC2, the available de-repressed AtMYC2D105N mutant (Goossens et al., 2015) was capable to transactivate pBIS1 around eightfold (Fig. 3A), suggesting it is involved in direct regulation of BIS1 expression and that its wild-type activity is limited by interaction with JAZ proteins. Accordingly, transactivation of pORCA3 by AtMYC2D105N was fivefold stronger than by wild-type AtMYC2 (Fig. 3A). The JAZ/MYC2 regulatory complex is widely conserved among the plant kingdom (Chini et al., 2016; Wasternack and Strnad, 2018). Likewise, the negatively charged aspartate (D) at position 105 of AtMYC2 that causes the de-repression phenotype is strictly conserved among all Arabidopsis and *C. roseus* clade IIIe bHLH TFs and corresponds to D126 of CrMYC2a, D60 of CrMYC2b and D53 of CrMyc2c (Fig. 5). To assess whether the same mutation in CrMYC2a-c would also trigger de-repression and transactivation of MIA pathway gene promoters, we engineered a de-repressed CrMYC2a (CrMYC2aD160N), CrMYC2b (CrMYC2bD60N) and CrMYC2c (CrMYC2cD50N). Surprisingly, only CrMYC2aD126N could transactivate pBIS1, pBIS2, pORCA2 and pORCA3 in BY-2 protoplasts (Fig. 3A; Fig. 3D-G), supporting a role for CrMYC2a, but not CrMYC2b or CrMYC2c, as an indirect regulator of MIA biosynthesis gene expression.

To confirm that the mutation in CrMYC2aD126N affects the interaction with JAZ proteins, as it does in AtMYC2D105N, we carried out a Y2H assay (Fig. 3B). CrMYC2a and AtMYC2 both interacted with all tested CrJAZ proteins. The mutant version of CrMYC2a lost its capacity to interact with CrJAZ3 and CrJAZ8, whereas the interaction with CrJAZ1, 2, 4 and 7 was maintained. In comparison, the mutant version of AtMYC2 did no longer interact with CrJAZ1, 2, 3 and 8, whereas interaction with CrJAZ4 and CrJAZ7 was also maintained. Generally, these Y2H results are in agreement with those published in other studies and that have shown a loss of interaction of AtMYC2D105N with most Arabidopsis JAZ proteins as well as a maintained ability to interact with some of them (Frerigmann et al., 2014; Goossens et al., 2015).

Originally, CrMYC2a was discovered in a yeast-one hybrid (Y1H) screen for activators of the STR promoter (pSTR), but has so far not been shown to transactivate pSTR in any plant system (Chatel et al., 2003; Paul et al., 2017; Pré et al., 2000; Zhang et al., 2011). Here, we corroborate these Y1H data by demonstrating that both the de-repressed AtMYC2D105N and CrMYC2aD126N can transactivate pSTR in tobacco protoplasts, contrary to the wild-type versions (Fig. 3A). Together, this also supports a role for CrMYC2a as a direct regulator of MIA biosynthesis gene expression. To assess whether this applies to the whole

MIA pathway, we tested transactivation of the available promoters of other MIA pathway genes. Indeed, CrMYC2aD126N, but not CrMYC2a was capable of transactivating the GES, 8HGO, IS, IO, SLS, TDC and SGD promoters (Fig. 3C).

Example 3. Transient expression of de-repressed AtMYC2 or CrMYC2a induces expression of MIA pathway genes in *C. roseus* flower petals and leads to increased strictosidine accumulation

To assess the potential of de-repressed AtMYC2 and CrMYC2a mutants to stimulate MIA biosynthesis gene expression *in planta*, they were transiently overexpressed in *C. roseus* flower petals using *Agrobacterium tumefaciens* infiltration. In accordance with the tobacco protoplast assays, overexpression of de-repressed AtMYC2 or CrMYC2a mutants resulted in a strong upregulation of MIA pathway genes in comparison to control petals transformed with a *pCaMV35S::GUS* construct (Fig. 6A and data not shown). Overall, overexpression of CrMYC2aD126N and AtMYC2D105N caused similar induction patterns, both in terms of the identity of the MIA pathway genes and the fold-induction, although CrMYC2aD126N had a particularly stronger effect on the induction of the BIS and ORCA genes than AtMYC2D105N. Both de-repressed MYC2 TFs also strongly induced the expression of genes encoding the NPF iridoid and strictosidine transporters (Fig. 6A and data not shown). Unexpectedly, also TABERSONINE 19-HYDROXYLASE (T19H), MINOVINCININE-19-O-ACETYLTRANSFERASE (MAT) and the recently published homolog TABERSONINE DERIVATIVE 19-O-ACETYLTRANSFERASE (TAT), normally specifically expressed in roots and involved in the biosynthesis of root-specific MIAs such as hörhammericine (Giddings et al., 2011; Laflamme et al., 2001), were pronouncedly induced by the de-repressed MYC2 TFs (Fig. 6A and data not shown). So far, no positive regulators have been described for these transporter or root-specific MIA biosynthesis genes. Notably, despite the strong upregulation of pathway genes for specific MIAs such as strictosidine, hörhammericine or ajmalicine, the expression of the vindoline pathway genes, i.e., TABERSONINE 16-HYDROXYLASE 2 (T16H2), TABERSONINE 16-O-METHYLTRANSFERASE (T16OMT), DESACETOXYVINDOLINE 4-HYDROXYLASE (D4H) or DEACETYLVINDOLINE O-ACETYLTRANSFERASE (DAT) was not induced by the overexpression of de-repressed MYC2s (data not shown). Qu et al. (2018) recently published the missing enzymatic steps between geissoschizine and tabersonine or catharanthine, i.e. Redox1, Redox2, SAT, HL1 and HL2 (Fig. 1A). Similar to the vindoline pathway, the genes coding for these newly discovered enzymes were not induced by the overexpression of de-repressed MYC2s (data not shown). This correlates with the observation that those genes are not JA-inducible either; hence, their expression may be subject to other regulatory cues.

Conversely, overexpression of wild-type AtMYC2 had no significant effect on MIA pathway gene expression, whereas that of wild-type CrMYC2a resulted in a transcriptional induction of the *BIS1* and

BIS2 genes, although 6- to 9-fold less than by CrMYC2aD126N overexpression (Fig. 6A and data not shown). Presumably as an indirect effect caused by BIS upregulation, also a minor induction of some IPAP-specific iridoid genes, e.g. of *GES*, *8HGO* and *IS*, was observed by *CrMYC2a* overexpression (Fig. 6A and data not shown). The expression profile of the epidermis-specific late iridoid and MIA pathway genes did not significantly differ between control flowers or those overexpressing wild-type MYC2s, which correlates with the lack of induction of ORCA2 or ORCA3 (Fig. 6 and data not shown).

Next, we carried out metabolite profiling by liquid chromatography Fourier-transform ion cyclotron resonance mass spectrometry (LC-FT-ICR-MS) to verify whether the transcriptional upregulation of the MIA pathway in the de-repressed AtMYC2 or CrMYC2a overexpressing petals leads to an increased accumulation of MIAs. This analysis revealed that strictosidine levels were increased approximately five-fold in flowers overexpressing de-repressed AtMYC2 or CrMYC2a. Levels of other detectable MIAs, such as catharanthine, vindoline and vinblastine, did not change (Fig. 6B). To further evaluate the impact of de-repressed MYC2s on the metabolome, the LC-FT-ICR-MS data were used for a principal component analysis (PCA) (Fig. 6C). The metabolomes of the various petal samples clustered in two distinguishable groups with one group containing the samples from flowers transformed with the pCaMV35S::GUS, pCaMV35S::CrMYC2a and pCaMV35S::AtMYC2 constructs and the other group containing those from flowers transformed with pCaMV35S::CrMYC2aD126N or pCaMV35S::AtMYC2D105N. The PCA loading control confirmed that strictosidine is the most important variable explaining the observed distribution pattern in the PCA (Fig. 6C). Four additional compounds show a comparable distribution pattern to strictosidine and strongly accumulated in pCaMV35S::CrMYC2aD126N and pCaMV35S::AtMYC2D105N expressing flowers (Fig. 6C). Based on accurate mass and MS^n fragmentation data, these peaks were shown to correspond to a strictosidine dimer, strictosidine hexoside, strictosidine secologanoside (a bahenioside A-like metabolite) and a low abundant, unknown metabolite (Fig. 6C).

Taken together, our findings show that overexpression of de-repressed CrMYC2aD126N and AtMYC2D105N upregulates the expression of the genes encoding the known iridoid and MIA regulators, early MIA enzymes and transporters, with the exception of the vindoline and catharanthine branch enzymes, in *C. roseus* petals, which finally leads to a marked increase in the accumulation of strictosidine.

Example 4. Transient overexpression of the CrMYC2a homologs CrMYC2b and CrMYC2c in *C. roseus* flower petals has no effect on the MIA pathway

We transiently overexpressed CrMYC2b and CrMYC2c as well as their de-repressed forms (CrMYC2bD60N and CrMYC2cD53N) in *C. roseus* flower petals. However, none of the overexpression constructs resulted in an upregulation of the vindoline or catharanthine pathway genes, nor any of the

other iridoid or MIA pathway genes as observed with CrMYC2aD126N (Table 1). These data are in line with the above observed difference between the de-repressed CrMYC2 mutants.

	35S:GUS	35S:MYC2a	35S:MYC2b	35S:MYC2c	35S:MYC2a^{D126N}	35S:MYC2b^{D60N}	35S:MYC2c^{D53N}
MYC2a	1 ± 0.1	11.87 ± 1.88**	1.38 ± 0.06*	0.79 ± 0.04	11.77 ± 0.75***	1.22 ± 0.38	0.8 ± 0.02
MYC2b	1 ± 0.15	3.17 ± 0.73*	1577.4 ± 148.31***	1.43 ± 0.37	8.27 ± 1.13***	717.5 ± 50.49***	1.18 ± 0.38
MYC2c	1 ± 0.28	2.15 ± 0.6	1.43 ± 0.23	135.66 ± 12.49***	1.77 ± 0.08*	1.25 ± 0.08	204.23 ± 12.89***
GES	1 ± 0.39	27.91 ± 10.92*	4.14 ± 2.07	4.32 ± 1.68	445.75 ± 131.05**	2.39 ± 0.5	4.51 ± 2.1
G8O	1 ± 0.05	2.56 ± 0.45**	0.95 ± 0.21	0.99 ± 0.15	34.7 ± 11.32*	1.01 ± 0.07	1.2 ± 0.16
8HGO	1 ± 0.07	5.59 ± 0.31***	0.7 ± 0.26	1.43 ± 0.11*	42.36 ± 11.72*	2.61 ± 1.49	1.11 ± 0.29
IS	1 ± 0.72	12.37 ± 3.19*	0.59 ± 0.01	6.05 ± 3.9	142.58 ± 46.04*	0.28 ± 0.03	1.89 ± 1.31
IO	1 ± 0.24	2.16 ± 0.66	1.96 ± 0.73	1.45 ± 0.28	13.96 ± 6.29	0.46 ± 0.03	0.88 ± 0.23
7DLGT	1 ± 0.09	1.14 ± 0.28	0.91 ± 0.03	0.9 ± 0.11	4.84 ± 0.67**	0.55 ± 0.13*	0.69 ± 0.08*
7DLH	1 ± 0.11	0.77 ± 0.26	1.39 ± 0.1	1.01 ± 0.14	0.99 ± 0.04	0.98 ± 0.04	0.71 ± 0.08
LAMT	1 ± 0.14	3.28 ± 0.43**	3.8 ± 0.56**	1.67 ± 0.11*	30.04 ± 3.09***	1.12 ± 0.19	1.34 ± 0.71
SLS	1 ± 0.01	1.45 ± 0.1**	2.47 ± 0.06***	1.13 ± 0.15	2.42 ± 0.08***	1.95 ± 0.3*	0.78 ± 0.36
TDC	1 ± 0.09	1.97 ± 0.12***	0.94 ± 0.11	0.89 ± 0.06	5.42 ± 0.84**	0.62 ± 0.18	0.71 ± 0.08
STR	1 ± 0.05	2.82 ± 0.48**	0.95 ± 0.04	1.29 ± 0.12	2.86 ± 0.29***	1.59 ± 0.41	1.06 ± 0.24
SGD	1 ± 0.08	3.11 ± 0.67*	2.41 ± 0.64	3.28 ± 0.5**	8.5 ± 1.24***	1.46 ± 0.42	1.04 ± 0.25
HYS	1 ± 0.12	1.81 ± 0.34	3.99 ± 0.86*	1.24 ± 0.19	9.95 ± 1.37***	0.88 ± 0.29	0.57 ± 0.02*
GS1	1 ± 0.02	1.08 ± 0.02*	1.06 ± 0.04	1 ± 0.04	2 ± 0.16**	0.99 ± 0.07	0.99 ± 0.03
GS2	1 ± 0.03	1.74 ± 0.14**	1.72 ± 0.16*	1.14 ± 0.09	7.66 ± 0.88**	1.3 ± 0.18	1.05 ± 0.02
GO	1 ± 0.08	2.55 ± 0.34*	1.35 ± 0.24	1.21 ± 0.06	7.98 ± 0.65***	0.99 ± 0.12	0.96 ± 0.17
Redox1	1 ± 0.08	1.71 ± 0.11**	1.08 ± 0.05	1.12 ± 0.12	2.3 ± 0.3*	0.95 ± 0.05	1.04 ± 0.07
Redox2	1 ± 0.12	1.28 ± 0.14	1.06 ± 0.13	1.32 ± 0.11	1.12 ± 0.08	1.19 ± 0.05	1.4 ± 0.12
SAT	1 ± 0.08	1.13 ± 0.11	0.84 ± 0.06	1.04 ± 0.11	1.14 ± 0.06	0.95 ± 0.06	1.18 ± 0.16
HL1	1 ± 0.13	1.18 ± 0.17	1.04 ± 0.09	0.93 ± 0.14	1.14 ± 0.14	1.76 ± 0.23*	0.95 ± 0.07
HL2	1 ± 0.13	1.55 ± 0.1*	0.99 ± 0.04	1.12 ± 0.1	2.29 ± 0.2**	0.93 ± 0.03	0.93 ± 0.1
T16H1	1 ± 0.04	1.53 ± 0.11**	1.28 ± 0.05**	1.12 ± 0.09	1.22 ± 0.05*	0.79 ± 0.17	0.95 ± 0.07
T16H2	1 ± 0.1	1.29 ± 0.08	0.76 ± 0.06	1 ± 0.15	0.33 ± 0.09**	0.78 ± 0.2	0.67 ± 0.11
16OMT	1 ± 0.09	1.03 ± 0.16	0.99 ± 0.13	1 ± 0.08	0.49 ± 0.03**	0.65 ± 0.24	0.8 ± 0.03
T3O	1 ± 0.05	1.83 ± 0.17**	2.54 ± 0.14***	1.28 ± 0.19	2.64 ± 0.3**	1.73 ± 0.36	1.07 ± 0.11
T3R	1 ± 0.1	0.56 ± 0.11*	0.95 ± 0.04	0.85 ± 0.03	0.47 ± 0.04**	0.55 ± 0.1*	0.73 ± 0.1
NMT	1 ± 0.08	0.93 ± 0.17	0.69 ± 0.06*	0.96 ± 0.09	0.49 ± 0.05**	1.48 ± 0.12*	0.89 ± 0.07
D4H	1 ± 0.22	1.74 ± 0.05*	0.95 ± 0.14	0.67 ± 0.11	0.95 ± 0.39	1.12 ± 0.13	0.44 ± 0.02*
DAT	1 ± 0.05	0.88 ± 0.1	0.7 ± 0.09*	0.77 ± 0.1	0.62 ± 0.09*	1.53 ± 0.37	0.57 ± 0.09**
TAT	1 ± 0.29	1.61 ± 0.15	1.5 ± 0.14	0.52 ± 0.08	11.47 ± 1.27***	0.88 ± 0.36	0.61 ± 0.34

C. roseus agro-infiltrated flower petals from four biological replicates were analyzed by qPCR. Values ± SE of the mean are normalized fold-changes relative to a *35S::GUS* control infiltration. Statistical significance was determined by the Student's t-test (*P < 0.05, **P < 0.01, ***P < 0.001).

Table 1. qPCR showing the effect of transient overexpression of CrMYC2b, CrMYC2c, CrMYC2bD60N, and CrMYC2cD53N on the expression of MIA biosynthesis genes in C. roseus petals. CrMYC2a and CrMYC2aD126N constructs were included as positive controls.

Example 5. Combinatorial overexpression of CrMYC2aD126N, BIS1 and ORCA3 in flower petals has synergistic effects on MIA pathway gene induction

This and previous studies start to shape (part of) the regulatory network of MIA biosynthesis, and accredit defined roles to CrMYC2a, BIS and ORCA TFs (Fig. 1A) (Montiel et al., 2011; Van Moerkercke et al., 2015; Zhang et al., 2011; Zhou and Memelink, 2016). We next wondered whether a combination of overexpression of BIS, ORCA and de-repressed MYC2 TFs would amplify MIA biosynthesis additively or rather surprisingly synergistically.

Therefore, we transiently transformed *C. roseus* flowers by infiltration with *A. tumefaciens* containing BIS1, ORCA3 or CrMYC2aD126N overexpression cassettes, either in a single, a double (pairwise), or the triple combination. First, we assessed the effect on gene expression. Generally, overexpression of either BIS1 or ORCA3 confirmed the previously reported specificities for the MIA pathway branches. BIS1 strongly upregulated all IPAP-specific iridoid pathway genes as well as the BIS2 TF, whereas no effect was observed for MIA pathway genes, hence mirroring the patterns previously obtained in other expression systems (Van Moerkercke et al., 2015) (Fig. 7A; Table 2). Unexpectedly, overexpression of ORCA3 also led to a marked upregulation of several of the iridoid pathway genes, but the overall effect was clearly weaker when compared to flowers expressing BIS1 or CrMYC2aD126N, suggesting this may be an indirect effect, for instance by an (indirect) induction of BIS expression (Fig. 7A; Table 2). In contrast, ORCA3 overexpression, similar to CrMYC2aD126N overexpression but distinct from BIS1 overexpression, clearly induced the seco-iridoid and MIA pathway genes downstream from 7DLH, as well as of the (seco-)iridoid and strictosidine transporters NPF2.6 and NPF2.9 and MIA pathway genes such as GS1, GS2, T19H and MAT (Fig. 7A). Further comparison of expression profiles revealed that only CrMYC2aD126N overexpression led to induction of the MIA pathway genes SGD, HYS, GO and TAT and the iridoid transporter NPF2.5, and that these genes were not affected by BIS1 or ORCA3 overexpression (Fig. 7A).

Combined overexpression of CrMYC2aD126N and BIS1 or of ORCA3 and BIS1 did not have marked synergistic effects, in contrast to combined overexpression of CrMYC2aD126N and ORCA3, which led to a synergistic upregulation of iridoid and MIA pathway genes such as GES, 8HGO, G8O, IS, IO, GS2, T19H, MAT and TAT (Fig. 7A; Table 2). However, synergistic effects were most pronounced in the triple-TF combination, in particular for the gene set mentioned above (Fig. 7A; Table 2). Synergistic activity was most pronounced (over 1,100-fold induction) for T19H and MAT, two genes that are normally expressed at very low level in flowers. Notably, we did not observe any induction of vindoline pathway genes or the recently described Redox1, Redox2, SAT, HL1 and HL2 genes with any of the TF combinations (Table 2).

	35S::GUS	*35S::CrMYC2a^{D126N}*	*35S::ORCA3*	*35S::BIS1*	*35S::ORCA3 35S::CrMYC2a^{D126N}*	*35S::ORCA3 35S::BIS1*	*35S::BIS1 35S::CrMYC2a^{D126N}*	*Triple*
MYC2a	1 ± 0.04	6.9 ± 0.48***	1.16 ± 0.12**	1.13 ± 0.15**	5.56 ± 0.48***	1.26 ± 0.02	7.01 ± 0.66***	6.02 ± 0.45***
BIS1	1 ± 0.16	49.06 ± 3.42***	6.57 ± 1.35**	413.86 ± 16.95***	51.65 ± 2.29***	224.97 ± 19.83***	388.51 ± 40.47***	258.84 ± 34.98**
BIS2	1 ± 0.28	155.52 ± 26.3**	34.44 ± 6.71**	207.66 ± 31.1***	632.42 ± 19.14***	95.96 ± 9***	157.04 ± 29.01	941.66 ± 142.09**
ORCA2	1 ± 0.08	2.71 ± 0.14***	0.85 ± 0.1	0.84 ± 0.04**	2.11 ± 0.11***	1.02 ± 0.08	3.62 ± 0.37***	1.62 ± 0.17
ORCA3	1 ± 0.11	36.31 ± 7.76**	1964.04 ± 162.58***	4.17 ± 1.75*	1638.36 ± 312.84**	1696.22 ± 123.82***	20.5 ± 5.89	1963.22 ± 261.75**
ORCA4	1 ± 0.18	6.97 ± 1.08**	1.49 ± 0.44	1.9 ± 0.44*	4.88 ± 0.54***	1.26 ± 0.33	3.56 ± 0.89	5.33 ± 1.51
ORCA5	1 ± 1	254.83 ± 57.98**	61.75 ± 26.71	60.69 ± 13.74*	142.41 ± 32.76**	13.83 ± 5.18	72.16 ± 21.12	146.38 ± 31.69
GES	1 ± 0.23	1271.92 ± 36.56***	161.75 ± 42.26**	1082.59 ± 96.3***	2510.8 ± 330.03***	591.11 ± 104.71	914.87 ± 100.33***	4462.52 ± 545.62**
8HGO	1 ± 0.07	12.17 ± 3.25**	3.07 ± 0.43**	26.02 ± 5.75**	48.43 ± 3.54***	10.93 ± 2.31	23.24 ± 4.11	85.25 ± 20.17
G8O	1 ± 0.04	50.66 ± 4.46***	9.89 ± 2.22**	87.4 ± 7.85***	167.88 ± 20.67***	37.72 ± 1.93***	48.57 ± 4.75***	189.85 ± 42.37
IS	1 ± 0.24	1677.04 ± 88.06***	270.65 ± 47.1**	3713.16 ± 283.8***	4148.33 ± 370.24***	1389.23 ± 283.14	2297.63 ± 214.4***	5365.99 ± 1469.1
IO	1 ± 0.02	50.5 ± 5.45***	9.76 ± 2.13**	76.3 ± 3.95***	140.43 ± 16.58***	42.46 ± 1.73***	36.49 ± 3.8***	163.72 ± 11.69**
7DLGT	1 ± 0.07	4.81 ± 0.5***	1.63 ± 0.09**	6.03 ± 0.48***	9.65 ± 0.39***	3.67 ± 0.05***	4.98 ± 0.35***	15.32 ± 0.67***
7DLH	1 ± 0.06	0.69 ± 0.04	1.22 ± 0.04	0.87 ± 0.07**	1.07 ± 0.02	1.03 ± 0.07	0.71 ± 0.05	1.03 ± 0.08
NPF2.4	1 ± 0.07	0.87 ± 0.09	1.01 ± 0.07	0.8 ± 0.04*	1.13 ± 0.04	0.99 ± 0.08	1.45 ± 0.12	0.93 ± 0.08
NPF2.5	1 ± 0.19	13.1 ± 1.74***	0.76 ± 0.04	1.24 ± 0.27**	6.32 ± 0.7***	0.78 ± 0.16	13.14 ± 2.47	8.73 ± 1.59
NPF2.6	1 ± 0.02	2.39 ± 0.42**	2.08 ± 0.14***	0.85 ± 0.07**	4.75 ± 0.24***	2.03 ± 0.23	3.14 ± 0.48	5.05 ± 0.95
LAMT	1 ± 0.13	25.38 ± 2.07***	6.13 ± 0.4***	1.42 ± 0.25**	32.01 ± 3.43***	5.7 ± 0.24***	20.19 ± 3.89	39.62 ± 7.7
SLS	1 ± 0.05	3.1 ± 0.24***	3.76 ± 0.29***	1.16 ± 0.1**	6.92 ± 0.51***	3.63 ± 0.36***	3.33 ± 0.31***	7.38 ± 1.51
TDC	1 ± 0.13	5.58 ± 0.21***	6.33 ± 0.25***	1.29 ± 0.11**	10.71 ± 0.5***	5.98 ± 0.36***	5.22 ± 0.44***	10.77 ± 1.26***
STR	1 ± 0.1	2.64 ± 0.29**	5.08 ± 0.42***	1.19 ± 0.12**	7.35 ± 0.74***	5.43 ± 0.75	2.99 ± 0.19***	7.8 ± 1.27
NPF2.9	1 ± 0.11	4.54 ± 0.33***	3.07 ± 0.35**	0.82 ± 0.05**	4.41 ± 0.17***	2.77 ± 0.21***	2.84 ± 0.2***	6.91 ± 0.49***
SGD	1 ± 0.09	3.19 ± 0.34***	1.5 ± 0.21	1.47 ± 0.34**	6.4 ± 0.64***	1.42 ± 0.28	3.58 ± 0.36***	5.39 ± 1.01
HYS	1 ± 0.12	4.77 ± 0.74**	0.69 ± 0.03*	1.27 ± 0.26**	3.96 ± 0.76**	0.62 ± 0.02	5.67 ± 0.78***	3.13 ± 0.26***
GS1	1 ± 0.04	1.65 ± 0.08***	2.27 ± 0.07***	0.97 ± 0.11**	4.11 ± 0.21***	2.15 ± 0.13***	1.82 ± 0.19	4.4 ± 0.28***
GS2	1 ± 0.07	5.08 ± 0.57***	6.23 ± 0.36***	1.14 ± 0.05**	16.85 ± 2.13***	5.86 ± 0.56***	5.02 ± 1.05	18.21 ± 1.28***

	35S::GUS	35S::CrMYC2a^{D126N}	35S::ORCA3	35S::BIS1	35S::ORCA3 35S::CrMYC2a^{D126N}	35S::ORCA3 35S::BIS1	35S::BIS1 35S::CrMYC2a^{D126N}	Triple
GO	1 ± 0.03	5.14 ± 0.44***	1.54 ± 0.07***	0.96 ± 0.12**	4.25 ± 0.21***	1.26 ± 0.19	5.05 ± 0.43***	4.94 ± 0.22***
Redox1	1 ± 0.06	1.03 ± 0.15**	1.01 ± 0.21*	0.7 ± 0.04	1.97 ± 0.13***	1.35 ± 0.12	0.89 ± 0.2	1.59 ± 0.38
Redox2	1 ± 0.11	0.6 ± 0.03*	1.49 ± 0.09*	0.72 ± 0.05*	1.43 ± 0.09	1.66 ± 0.15	0.55 ± 0.07	1.17 ± 0.31
SAT	1 ± 0.08	0.76 ± 0.1	1.75 ± 0.24**	0.89 ± 0.05	1.41 ± 0.27	1.61 ± 0.09	0.86 ± 0.06	1.52 ± 0.36
HL1	1 ± 0.07	0.37 ± 0.04***	0.79 ± 0.09	0.54 ± 0.03***	0.89 ± 0.12	0.75 ± 0.1	0.28 ± 0.04***	0.77 ± 0.13
HL2	1 ± 0.1	1.3 ± 0.11	0.82 ± 0.05	0.68 ± 0.06	1.11 ± 0.03	0.85 ± 0.04	1.24 ± 0.14	1.03 ± 0.09
T16H1	1 ± 0.1	0.99 ± 0.09	0.8 ± 0.03	1.17 ± 0.05	1.05 ± 0.06	0.96 ± 0.05	1.3 ± 0.07	0.96 ± 0.04
T16H2	1 ± 0.01	0.45 ± 0.04***	1.36 ± 0.11*	1.05 ± 0.11**	0.99 ± 0.1	1.44 ± 0.17	0.7 ± 0.03***	0.74 ± 0.13
16OMT	1 ± 0.17	0.52 ± 0.03	2.67 ± 0.22*	1.62 ± 0.05	1.64 ± 0.3	2.06 ± 0.23	0.67 ± 0.15	1.39 ± 0.34
T3O	1 ± 0.11	1.59 ± 0.41	2.53 ± 0.27*	1.1 ± 0.08	1.27 ± 0.18	2.34 ± 0.44	0.88 ± 0.15	1.17 ± 0.24
T3R	1 ± 0.06	0.91 ± 0.05	1.51 ± 0.14*	1 ± 0.07**	1.18 ± 0.07	1.56 ± 0.17	0.57 ± 0.02***	1.37 ± 0.14
NMT	1 ± 0.06	0.12 ± 0.03***	1.38 ± 0.11	0.44 ± 0.07**	0.72 ± 0.17**	0.87 ± 0.15	0.16 ± 0.03***	0.29 ± 0.07***
D4H	1 ± 0.12	0.63 ± 0.16	0.79 ± 0.08	0.8 ± 0.14	0.75 ± 0.2	0.74 ± 0.22	0.69 ± 0.19	0.96 ± 0.18
DAT	1 ± 0.07	0.17 ± 0.05***	0.92 ± 0.07	0.43 ± 0.06***	0.88 ± 0.08	0.63 ± 0.14	0.3 ± 0.07***	0.8 ± 0.09
T19H	1 ± 0.13	9.58 ± 0.32***	126.65 ± 10.24***	2.64 ± 0.6**	1050.8 ± 195.03*	147.35 ± 28	6.96 ± 0.62***	1138.76 ± 363.3
MAT	1 ± 0.1	2.39 ± 0.17***	6.09 ± 0.13***	0.82 ± 0.07*	120.56 ± 19.68***	5.91 ± 0.31***	1.77 ± 0.23	133.44 ± 7***
TAT	1 ± 0.24	5.22 ± 0.84**	2.33 ± 0.85	1.26 ± 0.31**	21 ± 3.12***	1.48 ± 0.27	3.6 ± 0.42	26.42 ± 2.5***

C. roseus agro-infiltrated flower petals from four biological replicates were analyzed by qPCR. Values ± SE of the mean are normalized fold-changes relative to a *35S::GUS* control infiltration. Statistical significance was determined by the Student's t-test (*P < 0.05, **P < 0.01, ***P < 0.001). "Triple" stands for simultaneous overexpression of *CrMYC2a*^{D126N}, *ORCA3* and *BIS1*.

Table 2. qPCR showing the effect of single, double or triple transient overexpression of CrMYC2aD126N, ORCA3 and BIS1 on the expression of MIA biosynthesis genes in C. roseus petals.

As we were also interested in potential amplification loops, we assessed expression of the TFs themselves, including CrMYC2a, the two BIS homologs, and the four ORCA homologs, i.e. ORCA2 to ORCA5 (Paul et al., 2017) (Fig. 8). qPCR targeting the 3' untranslated region (3'UTR) of CrMYC2a suggested no autoregulation or cross-regulation by ORCA3 or BIS1 overexpression. In agreement with our previous reports, BIS2 expression was markedly upregulated by BIS1 overexpression, over 200-fold (Van Moerkercke et al., 2016). Likewise, overexpression of CrMYC2aD126N led to a strong upregulation of the expression of both BIS TFs, ca. 50- and 150-fold, respectively. ORCA3 overexpression also induced the expression of BIS1 and BIS2, but weaker, respectively less than 10- and 40-fold, respectively. Notably, combined overexpression of CrMYC2aD126N and ORCA3 had a strong synergistic effect on the expression of BIS2 (over 600-fold) but not on that of BIS1. Expression of endogenous ORCA3 was ca. 5-fold upregulated by either BIS1 or ORCA3 overexpression. This can be considered as only a minor, possibly indirect effect, given the almost 90-fold induction observed with CrMYC2aD126N overexpression, which is in agreement with the known direct regulation. A similar pattern was also observed for ORCA5, but not ORCA2 or ORCA4 (Fig. 8). In contrast to the pathway genes themselves, no combinatorial or synergistic effects on TF gene expression could be spotted in the double or triple TF combinations, except for BIS2.

To correlate CrMYC2aD126N mediated transactivation with the possible presence of potential cis-regulatory elements, we searched the promoters of all MIA pathway genes for canonical MYC2 binding motifs (Godoy et al 2011). In general, we noticed that promoter sequences (2 kb upstream of the start codon) of the MIA pathway and transporter genes contained on average 1.7 potential MYC2 binding boxes (data not shown). Promoters of genes that could not be induced by CrMYC2aD126N overexpression contained on average 1.5 binding elements whereas, CrMYC2aD126N inducible promoters contained on average 1.88 binding elements. Noteworthy, the promoters containing 3 or more potential MYC2-binding domains such as LAMT (5), STR (5), T19H (4), HYS (4), TAT (3), IO (3) and G8O (3) are all strongly up-regulated in CrMYC2aD126N overexpressing flowers, suggesting some correlation, although weak, between the number of potential MYC2 binding elements and inducibility by CrMYC2aD126N.

Taken together, these findings further support that the expression of MIA pathway genes, as well as the transporter and TF genes, are orchestrated by the combinatorial action of TFs such as CrMYC2a, BIS1 and ORCA3 (see the model proposed in Fig. 9).

Example 6. Combinatorial overexpression of CrMYC2aD126N, BIS1 and ORCA3 is necessary and sufficient to elicit MIA accumulation in flower petals

To further explore the individual and combinatorial effects of the three TFs and to correlate gene expression with metabolite production, we carried out metabolite profiling by LC-FT-ICR-MS. First, the positive effect of CrMYC2aD126N overexpression on strictosidine accumulation was further increased (though not statistically significant) by the triple-TF infiltration, in agreement with the synergistic effects on MIA biosynthesis gene expression. Overexpression of BIS1 or ORCA3 did not affect strictosidine levels (Fig. 7B and C; Table 3). Likewise, no effect was observed on the levels of other abundant MIAs such vindoline, catharanthine, serpentine and vinblastine by any of the TF infiltrations (Table 3), which is in agreement with the observed lack of induction of the vindoline pathway genes. However, chromatogram comparisons displayed a number of prominently increased peaks, besides that corresponding to strictosidine, in the triple-TF infiltrated flowers, for instance at retention time (RT) 13.46 and 14.07 (Fig. 7B). The first peak corresponded to a compound with a parent [M+H]+ ion at m/z 353.1858 and accumulated up to 10-fold more in triple-TF infiltrated flowers compared to the control GUS-infiltrated flowers (Fig. 7B; Table 3). Based on MS fragmentation, this compound was identified as 16-hydroxytabersonine, a vindoline precursor, the identity of which was further confirmed by a standard synthesized in yeast. 16-hydroxytabersonine also accumulated two- to three-fold more in flowers overexpressing ORCA3 and CrMYC2aD126N, respectively, but not in BIS1-overexpressing flowers (Fig. 7C; Table 3). The second peak corresponded to a compound with a parent [M+H]+ ion at m/z 625.1762 and was accumulating 4.5-fold higher in triple-TF infiltrated flowers. MS fragmentation suggested this compound to be a flavonoid, most likely an isomer of rhamnetin 3-rutinoside. Given that this putative flavonoid also showed increased accumulation in the ORCA3-overexpressing flowers but not in the CrMYC2aD126N- or BIS1-overexpressing flowers, this effect seems to be strictly related to the overexpression of ORCA3, hence we did not further focus on it in this study.

RT	m/z	compound	*35S::GUS*	*35S::CrMYC2a^{D126N}*	*35S::ORCA3*	*35S::BIS1*	*Triple*
5.9	385.1756622	16-hydroxyhörhammericine	0	0.76	54.7**	1.64	3138.5***
9.03	369.1807005	Hörhammericine	19.56	219.73**	427.97**	35	46855.44**
9.25	355.2014712	Minovincinine	3.61	0.91	0.14	0	1157.44**
10.32	353.1858429	19-hydroxytabersonine	191.83	287.17	253.13	152.56	17287.91**
10.32	399.1912563	16-methoxyhörhammericine	77.12	389.41***	6695.37**	50.69	180330.08**
12.53	339.1702073	Perivine	2629.07	36288.96***	3765.59	2425.07	18763.94**
12.77	355.2015529	Unknown 1	10379.72	16477.59*	14456.35	9016.4	37921.49**
13.41	369.1807363	Unknown 2	7321.48	16201.74**	35307.97	8745.72	78924.36***
13.44	411.1912605	16-hydroxy-19-O-acetyltabersonine	1.15	8.38	0	0	9651.33**
13.46	353.1858266	16-hydroxytabersonine	5561.81	16827.53***	11251.68*	4793.75	57445.55**
13.77	385.2119976	3-hydroxy-16-methoxy-2,3-dihydrotabersonine	4.8	16.72	56.26*	0.7	339**
13.89	355.2015357	Isositsirikine	4564.55	8520.22**	5369.45	4222.45	23638.37**
14.07	625.1761518	rhamnetin 3-rutinoside†	51554.42	17825.41	228004.08**	15040.48	235162.76**
14.26	531.2334599	Strictosidine	313975.62	740006.26**	285060.87	299576.42	1136572.05*
14.46	355.2014581	16-hydroxyvincadifformine	40.49	177.56*	37.64	28.29	9268.12**
15.39	397.2120112	O-acetylstemmadenine	5377.71	24386.49***	7338.19	5806.2	35154.84***
15.52	415.2225487	Deacetylvindoline	11211.16	11176.51	8470.31	8233.43	11696.03
15.55	323.1752583	Akuammicine	1795.36	18494.54***	4100.27*	1486.82	26268.25***
15.91	353.1926627	Ajmalicine / Tetrahydroalstonine / Geissoschizine?	297.05	985.1**	932.95	292.78	4062.07***
16.4	337.1909524	Catharanthine	302711.78	414283.42	242100.58	290839.24	343682.75
16.48	889.3596611	Strictosidine secologanatoside	40905.41	130246.54**	37647.91	50260.03	273030.75***
17.26	349.1545094	Serpentine	149625.58	172196.19	129725.99	103865.99	157820.75
17.4	413.2069449	Vincristine	1375.65	1822.9	923.76	2224.11	1897.34
17.6	339.2065363	Vincadifformine	8.96	15.2	7.79	6.3	1599.64**
18.35	903.3751804	Strictosidine secologanoside	38453.17	109147.93**	75237.01	40707.83	223811.26**
18.87	367.2014746	16-methoxytabersonine	47.67	200.46**	98.14	37.65	496.79*
18.87	811.4273487	Vinblastine	6500.17	7831.44	4845.72	7390.03	10077.88
19.1	457.2330113	Vindoline	1072371.487	1254203.45	869881.4343	989551.5491	1184040.503
19.44	427.2224893	Demethoxyvindoline	10793.74	12398.13	20315.59	8783.75	24660*
19.85	399.211155	Desacetoxyvindoline	4.24	34.94	9.82	3.02	52.11*
21.03	793.4167214	Anhydrovinblastine	38051.39	92011.48	33880.52	78144.26	31663.22

Table 3. Triple overexpression of CrMYC2aD126N, ORCA3 and BIS1 boosts the production of MIAs in C. roseus petals. *C. roseus* agro-infiltrated flower petals from four biological replicates were analyzed by LC-ESI-FT-ICR-MS. Numbers indicate average total ion current (TIC) of the peaks. Statistical significance was determined by the Student's t-test (*P < 0.05, **P < 0.01, ***P < 0.001). Peak identity was determined by standards or inferred from MS_n fragmentation (see supplemental data). "Triple" stands for simultaneous overexpression of *CrMYC2a^{D126N}*, *ORCA3* and *BIS1*. † Compound is not a MIA.

Two other peaks, with RTs at 9.03 and 10.32 and parent [M+H]+ ions at m/z 369.1807 and 399.1913, respectively, were dramatically increased in the triple-TF infiltrated flowers. The increased accumulation amounted up to over 2,300-fold in the triple-TF infiltrated flowers and were barely detectable in samples derived from the GUS control or single-TF infiltrated flowers (Fig. 7B and C; Table 3). Surprisingly, based on MS fragmentation, these compounds were identified as the postulated root-specific MIAs hörhammericine (identity confirmed with an authentic standard) and 16-methoxyhörhammericine, also known as hörhammerinine (Fig. 7B and C; Table 3) (Abraham and Farnsworth, 1969). Additional novel peaks, potentially corresponding to strictosidine derivatives (including bis(monoterpenoid) indole alkaloids such as strictosidine secologanatoside and strictosidine secologanoside) and other MIAs, showed a similar marked increase in the triple-TF infiltrated flowers, but their identity could not unambiguously be determined (Table 3).

To further investigate the effects of the single- and triple-TF combinations on the flower metabolome, PCA was carried out (Fig. 7D). Samples of BIS1-overexpressing flowers largely overlapped with those of the control, but all others formed distinct groups, with the triple-TF infiltrated flowers showing the most divergent distribution pattern. According to the PCA loading plot, distribution of PC1 was mainly explained by variables corresponding to the above-mentioned differentially accumulating MIAs. Conversely, PC2 seemed to be mostly influenced by compounds related to flavonoids such as the tentatively annotated rhamnetin 3-rutinoside, explaining the distribution pattern of the samples infiltrated with the ORCA3 TF.

The observation that combinatorial overexpression of TFs in flowers had no effect on the accumulation of catharanthine and vindoline, but could boost the production of certain MIAs, encouraged us to take a closer look at potential MIA intermediates and final MIA products (Fig. 1A and B). Based on MS fragmentation and standards, we identified and quantified 18 additional MIAs (Table 3). As recently published, most intermediates between the strictosidine aglycone and tabersonine or catharanthine are unstable, hypothetical or not detectable in planta, which explains why we were not able to detect for instance deshydroxymethylstemmadenine or the critical intermediate stemmadenine (Qu et al 2018). Nevertheless, we observed an important increase of O-acetylstemmadenine (>6-fold) and 'early-branching' MIAs such as akuammicine (>14-fold), perivine (>7-fold) and isositsirikine (>5-fold) in triple-TF infiltrated flowers and to lesser extent in flowers infiltrated with 35S::CrMYC2aD126N, which correlated with the high induction observed for GS1, GS2 and GO (Fig. 1A and 7A; Table 3). Besides the above-mentioned 16-hydroxytabersonine, the subsequent downstream intermediates of the vindoline pathway, 16-methoxytabersonine and 3-hydroxy-16-methoxy-2,3-dihydrotabersonine, accumulated only poorly in control samples, but showed nevertheless slightly increased levels in triple-TF infiltrated

flowers (Table 3). The last precursors and derivate compounds in the vindoline pathway, like deacetylvindoline, desacetoxyvindoline and demethoxyvindoline, were relatively highly abundant, but remained unchanged among all samples (Table 3). In contrast to the 16-hydroxytabersonine-derived MIAs that are normally produced in the aerial plant parts, the 19-hydroxytabersonine-derived MIAs, such as hörhammericine, are normally produced in roots (Giddings et al., 2011). Considering the dramatic increase of root-specific MIAs, we screened our data for other MIAs in this specific branch (Fig. 7C and 1B). Similar to hörhammericine, most root-specific MIAs were barely detectable in the control flowers, but displayed a massively increased accumulation in triple-TF infiltrated flowers, including 19-hydroxytabersonine (>90-fold), 16-hydroxyhörhammericine (not detectable in control flowers), minovincinine (>320-fold), vincadifformine (>170-fold) and 16-hydroxyvincadifformine (>220-fold) (Table 3). Surprisingly, although expression of MAT and TAT was strongly increased in triple-TF infiltrated samples, we were not able to detect any previously described acetylated MIAs such as echitovenine, 19-O-acetylhörhammericine or 19-O-acetyltabersonine (Giddings et al., 2011; Laflamme et al., 2001). Instead we detected an over 8,300-fold increased accumulation of an acetylated compound in triple-TF infiltrated samples that we tentatively identified as 16-hydroxy-19-O-acetyltabersonine (Table 3). This acetylated MIA, at the detection limit in all single TF infiltrated samples, is most likely the result of a concerted reaction involving T16H and the highly expressed T19H and TAT in a yet undiscovered order. 16-methoxy-19-O-acetyltabersonine, a similarly acetylated MIA could be produced in yeast in a previous study, confirming that C-16 hydroxylation and C-19 hydroxylation/acetylation can occur on the same molecule (Carqueijeiro et al 2018 TPJ 94:469-484).

Taken together, metabolite profiling reveals that combinatorial overexpression of BIS1, ORCA3 and CrMYC2aD126N has a tremendous effect on the general MIA profile, an effect which cannot be obtained with single TFs, and correlating with the MIA pathway gene induction that they trigger together.

Example 7. Identification of orthologous genes of BIS1

As shown by the phylogenetic analysis in Figure 4, BIS1 (Caros001862.1) and BIS2 (Caros006385.1) are clade IVa bHLH transcription factors. This clade is characterized by the absence of a JAZ interacting domain. Orthologues in *Arabidopsis thaliana* are At2g22750 (bHLH18), At4g37850 (bHLH25), At2g22760 (bHLH19) and At2g22770 (bHLH20). Orthologues *Medicago truncatula* are MtTSAR1 (SEQ ID No. 29) and MtTSAR2 (SEQ ID No. 30) (see Mertens et al 2016 Plant & Cell Physiology 57:2564-2575).

Example 8. Identification of orthologous genes of ORCA3

ORCA3 belongs to the AP2/ERF (formerly called AP2/EREBP) subfamily of AP2-domain proteins (Fujimoto et al., 2000 Plant Cell 12:393-404; Riechmann and Meyerowitz, 1998 Biol Chem 379:633-646) with a

single AP2 domain (SEQ ID No. 24). In contrast, the AP2 subfamily contains two AP2 domains. In addition, the amino acid sequences of the AP2 domain and the AP2/ERF domain are quite divergent (for details see Fujimoto et al., 2000 Plant Cell 12:393-404). When ORCA3 is aligned with a number of functionally characterized AP2/ERF-type proteins from different plant species (NtEREBP-1, NtEREBP-2, NtEREBP-3, NtEREBP-4, AtERF1, AtERF2, AtERF3, AtERF4, AtERF5, LePti4, LePti5, LePti6, CrORCA1, CrORCA2), a high level of homology within the AP2/ERF domain is found (see Figure 2a in van der Fits and Memelink 2001 TPJ 25:43-53). Based on sequence and structural homology, several conserved domains can be pointed out in the ORCA3 protein. ORCA3 forms part of a subgroup of AP2/ERF-domain proteins having a conserved C-terminal extension of the AP2/ERF domain of about 20 amino acids (SEQ ID No. 25). Outside the AP2/ERF domain, little or no homology with other AP2/ERF-domain proteins is observed. The AP2/ERF domain of ORCA3 contains a putative bipartite nuclear localization signal (SEQ ID No. 26). In the N-terminus of the ORCA3 protein, an acidic domain can be pointed out (SEQ ID No. 27). Furthermore, a serine-rich region is present in the C-terminus of the ORCA3 protein (SEQ ID No. 28).

Homologues in *C. roseus* are the ORCA2 protein (GenBank: CAB93940.1) and the ORCA3-like protein 1 from *C. roseus* with accession or sequence ID ANC60170.1 (51% identity).

Orthologues of ORCA3 in *Datura metel* are the AP2-binding DNA-binding domain protein with accession or sequence ID ABU40945.1 (98% identity).

Conclusion

Metabolic pathways are regulated and fine-tuned in a complicated way and increased induction of pathway gene expression does not necessarily lead to increased flux through the pathway or increased accumulation of the desired metabolites. This is particularly applicable to the MIA pathway with its numerous branches and high degree of cellular and organellar compartmentalization. For instance, overexpression of BIS1 or CrMYC2a in *C. roseus* flower petals had no effect on the level of the 24 flower MIAs identified here. Overexpression of ORCA3 and CrMYC2aD126N resulted in a statistically significant increase in the levels of six and nine of these MIAs, respectively. However, a MIA pathway-encompassing effect, leading to a dramatic increase in the levels of 23 MIAs was only obtained by the triple-TF combination, i.e. when wild-type BIS1 and ORCA3, and the de-repressed CrMYC2aD126N were concomitantly transiently overexpressed. This demonstrates that an intelligent and inventive design of combinatorial TF cassettes allows to create superior engineering tools, even encompassing transcriptional regulation across multiple cell types in a single engineering event.

In conclusion, we here provide a broad view on the hierarchy, range, and cell type- and pathway branch-specificity of and within a specific JA-modulated regulatory module that involves CrMYC2a, BIS and ORCA TFs. This view is visualized in a simplified model (Fig. 9). Elements of the network that have been exposed

include: (i) the first direct regulator of SGD expression, namely CrMYC2a, which is particularly valuable because SGD represents the gateway towards all MIAs, (ii) the TFs regulating directly and indirectly the expression of known (seco)iridoid MIA transporters, which is crucial given the spatial distribution of the MIA pathway across cell types and compartments, (iii) the capacity of a combinatorial TF cassette to induce a genuine MIA pathway, thereby overruling previously anticipated organ-specific boundaries such as the unexpected induction of the postulated root-specific T19H, MAT and TAT genes, (iv) the existence of amplification loops or TF cascades, e.g. positioning CrMYC2a as a strong and direct inducer of BIS and ORCA expression, and (v) the existence of negative feedback loops as evidenced by the fact that an engineered, de-repressed CrMYC2a variant needed to be employed to enable its activity in the absence of the JA signal.

Materials and methods

Plant material and transformation

Flower petals of one-year-old *Catharanthus roseus* var. "Little bright eyes" plants (grown under greenhouse conditions) were infiltrated with *Agrobacterium tumefaciens* C58C1 harboring the respective constructs for overexpression (see below) as previously described (Van Moerkercke et al., 2016), with slight modifications. One day prior to the infiltration experiment, all open flowers were removed from the plants. The day of infiltration, liquid *A. tumefaciens* cultures were washed, resuspended and diluted to a final OD=0.3 in infiltration buffer (100 µM acetosyringone, 10 mM MgCl2, 20 mM MES) and left shaking for 3 h. Petals of freshly opened flowers were pierced with a needle and then infiltrated with a 1-mL syringe containing the *A. tumefaciens* solution. After 48 h of incubation, petals from five different flowers per sample were cut in two and immediately flash-frozen in liquid nitrogen. One sample was used for RNA extraction for gene expression analysis and the other sample was used for metabolite profiling.

Generation of DNA constructs

Constructs were made with the Gateway® Technology (Invitrogen). ENTRY clones containing AtMYC2, AtMYC2D105N, CrMYC2a, BIS1 and ORCA3 had been generated previously (Goossens et al., 2015; Van Moerkercke et al., 2015). For all other transcription factors used in this study, the coding sequences were amplified from *C. roseus* var. "Little bright eyes" cDNA with Q5® High-Fidelity DNA Polymerase (New England BioLabs®) and gene-specific oligonucleotides containing attB Gateway® recombination sites. The fragments were gel-purified (GeneJET Gel Extraction Kit, ThermoFisher) and recombined into pDONRTM221 using Gateway® BP clonaseTM II enzyme mix (ThermoFisher); the resulting ENTRY clones were sequenced. CrMYC2aD126N, CrMYC2bD60N, and CrMYC2cD53N were constructed by overlap

extension PCR using the ENTRY clones as templates. For expression in flower petals under control of the CaMV35S promoter or promoter transactivation assays, ENTRY clones were recombined into pK7WG2D or p2GW7 (Karimi et al., 2002), respectively, using LR ClonaseTM enzyme mix (ThermoFisher). The promoter fragments of GES, G8O, 8HGO, IS, IO, 7DLGT, 7DLH, LAMT, STR and BIS1 had been cloned previously (Van Moerkercke et al., 2016; Van Moerkercke et al., 2015; Vom Endt et al., 2007). The promoter fragments (including 5' UTRs) of SLS (1413bp), TDC (2353 bp) and SGD (1651 bp) were amplified from C. roseus var. "Little bright eyes" genomic DNA and recombined into ENTRY vectors. ENTRY clones containing promoter fragments were recombined into pGWL7 (Karimi et al., 2002) for transient expression assays.

Phylogenetic analysis

Amino acid sequences of full-length proteins of *C. roseus* and *A. thaliana* were aligned using ClustalW. A neighbor-joining tree was constructed in MEGA6 (Tamura et al., 2013). Evolutionary distances were calculated using the Jones, Taylor, and Thornton (JTT) model. Bootstrap analyses were performed with 1,000 bootstrap replicates. bHLH clades were defined as in Heim et al. (2003).

Protein interaction assays in yeast

Y2H assays were performed as described (Cuéllar Pérez et al., 2013). Bait and prey were fused to the GAL4-AD or GAL4-BD via cloning into pGAD424gate or pGBT9gate, respectively. The *Saccharomyces cerevisiae* PJ69-4A yeast strain was co-transformed with bait and prey using the polyethylene glycol (PEG)/lithium acetate method. Positive transformants were selected on Synthetic Defined (SD) medium lacking Leu and Trp (-2) (Clontech). Colonies were grown overnight in liquid cultures (-2) at 30°C and 10- or 100-fold dilutions were dropped on control medium (-2) and selective medium lacking Leu, Trp and His (-3) (Clontech).

Transient expression assays in N. tabacum *protoplasts*

Transactivation of *C. roseus* gene promoters by *C. roseus* TFs was assessed in transient expression assays in *N. tabacum* 'Bright Yellow-2' protoplasts as described (Vanden Bossche et al., 2013). Protoplasts were co-transfected with reporter, effector and normalizer plasmids. The reporter plasmid consists of a fusion between the promoter of interest and the FIREFLY LUCIFERASE (fLUC) gene. The effector plasmid contains the TF of interest driven by the CaMV35S promoter. The normalizer plasmid contains the RENILLA LUCIFERASE (rLUC) driven by the CaMV35S promoter. fLUC and rLUC readouts were collected after overnight incubation and lysis using the Dual-Luciferase Reporter Assay System (Promega). Each assay incorporated four or eight biological repeats. Normalized values of the promoter activities were

obtained by dividing the fLUC values by the corresponding rLUC values. The average of the normalized fLUC values was calculated and set out relative to the control fLUC values (i.e. with an effector plasmid carrying the GUS gene instead of a TF gene).

Transcript profiling by qPCR

RNA from agro-infiltrated flower petals was extracted using the RNeasy Plant Mini kit (Qiagen) including an additional DNA on-column digest with RQ1 DNase (Promega). Up to 500 ng of RNA was used as a template for reverse transcription with the iScriptTM cDNA synthesis kit (Bio-Rad). qPCR was performed, on a LightCycler® 480 (Roche) using SYBR Green qPCR master Mix (Agilent) and gene-specific oligonucleotides (Table S6) in two technical replicates. The data was normalized to two reference genes, N2227 and SAND (Pollier et al., 2014) using qBase (Hellemans et al., 2007).

Metabolite profiling

For metabolite profiling, infiltrated petals were snap-frozen and ground in liquid nitrogen. After lyophilization, around 5 mg of dry petal tissue was weighed and extracted with 1 mL of methanol containing 100 mM caffeine as an internal standard. The samples were extracted at room temperature for 10 min and centrifuged for 10 min at 20,800 x g. The resulting supernatant was evaporated to dryness under vacuum, and the residue was dissolved in 800 µL of water/cyclohexane (1:1, v/v). The samples were centrifuged (10 min at 20,800 x g), and 200 µL of the aqueous phase was retained for LC-MS analysis. For LC-MS, 10 µL of the sample was injected in a ZORBAX RRHD Eclipse XDB-C18 column (2.1 x 150 mm, 1.8 µm) mounted on a thermo instrument equipped with an LTQ FT Ultra and an electrospray ionization source. The following gradient was run using acidified (0.1% (v/v) formic acid) solvents A (water/acetonitrile (99:1, v/v)) and B (acetonitrile/water; 99:1, v/v)): time 0 min, 5% B; 30 min, 55% B; 35 min, 100% B. FT-MS spectra between m/z 120-1400 were recorded at a resolution of 100,000. For identification, full MS spectra were interchanged with a dependent MS2 scan event in which the most abundant ion in the previous full MS scan was fragmented, two dependent MS3 scan events in which the two most abundant daughter ions were fragmented and a dependent MS4 scan event in which the most abundant granddaughter ion was fragmented. The collision energy was set at 35%. Peak areas were calculated using the Progenesis QI software and normalized by petal sample mass and the peak area of caffeine. Compounds were identified based on their exact mass, fragmentation and co-elution with authentic standards. Compounds for which no authentic standards are available were tentatively identified based on their exact mass and MSn fragmentation.

The production of the standards 16-hydroxytabersonine and 19-hydroxytabersonine was carried out based in yeast on the biotransformation method reported by Qu et al. (2015) with modifications. Briefly,

T16H2 or T19H were cloned via Gateway recombination into the yeast vector pAG423GAL-ccdB (Alberti et al., 2007) (AddGene plasmid 14149) and co-transformed in the wild-type yeast strain PA59 (Arendt et al., 2017) (MATa; his3Δ1; leu2Δ0; ura3Δ0; lys2Δ0; trp-1) with the P450 reductase MTR1 (Miettinen et al., 2017) cloned in pAG415GAL-ccdB (Alberti et al., 2007) (AddGene plasmid 14145). Transformed colonies were selected on Synthetic Defined (SD) medium (Clontech) with -His, -Leu (Clontech) amino acid dropouts (DO). For compound production, pre-cultures were grown at 30°C with shaking at 250 rpm for 24 h in 5 mL of liquid SD medium with -His, -Leu DO. To induce gene expression, the precultures were washed and inoculated into 50 mL of SD Gal/Raf medium with -His, -Leu DO supplements to a starting optical density of 0.25. The cultures were incubated for 24 h before addition of 125 µM of tabersonine (dissolved in ethanol) and further incubated for another 48 h. The produced 16-hydroxytabersonine or 19-hydroxytabersonine were extracted from the yeast medium using ethyl acetate and analyzed by LC-FT-ICR-MS as described above.

For details on all metabolite identifications, see Supplemental Dataset S1. The PCA was performed with MetabolAnalyst 3.0 with Pareto-scaled mass spectrometry data and standard settings (http://www.metaboanalyst.ca/) (Xia and Wishart, 2016).

Data deposition

The sequences reported in this paper have been deposited in the GenBank database [accession nos. MG676658 (CrICE1a), MG676659 (CrICE1b), MG676660 (CrMYC2b), MG676661 (CrMYC2c), MG676662 (CrJAM2), MG676663 (CrJAM3), MG676664 (CrGL3), MG676665 (CrMYBETC1a), MG676666 (CrMYBETC1b), MG676667 (CrR2R3MYB63). MG676668 (CrR2R3MYB116), MG676669 (CrR2R3MYBTT2a), MG676670 (CrR2R3MYBTT2b), MG676671 (CrANAC002), MG676672 (CrANAC055), MG676673 (CrANAC056), MG676674 (CrWRKY40), MG676675 (CrWRKY75), MG676676 (CrJAZ7), MG676677 (CrJAZ4), MG676678 (pSGD), MG676679 (pSLS1)].

Sequences

The following sequences can be found in the sequence list

SEQ ID No. 1: WT CrMYC2a protein (Catharanthus roseus)

SEQ ID No. 2: Mutant CrMYC2a protein (Catharanthus roseus)

SEQ ID No. 3: WT CrMYC2b protein (Catharanthus roseus)

SEQ ID No. 4: Mutant CrMYC2b protein (Catharanthus roseus)

SEQ ID No. 5: WT CrMYC2c protein (Catharanthus roseus)

SEQ ID No. 6: Mutant CrMYC2c protein (Catharanthus roseus)

SEQ ID No. 7: JAZ interacting domain of CrMYC2a (Catharanthus roseus)

SEQ ID No. 8: JAZ interacting domain of CrMYC2a (Catharanthus roseus)

SEQ ID No. 9: JAZ interacting domain of CrMYC2a (Catharanthus roseus)

SEQ ID No. 10: JAZ interacting domain of CrMYC2a (Catharanthus roseus)

SEQ ID No. 11: WT CrMYC2a DNA (Catharanthus roseus)

SEQ ID No. 12: JAZ interacting domain of CrMYC2a (Catharanthus roseus)

SEQ ID No. 13: WT CrMYC2b DNA (Catharanthus roseus)

SEQ ID No. 14: JAZ interacting domain of CrMYC2b (Catharanthus roseus)

SEQ ID No. 15: WT CrMYC2c DNA (Catharanthus roseus)

SEQ ID No. 16: JAZ interacting domain of CrMYC2c (Catharanthus roseus)

SEQ ID No. 17: WT AtMYC2 DNA (Arabidopsis thaliana)

SEQ ID No. 18: WT BIS1 protein (Catharanthus roseus)

SEQ ID No. 19: WT BIS2 protein (Catharanthus roseus)

SEQ ID No. 20: WT BIS1 DNA (Catharanthus roseus)

SEQ ID No. 21: WT BIS2 DNA (Catharanthus roseus)

SEQ ID No. 22: WT ORCA3 protein (Catharanthus roseus)

SEQ ID No. 23: WT ORCA3 DNA (Catharanthus roseus)

SEQ ID No. 24: AP2/ERF domain of ORCA3 (Catharanthus roseus)

SEQ ID No. 25: C-terminal domain of ORCA3 (Catharanthus roseus)

SEQ ID No. 26: NLS of CrORCA3 of ORCA3 (Catharanthus roseus)

SEQ ID No. 27: Acidic domain of ORCA3 (Catharanthus roseus)

SEQ ID No. 28: Serine-rich domain of ORCA3 (Catharanthus roseus)

SEQ ID No. 29: MtTSAR1 (Medicago truncatula)

SEQ ID No. 30: MtTSAR2 (Medicago truncatula)

SEQ ID No. 31: ORCA2 protein (Catharanthus roseus)

SEQ ID No. 32: WT AtMYC2 protein (Arabidopsis thaliana)

SEQ ID No. 33: Mutant AtMYC2 protein (Arabidopsis thaliana)

SEQ ID No. 34: JAZ interacting domain of CrMYC2aD126N (Catharanthus roseus)

SEQ ID No. 35: JAZ interacting domain of AtMYC2 (Catharanthus roseus)

SEQ ID No. 36: JAZ interacting domain of AtMYC3 (Catharanthus roseus)

SEQ ID No. 37: JAZ interacting domain of AtMYC4 (Catharanthus roseus)

SEQ ID No. 38: JAZ interacting domain of AtMYC5 (Catharanthus roseus)

SEQ ID No. 39: JAZ interacting domain of AtMYC2D105N (Catharanthus roseus)

SEQ ID No. 40: JAZ interacting domain of AtMYC3/atr2D (Catharanthus roseus)

References

Abraham, D.J., Farnsworth, N.R., 1969. Structure elucidation and chemistry of *Catharanthus alkaloids*. III. Structure of leurosine, an active anticancer alkaloid. J. Pharm. Sci. 58, 694-698.

Aerts, R.J., Gisi, D., De Carolis, E., De Luca, V., Baumann, T.W., 1994. Methyl jasmonate vapor increases the developmentally controlled synthesis of alkaloids in *Catharanthus* and *Cinchona* seedlings. Plant J. 5, 635–643.

Alberti, S., Gitler, A.D., Lindquist, S., 2007. A suite of Gateway® cloning vectors for high-throughput genetic analysis in *Saccharomyces cerevisiae*. Yeast. 24, 913-919.

Arendt, P., Miettinen, K., Pollier, J., De Rycke, R., Callewaert, N., Goossens, A., 2017. An endoplasmic reticulum-engineered yeast platform for overproduction of triterpenoids. Metab. Eng. 40, 165-175.

Asada, K., Salim, V., Masada-Atsumi, S., Edmunds, E., Nagatoshi, M., Terasaka, K., Mizukami, H., De Luca, V., 2013. A 7-deoxyloganetic acid glucosyltransferase contributes a key step in secologanin biosynthesis in Madagascar periwinkle. Plant Cell. 25, 4123-4134.

Besseau, S., Kellner, F., Lanoue, A., Thamm, A.M.K., Salim, V., Schneider, B., Geu-Flores, F., Höfer, R., Guirimand, G., Guihur, A., Oudin, A., Glevarec, G., Foureau, E., Papon, N., Clastre, M., Giglioli-Guivarc'h, N., St-Pierre, B., Werck-Reichhart, D., Burlat, V., De Luca, V., O'Connor, S.E., Courdavault, V., 2013. A pair of tabersonine 16-hydroxylases initiates the synthesis of vindoline in an organ-dependent manner in *Catharanthus roseus*. Plant Physiol. 163, 1792-1803.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., Benfey, P.N., 2003. A gene expression map of the *Arabidopsis* root. Science. 302, 1956-1960.

Brkljacic, J., Grotewold, E., 2017. Combinatorial control of plant gene expression. Biochim. Biophys. Acta - Gene Regul. Mech. 1860, 31-40.

Brown, S., Clastre, M., Courdavault, V., O'Connor, S.E., 2015. De novo production of the plant-derived alkaloid strictosidine in yeast. Proc. Natl. Acad. Sci. USA. 112, 3205-3210.

Chatel, G., Montiel, G., Pré, M., Memelink, J., Thiersault, M., Saint-Pierre, B., Doireau, P., Gantet, P., 2003. CrMYC1, a *Catharanthus roseus* elicitor- and jasmonate-responsive bHLH transcription factor that binds the G-box element of the strictosidine synthase gene promoter. J. Exp. Bot. 54, 2587-2588.

Chini, A., Gimenez-Ibanez, S., Goossens, A., Solano, R., 2016. Redundancy and specificity in jasmonate signalling. Curr. Opin. Plant Biol. 33, 147-156.

Courdavault, V., Papon, N., Clastre, M., Giglioli-Guivarc'h, N., St-Pierre, B., Burlat, V., 2014. A look inside an alkaloid multisite plant: the *Catharanthus* logistics. Curr. Opin. Plant Biol. 19C, 43-50.

Cuéllar Pérez, A., Pauwels, L., De Clercq, R., Goossens, A., 2013. Yeast two-hybrid analysis of jasmonate signaling proteins. Methods Mol. Biol. 1011, 173-185.

De Geyter, N., Gholami, A., Goormachtig, S., Goossens, A., 2012. Transcriptional machineries in jasmonate-elicited plant secondary metabolism. Trends Plant Sci. 17, 349-359.

Deepthi, S., Satheeshkumar, K., 2017. Cell line selection combined with jasmonic acid elicitation enhance camptothecin production in cell suspension cultures of *Ophiorrhiza mungos* L. Appl. Microbiol. Biotechnol. 101, 545-558.

Dugé de Bernonville, T., Carqueijeiro, I., Lanoue, A., Lafontaine, F., Sánchez Bel, P., Liesecke, F., Musset, K., Oudin, A., Glévarec, G., Pichon, O., Besseau, S., Clastre, M., St-Pierre, B., Flors, V., Maury, S., Huguet, E., O'Connor, S.E., Courdavault, V., 2017. Folivory elicits a strong defense reaction in *Catharanthus roseus*: metabolomic and transcriptomic analyses reveal distinct local and systemic responses. Sci. Rep. 7, 40453.

Dugé de Bernonville, T., Clastre, M., Besseau, S., Oudin, A., Burlat, V., Glévarec, G., Lanoue, A., Papon, N., Giglioli-Guivarc'h, N., St-Pierre, B., Courdavault, V., 2015. Phytochemical genomics of the Madagascar periwinkle: Unravelling the last twists of the alkaloid engine. Phytochemistry. 113, 9-23.

Eastman, P., 2011. Oncology drug shortages worsening, threatening patient care. Oncology Times. 33, 14-15.

Facchini, P.J., De Luca, V., 2008. Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. Plant J. 54, 763-784.

Frerigmann, H., Berger, B., Gigolashvili, T., 2014. bHLH05 is an interaction partner of MYB51 and a novel regulator of glucosinolate biosynthesis in Arabidopsis. Plant Physiol. 166, 349-369.

Gasperini, D., Chételat, A., Acosta, I.F., Goossens, J., Pauwels, L., Goossens, A., Dreos, R., Alfonso, E., Farmer, E.E., 2015. Multilayered organization of jasmonate signalling in the regulation of root growth. PLoS Genet. 11, e1005300.

Geu-Flores, F., Sherden, N.H., Courdavault, V., Burlat, V., Glenn, W.S., Wu, C., Nims, E., Cui, Y., O'Connor, S.E., 2012. An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. Nature. 492, 138-142.

Giddings, L.-A., Liscombe, D.K., Hamilton, J.P., Childs, K.L., DellaPenna, D., Buell, C.R., O'Connor, S.E., 2011. A stereoselective hydroxylation step of alkaloid biosynthesis by a unique cytochrome P450 in *Catharanthus roseus*. Journal of biological chemistry. 286, 16751-16757.

Goossens, J., Fernández-Calvo, P., Schweizer, F., Goossens, A., 2016. Jasmonates: signal transduction components and their roles in environmental stress responses. Plant Mol. Biol. 91, 673-689.

Goossens, J., Mertens, J., Goossens, A., 2017. Role and functioning of bHLH transcription factors in jasmonate signalling. J. Exp. Bot. 68, 1333-1347.

Goossens, J., Swinnen, G., Vanden Bossche, R., Pauwels, L., Goossens, A., 2015. Change of a conserved amino acid in the MYC2 and MYC3 transcription factors leads to release of JAZ repression and increased activity. New Phytol. 206, 1229-1237.

Guirimand, G., Guihur, A., Poutrain, P., Héricourt, F., Mahroug, S., St-Pierre, B., Burlat, V., Courdavault, V., 2011. Spatial organization of the vindoline biosynthetic pathway in *Catharanthus roseus*. J. Plant Physiol. 168, 549-557.

Gundlach, H., Müller, M.J., Kutchan, T.M., Zenk, M.H., 1992. Jasmonic acid is a signal transducer in elicitor-induced plant cell cultures. Proc. Natl. Acad. Sci. USA. 89, 2389-2393.

Heim, M.A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B., Bailey, P.C., 2003. The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. Mol. Biol. Evol. 20, 735-747.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., Vandesompele, J., 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol. 8, R19.

Karimi, M., Inzé, D., Depicker, A., 2002. GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation. Trends Plant Sci. 7, 193-195.

Kazan, K., Manners, J.M., 2012. JAZ repressors and the orchestration of phytohormone crosstalk. Trends Plant Sci. 17, 22-31.

Laflamme, P., St-Pierre, B., De Luca, V., 2001. Molecular and biochemical analysis of a Madagascar periwinkle root-specific minovincinine-19-hydroxy-*O*-acetyltransferase. Plant Physiol. 125, 189-198.

Larsen, B., Fuller, V.L., Pollier, J., Van Moerkercke, A., Schweizer, F., Payne, R., Colinas, M., O'Connor, S.E., Goossens, A., Halkier, B.A., 2017. Identification of iridoid glucoside transporters in *Catharanthus roseus*. Plant Cell Physiol. 58, 1507-1518.

Levac, D., Murata, J., Kim, W.S., De Luca, V., 2008. Application of carborundum abrasion for investigating the leaf epidermis: molecular cloning of *Catharanthus roseus* 16-hydroxytabersonine-16-*O*-methyltransferase. Plant J. 53, 225-236.

Liscombe, D.K., Usera, A.R., O'Connor, S.E., 2010. Homolog of tocopherol C methyltransferases catalyzes *N* methylation in anticancer alkaloid biosynthesis. Proc. Natl. Acad. Sci. USA. 107, 18793-13798.

Memelink, J., 2009. Regulation of gene expression by jasmonate hormones. Phytochemistry. 70, 1560-1570.

Menke, F.L.H., Champion, A., Kijne, J.W., 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.

Miettinen, K., Dong, L., Navrot, N., Schneider, T., Burlat, V., Pollier, J., Woittiez, L., van der Krol, S., Lugan, R., Ilc, T., Verpoorte, R., Oksman-Caldentey, K.-M., Martinoia, E., Bouwmeester, H., Goossens, A., Memelink, J., Werck-Reichhart, D., 2014. The seco-iridoid pathway from *Catharanthus roseus*. Nat. Commun. 5, 3606.

Miettinen, K., Pollier, J., Buyst, D., Arendt, P., Csuk, R., Sommerwerk, S., Moses, T., Mertens, J., Sonawane, P.D., Pauwels, L., Aharoni, A., Martins, J., Nelson, D.R., Goossens, A., 2017. The ancient CYP716 family is a major contributor to the diversification of eudicot triterpenoid biosynthesis. Nat. Commun. 8, 14153.

Montiel, G., Zarei, A., Körbes, A.P., Memelink, J., 2011. The jasmonate-responsive element from the *ORCA3* promoter from *Catharanthus roseus* is active in Arabidopsis and is controlled by the transcription factor AtMYC2. Plant Cell Physiol. 52, 578-587.

O'Connor, S.E., 2015. Engineering of secondary metabolism. Annu. Rev. Genet. 49, 71-94.

Pan, Q., Saiman, M.Z., Mustafa, N.R., Verpoorte, R., Tang, K., 2016. A simple and rapid HPLC-DAD method for simultaneously monitoring the accumulation of alkaloids and precursors in different parts and different developmental stages of *Catharanthus roseus* plants. J. Chromatogr. B. 1014, 10-16.

Patra, B., Pattanaik, S., Schluttenhofer, C., Yuan, L., 2018. A network of jasmonate-responsive bHLH factors modulate monoterpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. New Phytol., in press.

Paul, P., Singh, S.K., Patra, B., Sui, X., Pattanaik, S., Yuan, L., 2017. A differentially regulated AP2/ERF transcription factor gene cluster acts downstream of a MAP kinase cascade to modulate terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. New Phytol. 213, 1107-1123.

Pauwels, L., Goossens, A., 2011. The JAZ proteins: a crucial interface in the jasmonate signaling cascade. Plant Cell. 23, 3089-3100.

Pauwels, L., Inzé, D., Goossens, A., 2009. Jasmonate-inducible gene: what does it mean? Trends Plant Sci. 14, 87-91.

Payne, R.M.E., Xu, D., Foureau, E., Teto Carqueijeiro, M.I.S., Oudin, A., Dugé de Bernonville, T., Novak, V., Burow, M., Olsen, C.-E., Jones, D.M., Tatsis, E.C., Pendle, A., Halkier, B.A., Geu-Flores, F., Courdavault, V., Nour-Eldin, H.H., O'Connor, S.E., 2017. An NPF transporter exports a central monoterpene indole alkaloid intermediate from the vacuole. Nat. Plants. 3, 16208.

Pollier, J., Vanden Bossche, R., Rischer, H., Goossens, A., 2014. Selection and validation of reference genes for transcript normalization in gene expression studies in *Catharanthus roseus*. Plant Physiol. Biochem. 83C, 20-25.

Pré, M., Sibéril, Y., Memelink, J., Champion, A., Doireau, P., Gantet, P., 2000. Isolation by the yeast one-hybrid system of cDNAs encoding transcription factors that bind to the G-box element of the *strictosidine synthase* gene promoter from *Catharanthus roseus*. Int. J. Biochromatogr. 5, 229-244.

Qu, Y., Easson, M.L.A.E., Froese, J., Simionescu, R., Hudlicky, T., De Luca, V., 2015. Completion of the seven-step pathway from tabersonine to the anticancer drug precursor vindoline and its assembly in yeast. Proc. Natl. Acad. Sci. USA. 112, 6224-6229.

Rodriguez, S., Compagnon, V., Crouch, N.P., St-Pierre, B., De Luca, V., 2003. Jasmonate-induced epoxidation of tabersonine by a cytochrome P-450 in hairy root cultures of *Catharanthus roseus*. Phytochemistry. 64, 401-409.

Roepke, J., Salim, V., Wu, M., Thamm, A.M.K., Murata, J., Ploss, K., Boland, W., De Luca, V., 2010. Vinca drug components accumulate exclusively in leaf exudates of Madagascar periwinkle. Proc. Natl. Acad. Sci. USA. 107, 15287-15292.

Salim, V., Wiens, B., Masada-Atsumi, S., Yu, F., De Luca, V., 2014. 7-deoxyloganetic acid synthase catalyzes a key 3 step oxidation to form 7-deoxyloganetic acid in *Catharanthus roseus* iridoid biosynthesis. Phytochemistry. 101, 23-31.

Simkin, A.J., Miettinen, K., Claudel, P., Burlat, V., Guirimand, G., Courdavault, V., Papon, N., Meyer, S., Godet, S., St-Pierre, B., Giglioli-Guivarc'h, N., Fischer, M.J.C., Memelink, J., Clastre, M., 2013. Characterization of the plastidial geraniol synthase from Madagascar periwinkle which initiates the

monoterpenoid branch of the alkaloid pathway in internal phloem associated parenchyma. Phytochemistry. 85, 36-43.

Smolen, G.A., Pawlowski, L., Wilensky, S.E., Bender, J., 2002. Dominant alleles of the basic helix-loop-helix transcription factor ATR2 activate stress-responsive genes in Arabidopsis. Genetics. 161, 1235-1246.

St-Pierre, B., Vazquez-Flota, F.A., De Luca, V., 1999. Multicellular compartmentation of *Catharanthus roseus* alkaloid biosynthesis predicts intercellular translocation of a pathway intermediate. Plant Cell. 11, 887-900.

Stavrinides, A., Tatsis, E.C., Caputi, L., Foureau, E., Stevenson, C.E.M., Lawson, D.M., Courdavault, V., O'Connor, S.E., 2016. Structural investigation of heteroyohimbine alkaloid synthesis reveals active site elements that control stereoselectivity. Nat. Commun. 7, 12116.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30, 2725-2729.

Tatsis, E.C., Carqueijeiro, I., Dugé de Bernonville, T., Franke, J., Dang, T.-T., Oudin, A., Lanoue, A., Lafontaine, F., Stavrinides, A.K., Clastre, M., Courdavault, V., O'Connor, S.E., 2017. A three enzyme system to generate the *Strychnos* alkaloid scaffold from a central biosynthetic intermediate. Nat. Commun. 8, 316.

van der Fits, L., Memelink, J., 2000. ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. Science. 289, 295-297.

van der Heijden, R., Jacobs, D.I., Snoeijer, W., Hallard, D., Verpoorte, R., 2004. The *Catharanthus* alkaloids: pharmacognosy and biotechnology. Curr. Med. Chem. 11, 607-628.

Van Moerkercke, A., Steensma, P., Gariboldi, I., Espoz, J., Purnama, P.C., Schweizer, F., Miettinen, K., Vanden Bossche, R., De Clercq, R., Memelink, J., Goossens, A., 2016. The basic helix-loop-helix transcription factor BIS2 is essential for monoterpenoid indole alkaloid production in the medicinal plant *Catharanthus roseus*. Plant J. 88, 3-12.

Van Moerkercke, A., Steensma, P., Schweizer, F., Pollier, J., Gariboldi, I., Payne, R., Vanden Bossche, R., Miettinen, K., Espoz, J., Purnama, P.C., Kellner, F., Seppänen-Laakso, T., O'Connor, S.E., Rischer, H., Memelink, J., Goossens, A., 2015. The bHLH transcription factor BIS1 controls the iridoid branch of the monoterpenoid indole alkaloid pathway in *Catharanthus roseus*. Proc. Natl. Acad. Sci. USA. 112, 8130-8135.

Vanden Bossche, R., Demedts, B., Vanderhaeghen, R., Goossens, A., 2013. Transient expression assays in tobacco protoplasts. Methods Mol. Biol. 1011, 227-239.

Vom Endt, D., Soares e Silva, M., Kijne, J.W., Pasquali, G., Memelink, J., 2007. Identification of a bipartite jasmonate-responsive promoter element in the *Catharanthus roseus ORCA3* transcription factor gene that interacts specifically with AT-hook DNA-binding proteins. Plant Physiol. 144, 1680-1689.

Wasternack, C., Hause, B., 2013. Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in *Annals of Botany*. Ann. Bot. 111, 1021-1058.

Wasternack, C., Strnad, M., 2018. Jasmonates are signals in the biosynthesis of secondary metabolites — Pathways, transcription factors and applied aspects — A brief review. New Biotechnol., in press.

Xia, J., Wishart, D.S., 2016. Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. Curr. Protoc. Bioinform. 55, 14.10. 1-14.10.91.

Zhang, F., Yao, J., Ke, J., Zhang, L., Lam, V.Q., Xin, X.-F., Zhou, X.E., Chen, J., Brunzelle, J., Griffin, P.R., Zhou, M., Xu, H.E., Melcher, K., He, S.Y., 2015. Structural basis of JAZ repression of MYC transcription factors in jasmonate signalling. Nature. 525, 269-273.

Zhang, H., Hedhili, S., Montiel, G., Zhang, Y., Chatel, G., Pré, M., Gantet, P., 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.

Zhou, M., Memelink, J., 2016. Jasmonate-responsive transcription factors regulating plant secondary metabolism. Biotechnol. Adv. 34, 441-449.

CLAIMS

1. A *Catharanthus roseus* MYC2 protein with a D to N mutation at a position relative to position 126 of SEQ ID No. 1.
2. The MYC2 protein according to claim 1 further comprising a JAZ interacting domain comprising SEQ ID No. 7 and/or SEQ ID No. 8.
3. The MYC2 protein according to claim 1 or 2, wherein said protein has at least 90% homology to SEQ ID No. 1.
4. A mutant *Catharanthus roseus* MYC2a protein as depicted in SEQ ID No. 2.
5. A nucleic acid molecule encoding the MYC2 or MYC2a protein according to any of claims 1-4.
6. A plant cell expressing the nucleic acid molecule of claim 5 or comprising the MYC2 or MYC2a protein according to any of claims 1-4.
7. The plant cell of claim 6 further expressing the transcriptional activators BIS1 and ORCA3.
8. The plant cell of claim 7 wherein the expression level of BIS1 and ORCA3 is at least 10% higher than the corresponding expression level in a control plant cell.
9. A regulatory module comprising the MYC2 or MYC2a protein of any of claims 1-4, further comprising the transcription factors BIS1 and ORCA3, said BIS1 and ORCA3 transcription factors are transcriptionally controlled by promoters different to the promoters to which they are naturally linked.
10. Use of the MYC2 or MYC2a protein according to any of claims 1-4 or the nucleic acid molecule of claim 5 or the regulatory module of claim 9 to produce at least one monoterpenoid indole alkaloid in a host cell.
11. Use of the plant cell according to claim 7 or 8 to enhance the production of at least one monoterpenoid indole alkaloid with at least 10% compared to a control plant cell.
12. The use according to claim 10 or 11 wherein the at least one monoterpenoid indole alkaloid is a 16-hydroxytabersonine-derived MIA and/or a 19-hydroxytabersonine-derived MIA.
13. The use according to any of claims 10-12 wherein the at least one monoterpenoid indole alkaloid is selected from the list consisting of strictosidine, catharanthine, 16-hydroxytabersonine, 16-methoxylhörhammericine, hörhammericine, serpentine, vindoline, vinblastine, akuammicine, isositsirikine, 16-methoxytabersonine, 3-hydroxy-16-methoxy-2,3-dihydrotabersonine, 19-hydroxytabersonine, 16-hydroxyhörhammericine, minovincinine, vincadifformine, 16-hydroxyvincadifformine, 16-hydroxy-19-O-acetyltabersonine, perivine and O-acetylstemmadine.
14. A method of producing at least one monoterpenoid indole alkaloid in a host cell, comprising:
    a. Expressing the MYC2 or MYC2a protein of any of claims 1-4 in said host cell;
    b. Optionally further expressing the BIS1 and/or ORCA3 transcription factors;

c. Selecting host cells that produce said at least one monoterpenoid indole alkaloid.

15. Method of increasing the production of at least one monoterpenoid indole alkaloid in a plant cell, comprising:

a. Expressing the MYC2 or MYC2a protein of any of claims 1-4 in said plant cell;

b. Optionally further expressing the BIS1 and/or ORCA3 transcription factors;

c. Selecting plant cells with an at least 10% increased production of said at least one monoterpenoid indole alkaloid compared to control plant cells not expressing the MYC2 or MYC2a protein of any of claims 1-4.

Figure 1

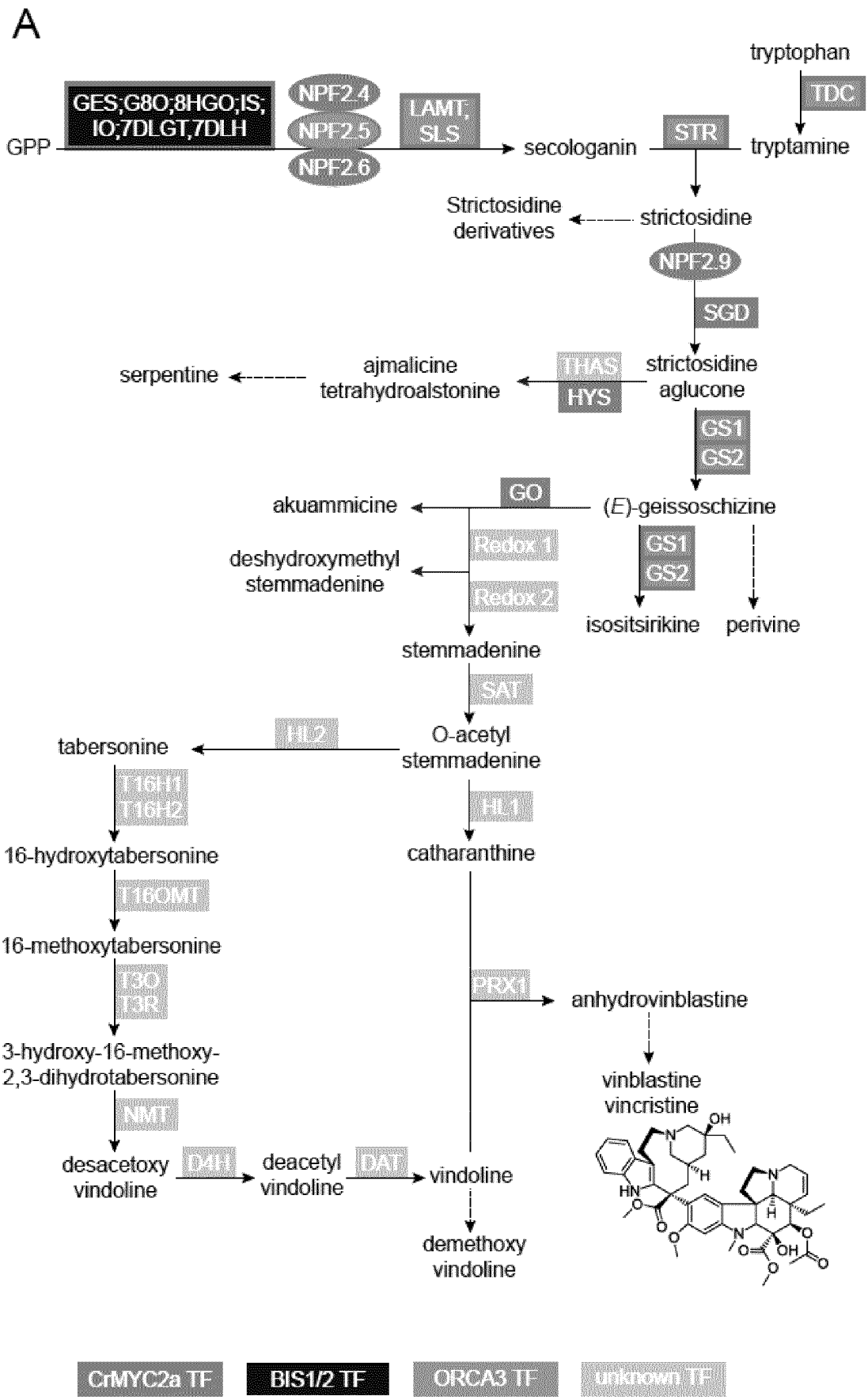

A
tryptophan
TDC
GES;G8O;8HGO;IS;
IO;7DLGT;7DLH
NPF2.4
NPF2.5
NPF2.6
LAMT,
SLS
GPP
secologanin
STR
tryptamine
Strictosidine
derivatives
strictosidine
NPF2.9
SGD
serpentine
ajmalicine
tetrahydroalstonine
THAS
HYS
strictosidine
aglucone
GS1
GS2
akuammicine
GO
(E)-geissoschizine
Redox 1
deshydroxymethyl
stemmadenine
Redox 2
GS1
GS2
isositsirikine
perivine
stemmadenine
SAT
O-acetyl
stemmadenine
tabersonine
T16H1
T16H2
16-hydroxytabersonine
T16OMT
16-methoxytabersonine
T3O
T3R
3-hydroxy-16-methoxy-
2,3-dihydrotabersonine
NMT
desacetoxy
vindoline
D4H
deacetyl
vindoline
DAT
vindoline
demethoxy
vindoline
HL2
HL1
catharanthine
PRX1
anhydrovinblastine
vinblastine
vincristine
CrMYC2a TF
BIS1/2 TF
ORCA3 TF
unknown TF

Figure 1 continued

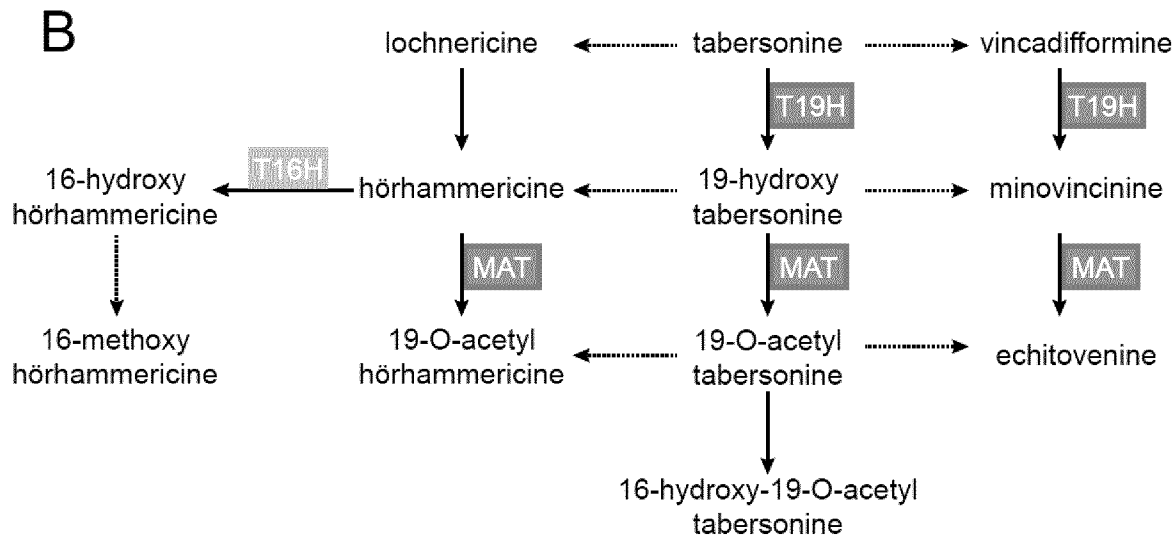


Figure 2

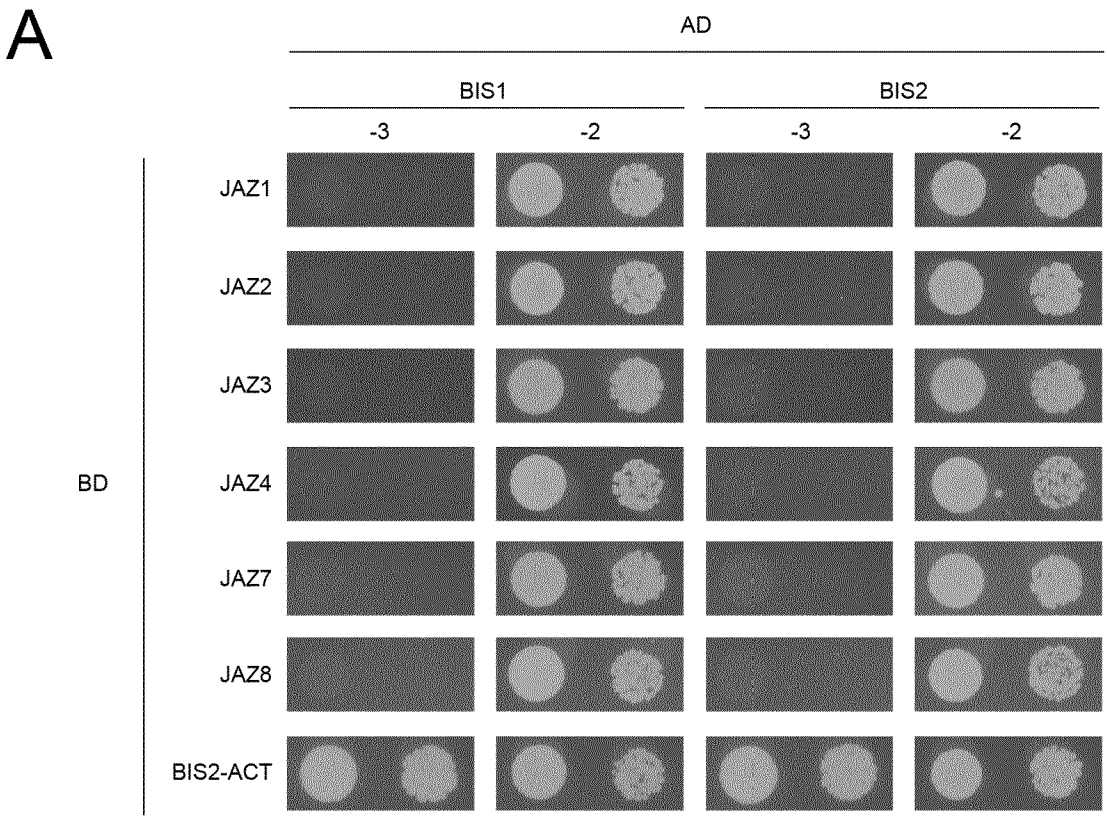
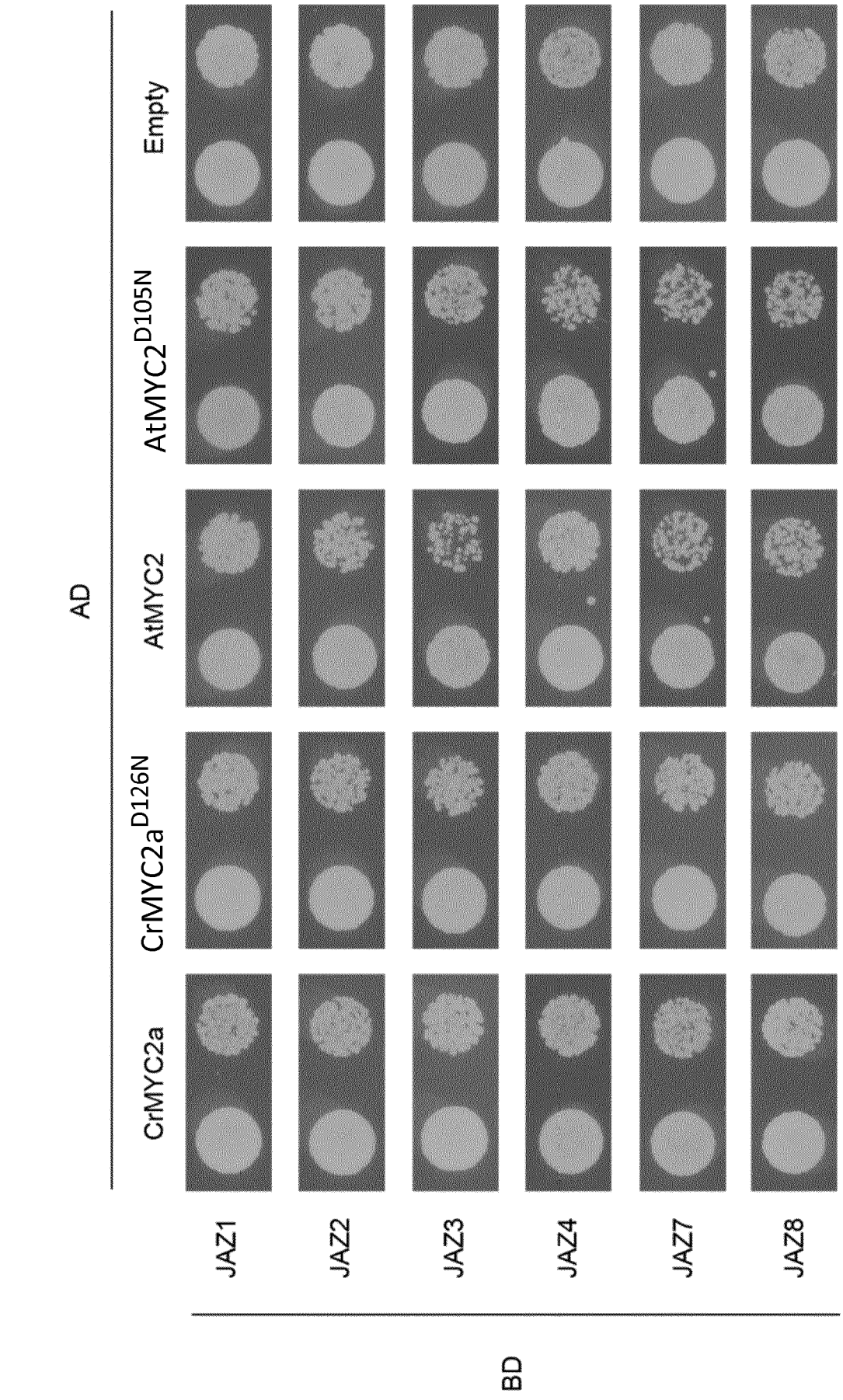


Figure 2 continued



B

Figure 3

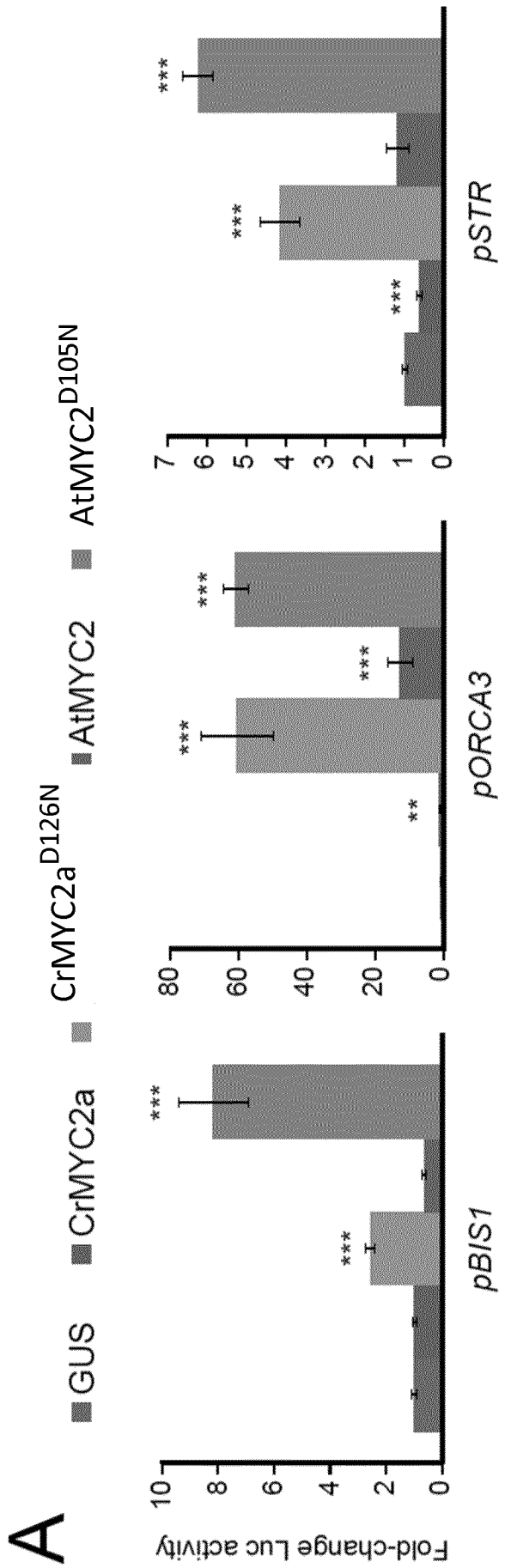

A
Fold-change Luc activity
GUS
CrMYC2a
CrMYC2a^D126N
AtMYC2
AtMYC2^D105N
pBIS1
pORCA3
pSTR

Figure 3 continued

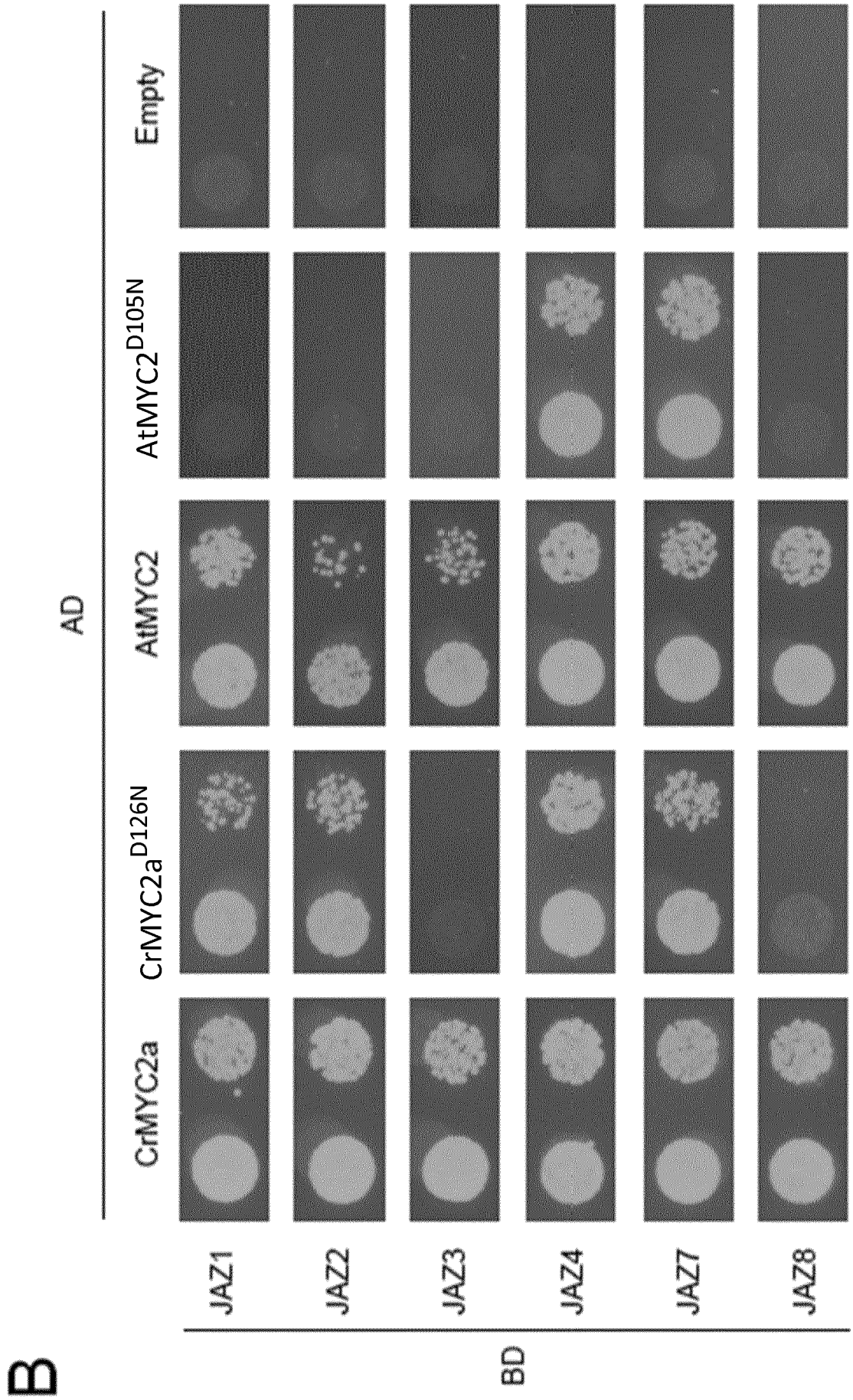


Figure 3 continued

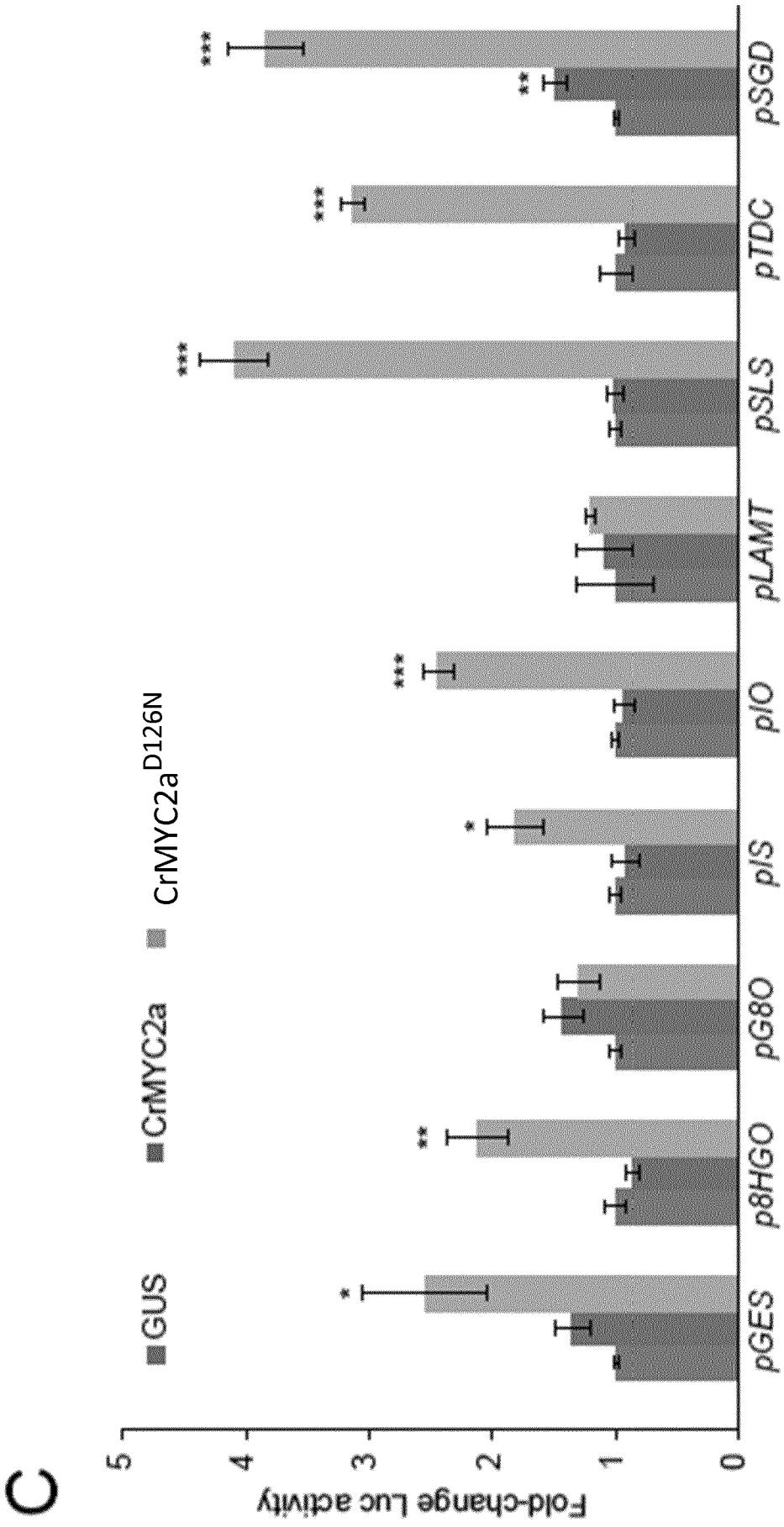
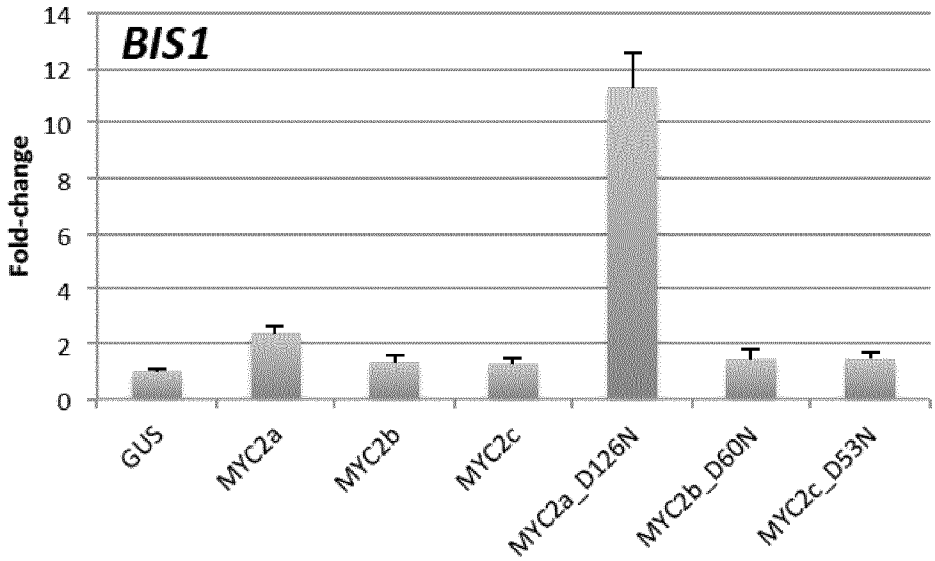


Figure 3 continued

D



E

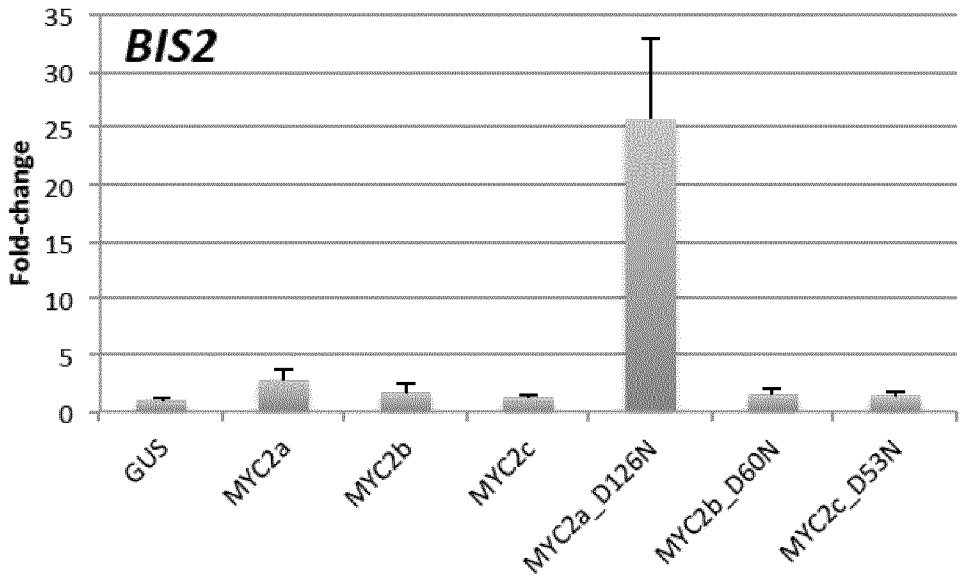
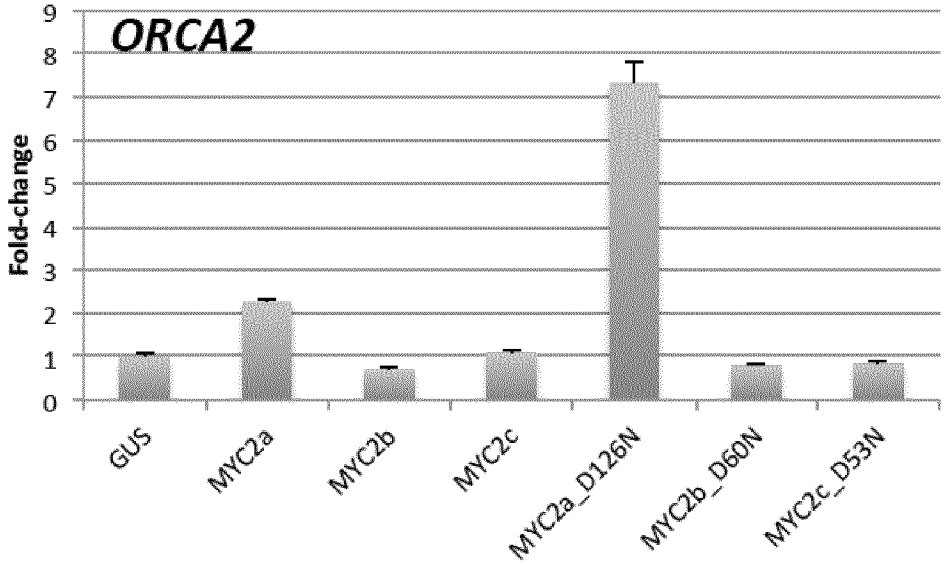


Figure 3 continued

F



G

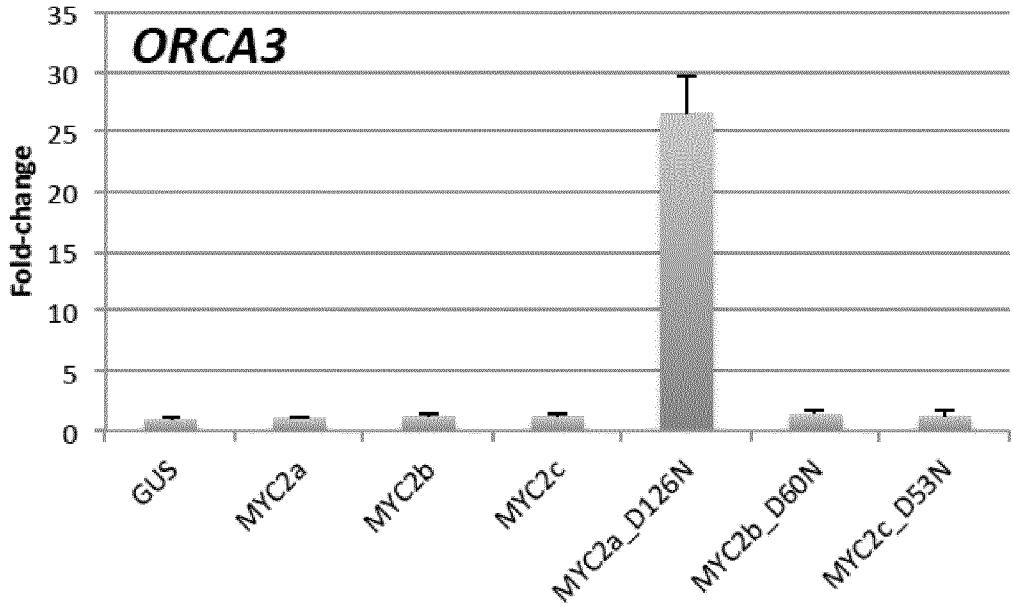


Figure 4

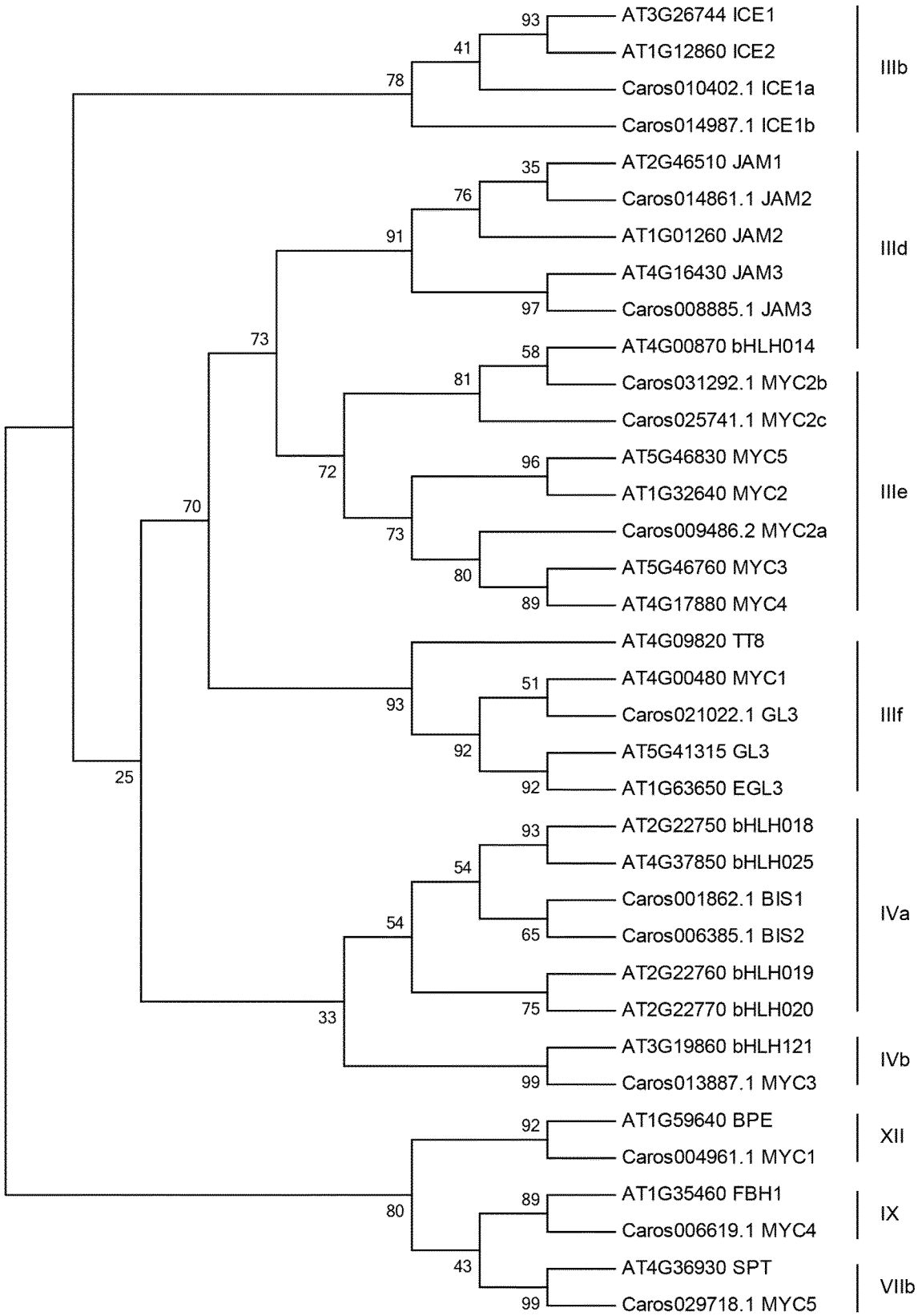

AT3G26744 ICE1
AT1G12860 ICE2
Caros010402.1 ICE1a
Caros014987.1 ICE1b
IIIb
AT2G46510 JAM1
Caros014861.1 JAM2
AT1G01260 JAM2
AT4G16430 JAM3
Caros008885.1 JAM3
IIId
AT4G00870 bHLH014
Caros031292.1 MYC2b
Caros025741.1 MYC2c
AT5G46830 MYC5
AT1G32640 MYC2
Caros009486.2 MYC2a
AT5G46760 MYC3
AT4G17880 MYC4
IIIe
AT4G09820 TT8
AT4G00480 MYC1
Caros021022.1 GL3
AT5G41315 GL3
AT1G63650 EGL3
IIIf
AT2G22750 bHLH018
AT4G37850 bHLH025
Caros001862.1 BIS1
Caros006385.1 BIS2
AT2G22760 bHLH019
AT2G22770 bHLH020
IVa
AT3G19860 bHLH121
Caros013887.1 MYC3
IVb
AT1G59640 BPE
Caros004961.1 MYC1
XII
AT1G35460 FBH1
Caros006619.1 MYC4
IX
AT4G36930 SPT
Caros029718.1 MYC5
VIIb
93
41
78
35
76
91
97
58
81
96
72
70
73
80
89
51
93
92
92
25
93
54
54
65
75
33
99
92
89
80
43
99

Figure 5

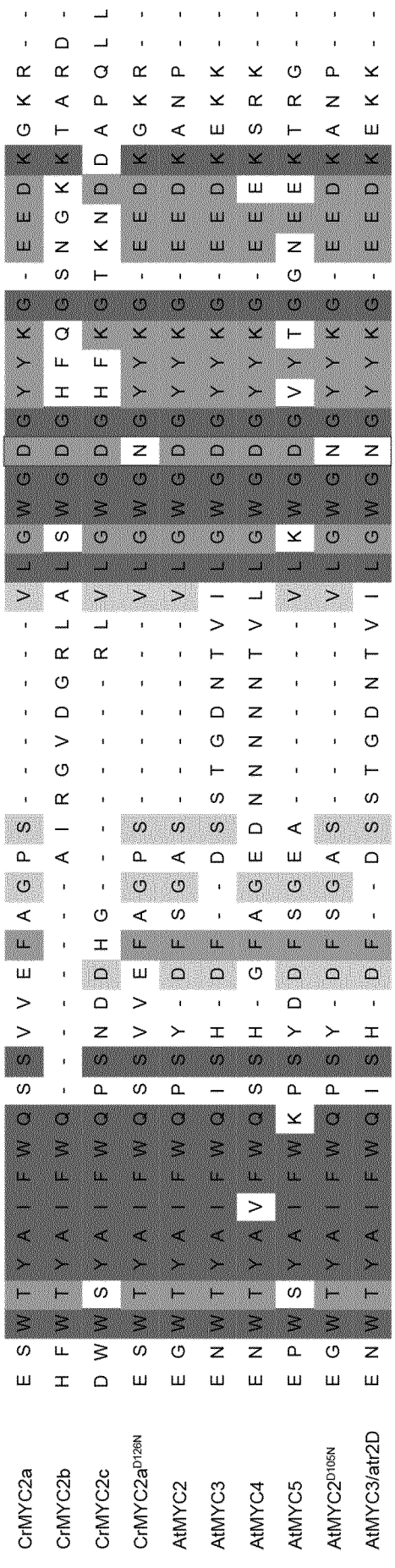

CrMYC2a
CrMYC2b
CrMYC2c
CrMYC2a^D126N
AtMYC2
AtMYC3
AtMYC4
AtMYC5
AtMYC2^D105N
AtMYC3/atr2D

Figure 6

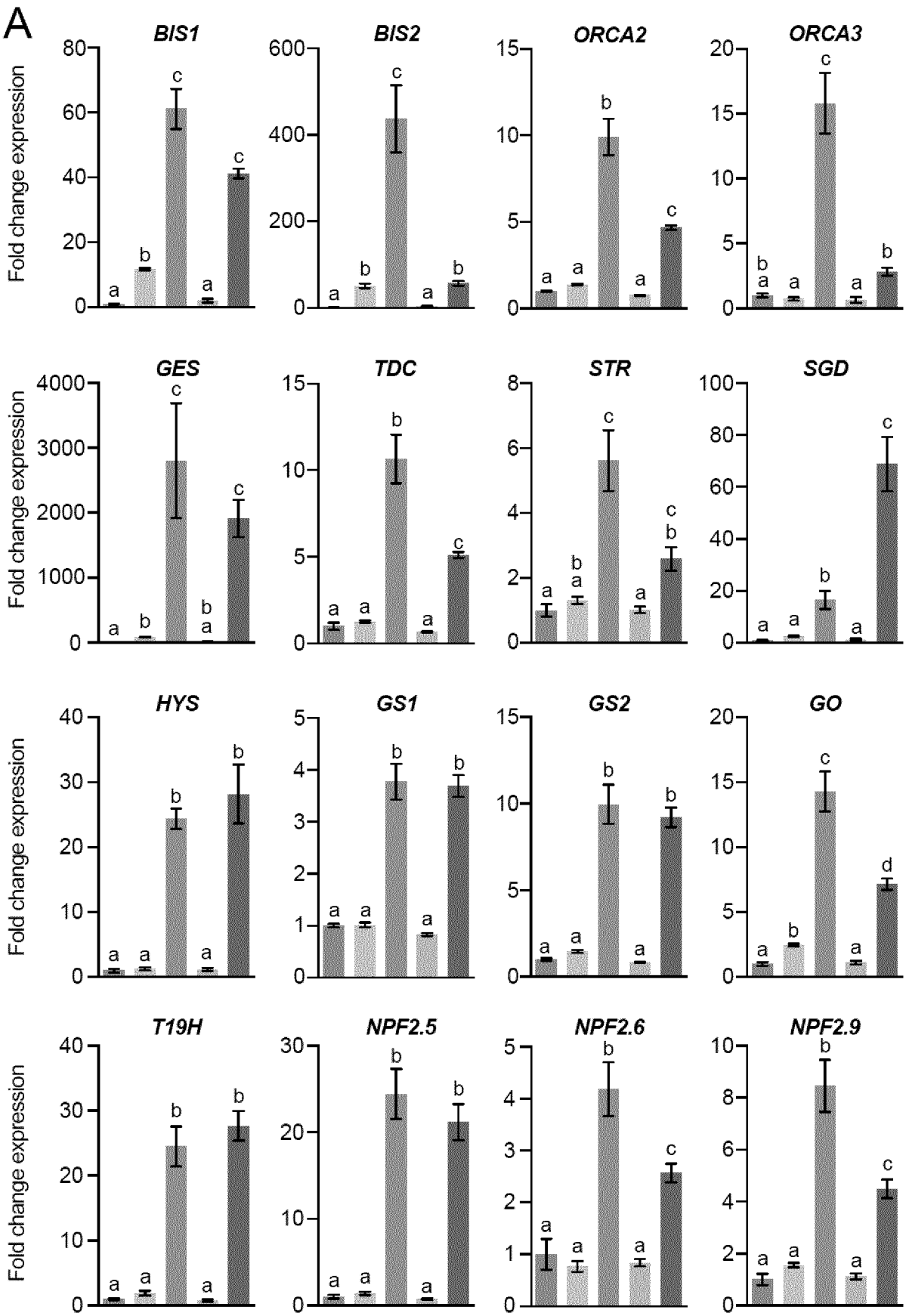

A
BIS1
BIS2
ORCA2
ORCA3
GES
TDC
STR
SGD
HYS
GS1
GS2
GO
T19H
NPF2.5
NPF2.6
NPF2.9
Fold change expression

Figure 6 continued

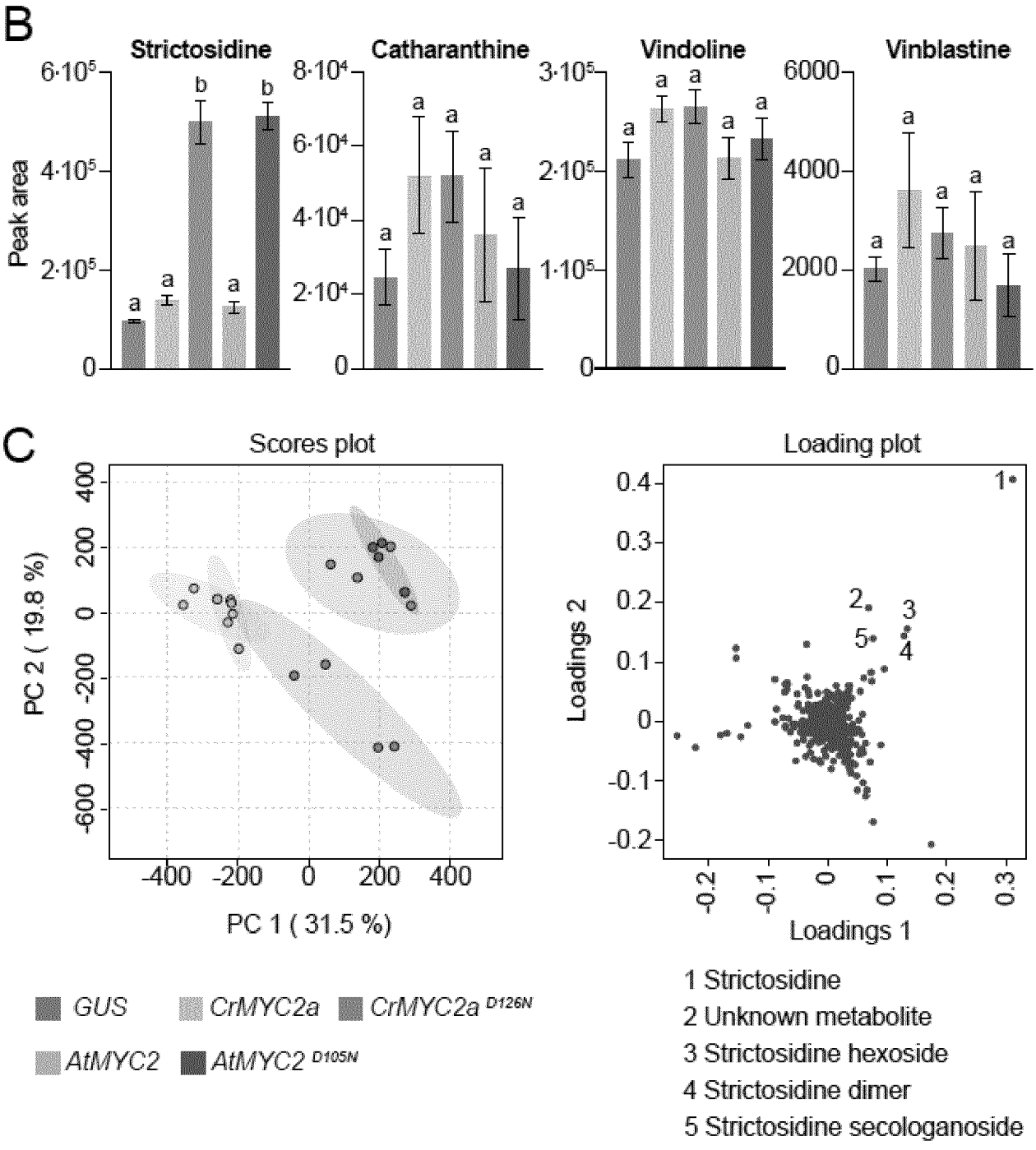


Figure 7

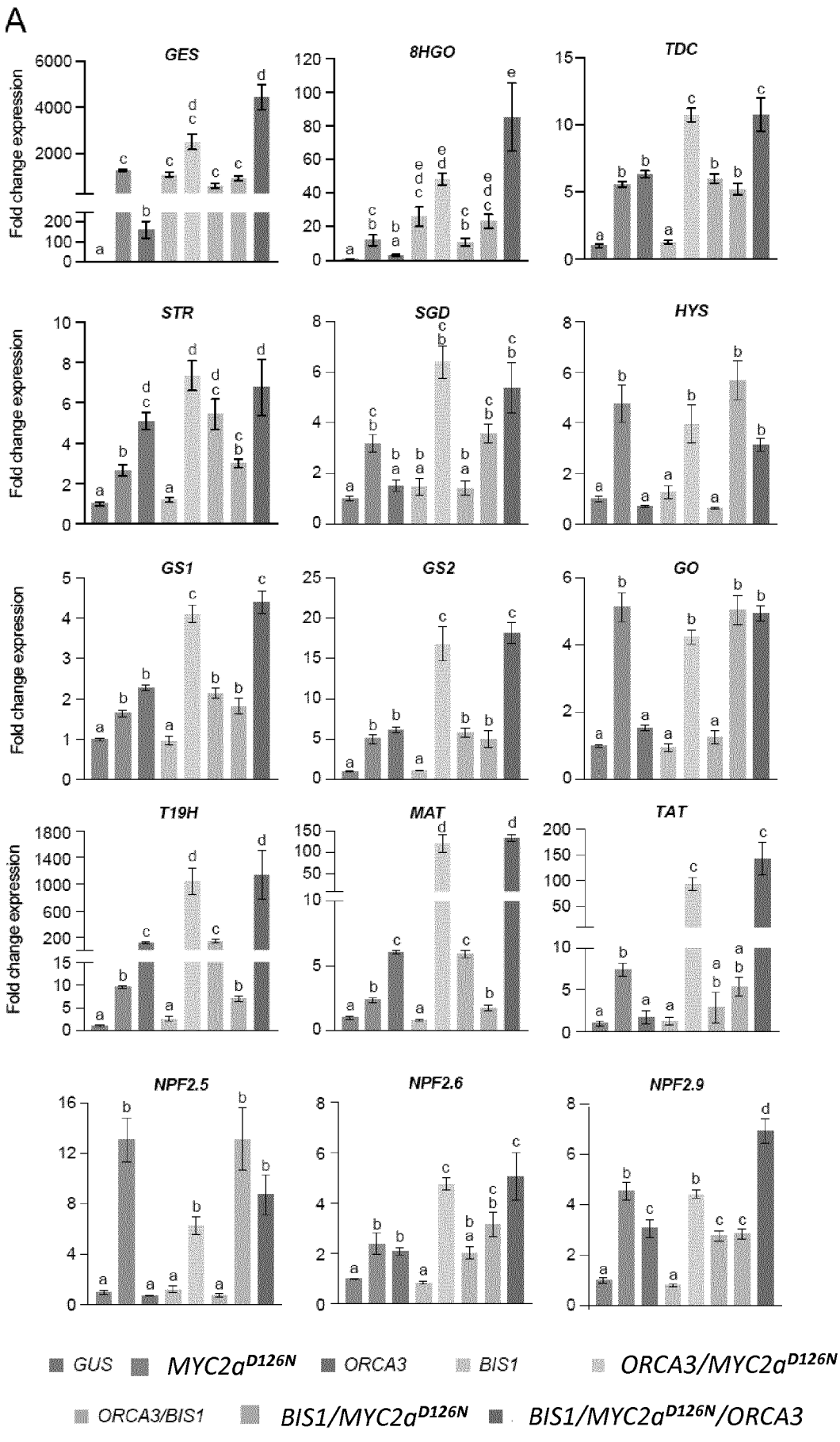

A
GES
8HGO
TDC
STR
SGD
HYS
GS1
GS2
GO
T19H
MAT
TAT
NPF2.5
NPF2.6
NPF2.9
Fold change expression
GUS
MYC2a^D126N
ORCA3
BIS1
ORCA3/MYC2a^D126N
ORCA3/BIS1
BIS1/MYC2a^D126N
BIS1/MYC2a^D126N/ORCA3

Figure 7 continued

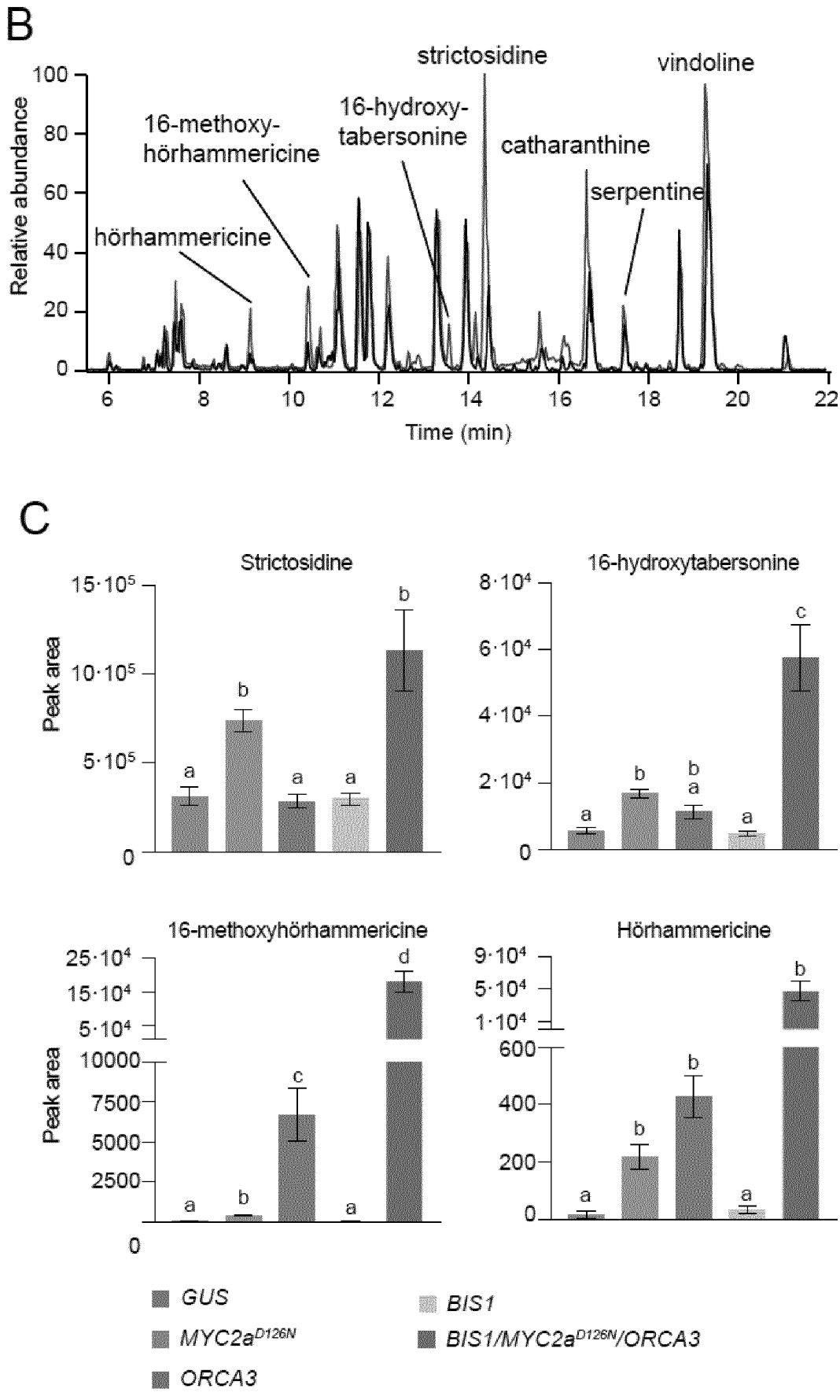


Figure 7 continued

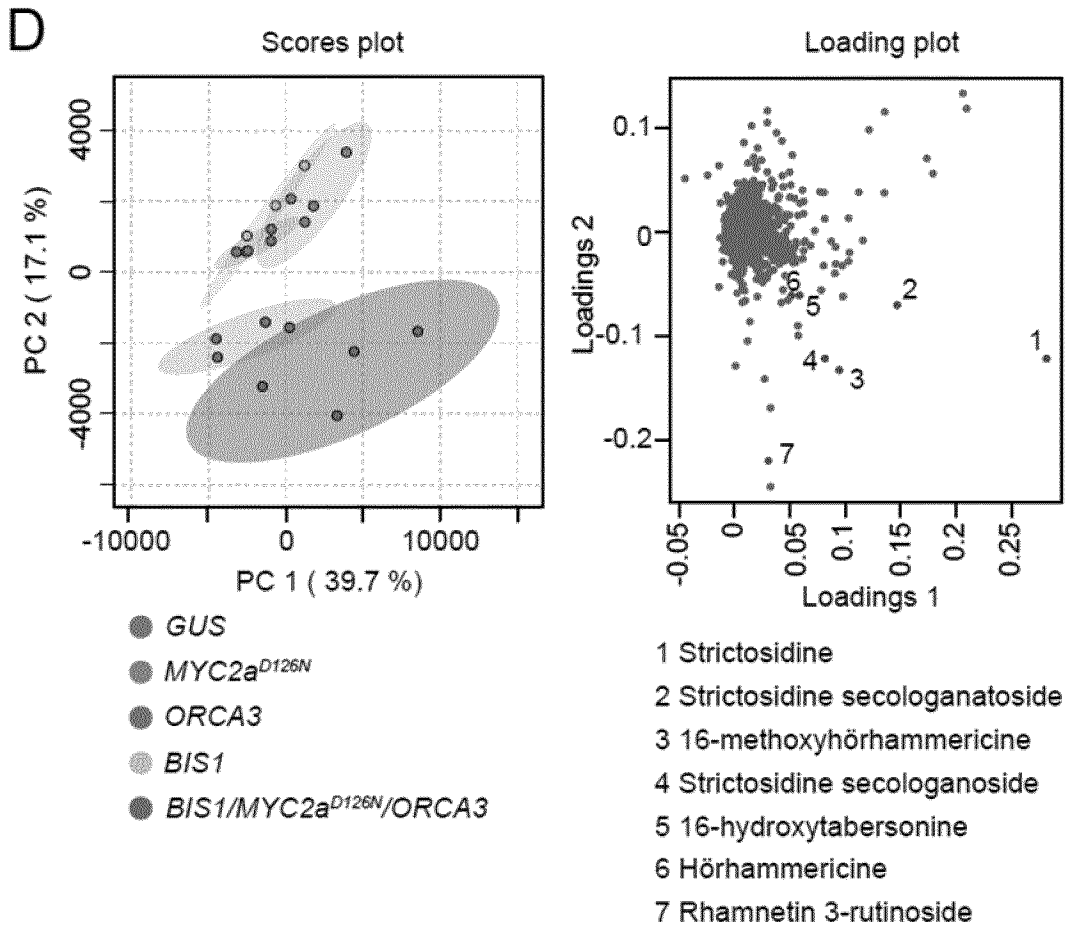


Figure 8

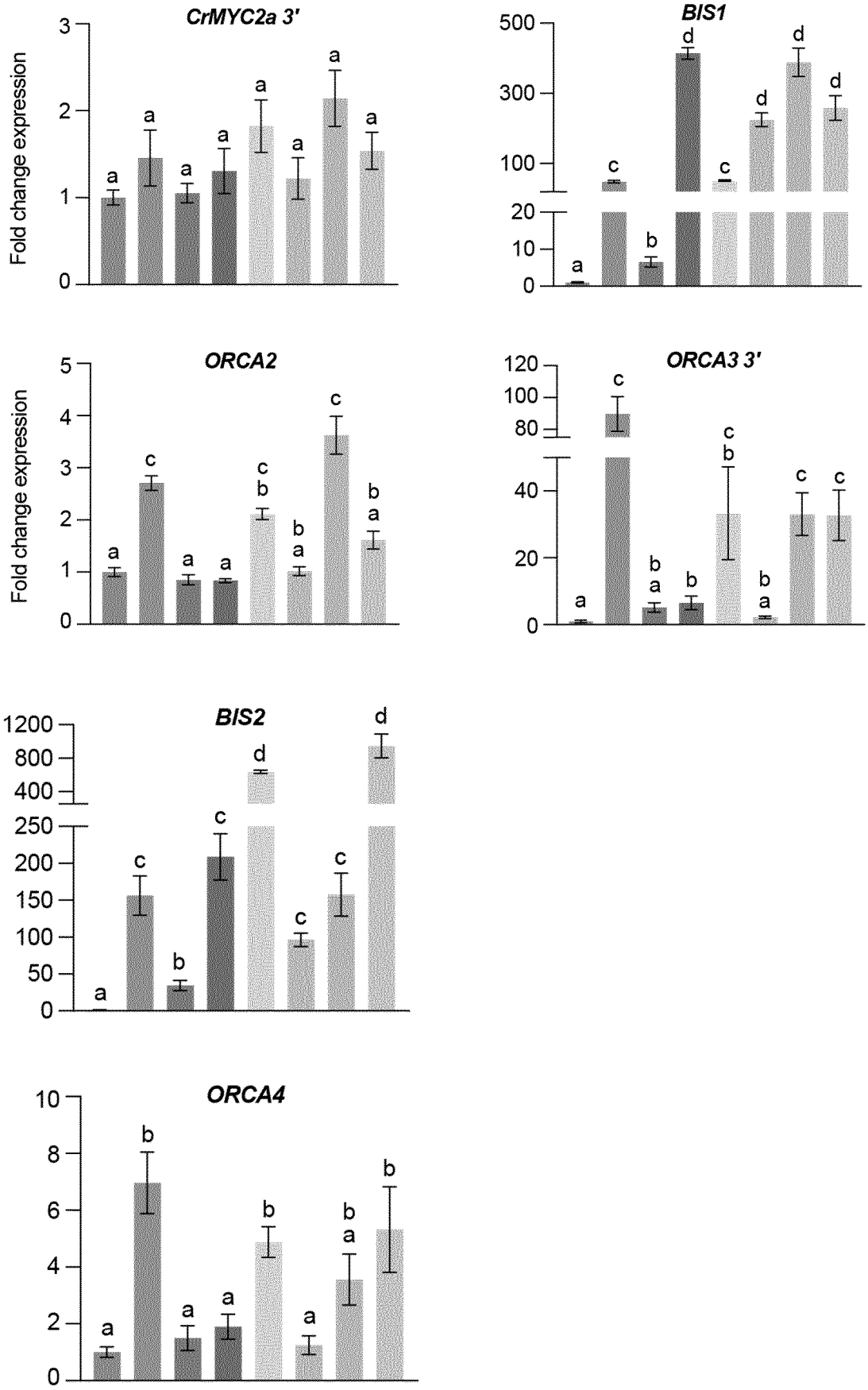

CrMYC2a 3'
Fold change expression
BIS1
ORCA2
Fold change expression
ORCA3 3'
BIS2
ORCA4

Figure 8 continued

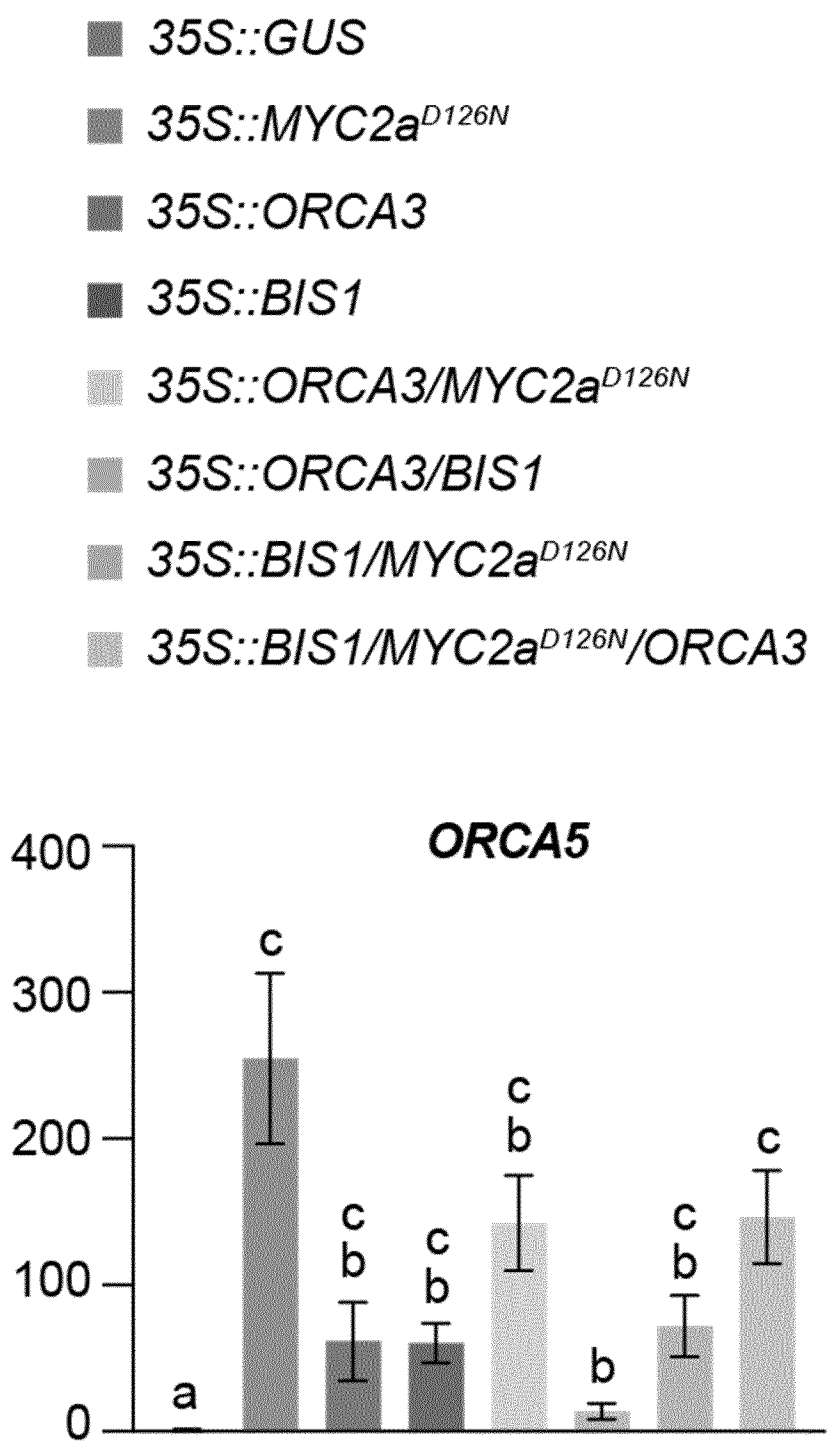
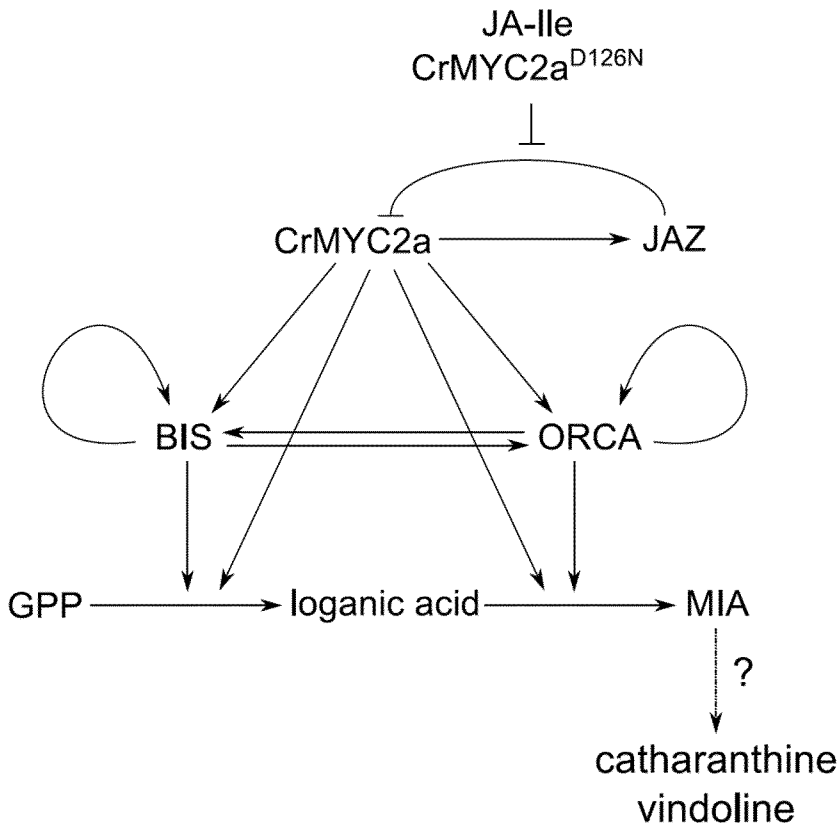


Figure 9


JA-Ile
CrMYC2a^D126N
CrMYC2a
JAZ
BIS
ORCA
GPP
loganic acid
MIA
?
catharanthine
vindoline

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2019/062848

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/415 C12N15/82
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2018/039590 A1 (UNIV MICHIGAN STATE [US]; HOWE GREGG A [US] ET AL.) 1 March 2018 (2018-03-01) page 34, line 12 - page 36, line 4 -----	1-8
Y	JONAS GOOSSENS ET AL: "Change of a conserved amino acid in the MYC2 and MYC3 transcription factors leads to release of JAZ repression and increased activity", NEW PHYTOLOGIST, vol. 206, no. 4, 1 June 2015 (2015-06-01), pages 1229-1237, XP055597086, GB ISSN: 0028-646X, DOI: 10.1111/nph.13398 cited in the application the whole document -----	1-8, 10-15
	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 June 2019	01/07/2019

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Kania, Thomas

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ZHANG HONGTAO ET AL: "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 JOURNAL, BLACKWELL SCIENCE, OXFORD, GB, vol. 67, no. 1, 1 July 2011 (2011-07-01), pages 61-71, XP002738507, ISSN: 1365-313X, DOI: 10.1111/J.1365-313X.2011.04575.X cited in the application the whole document	1-8, 10-15
A	NATHAN DE GEYTER ET AL: "Transcriptional machineries in jasmonate-elicited plant secondary metabolism", TRENDS IN PLANT SCIENCE, vol. 17, no. 6, 9 March 2012 (2012-03-09), pages 349-359, XP028520086, ISSN: 1360-1385, DOI: 10.1016/J.TPLANTS.2012.03.001 [retrieved on 2012-03-09] cited in the application	1-15
A	PAN QIFANG ET AL: "Monoterpenoid indole alkaloids biosynthesis and its regulation inCatharanthus roseus: a literature review from genes to metabolites", PHYTOCHEMISTRY REVIEWS, KLUWER, NL, vol. 15, no. 2, 1 April 2015 (2015-04-01), pages 221-250, XP035928057, ISSN: 1568-7767, DOI: 10.1007/S11101-015-9406-4 [retrieved on 2015-04-01]	1-15
T	SCHWEIZER FABIAN ET AL: "An engineered combinatorial module of transcription factors boosts production of monoterpenoid indole alkaloids inCatharanthus roseus", METABOLIC ENGINEERING, ACADEMIC PRESS, US, vol. 48, 28 May 2018 (2018-05-28), pages 150-162, XP085437894, ISSN: 1096-7176, DOI: 10.1016/J.YMBEN.2018.05.016 the whole document	

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2019/062848

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2018039590 A1	01-03-2018	CA 3035084 A1	01-03-2018
		EP 3504334 A1	03-07-2019
		WO 2018039590 A1	01-03-2018