UNIVERSITY OF COPENHAGEN



Københavns Universitet

Lessons learned from metabolic engineering in hairy roots

Transcriptome and metabolic profile changes caused by Rhizobium-mediated plant transformation in Cucurbitaceae species

Almeida, Aldo; Favero, Bruno Trevenzoli; Dong, Lemeng; Cárdenas, Pablo D.; Saenz-Mata, Jorge; Lütken, Henrik; Bak, Søren

Published in:
Plant Physiology and Biochemistry

DOI:
10.1016/j.plaphy.2023.107797

Publication date:
2023

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Almeida, A., Favero, B. T., Dong, L., Cárdenas, P. D., Saenz-Mata, J., Lütken, H., & Bak, S. (2023). Lessons learned from metabolic engineering in hairy roots: Transcriptome and metabolic profile changes caused by Rhizobium-mediated plant transformation in Cucurbitaceae species. *Plant Physiology and Biochemistry*, *201*, Article 107797. https://doi.org/10.1016/j.plaphy.2023.107797

Download date: 02. Oct. 2026

Plant Physiology and Biochemistry 201 (2023) 107797

Contents lists available at ScienceDirect

Plant Physiology and Biochemistry

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

Lessons learned from metabolic engineering in hairy roots: Transcriptome and metabolic profile changes caused by *Rhizobium*-mediated plant transformation in Cucurbitaceae species

Aldo Almeida [a,*,1], Bruno Trevenzoli Favero [a], Lemeng Dong [b], Pablo D. Cárdenas [a], Jorge Saenz-Mata [c], Henrik Lütken [a], Søren Bak [a,**]

[a] Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871, Frederiksberg, Denmark
[b] Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, Netherlands
[c] Facultad de Ciencias Biológicas, Universidad Juárez del Estado de Durango, Av. Universidad s/n, 35010, Gómez Palacio, Durango, Mexico

A B S T R A C T

Cucurbitaceae species are used in traditional medicine around the world. Cucurbitacins are highly oxygenated triterpenoids found in Cucurbitaceae species and exhibit potent anticancer activity alone and in combination with other existing chemotherapeutic drugs. Therefore, increasing production of these specialized metabolites is of great relevance. We recently showed that hairy roots of *Cucurbita pepo* can be used as a platform for metabolic engineering of cucurbitacins to modify their structure and increase their production. To study the changes in cucurbitacin accumulation upon formation of hairy roots, an empty vector (EV) control and *Cucurbitacin inducing bHLH transcription factor 1* (*CpCUCbH1*)-overexpressing hairy roots of *C. pepo* were compared to untransformed (WT) roots. Whilst *CpCUCbH1*-overexpression increased production of cucurbitacins I and B by 5-fold, and cucurbitacin E by 3-fold when compared to EV lines, this increase was not significantly different when compared to WT roots. This indicated that *Rhizobium rhizogenes* transformation lowered the cucurbitacins levels in hairy roots, but that increasing expression of cucurbitacin biosynthetic genes by *CpCUCbH1*-overexpression restored cucurbitacin production to WT levels. Subsequent metabolomic and RNA-seq analysis indicated that the metabolic profile and transcriptome of hairy roots was significantly changed when compared to WT roots. Interestingly, it was observed that 11% of the differentially expressed genes were transcription factors. It was noteworthy that the majority of transcripts showing highest Pearson correlation coefficients to the *Rhizobium rhizogenes* genes *rolB*, *rolC* and *ORF13a*, were predicted to be transcription factors. In summary, hairy roots are an excellent platform for metabolic engineering of plant specialized metabolites, but these extensive transcriptome and metabolic profile changes should be considered in subsequent studies.

1. Introduction

The current work arises from our previous article on metabolic engineering of cucurbitacins in *Cucurbita pepo* hairy roots in which we demonstrated that overexpression of the transcription factor *Cucurbitacin inducing bHLH transcription factor 1* (*CpCUCbH1*) can increase cucurbitacin levels in several Cucurbitaceae species (Almeida et al., 2022). Here we investigated in detail the effects of *Rhizobium*-mediated transformation on the metabolic profile and transcriptome that we observed when comparing stable and transiently transformed plant tissues to non-transformed tissues.

Bacterial species within the Rhizobiaceae are capable of establishing symbiotic root nodule nitrogen fixation or pathogenicity with host plants (Wang et al., 2020; Kim et al., 2002). These host-associated lifestyles are controlled by mobile DNA elements (plasmids), that encode nodulation or virulence factors (Unay and Perret, 2020; Lapham et al., 2021; Velázquez et al., 2005).

The species *Rhizobium rhizogenes* and *Rhizobium tumefaciens (Agrobacterium tumefaciens)* (Velázquez et al., 2005; Young and De Boer, 2001; Young, 2001; Young et al., 2003, 2008) have the capacity to transfer stretches of transfer-DNA (T-DNA), flanked by right-border and left-border sequences, into infected plant cells. Once inside the nucleus the T-DNA may be transiently expressed and/or integrate into the plant host genome (Nishizawa-Yokoi et al., 2021). However, the physiological bases of the diseases caused by infection of either bacteria differ significantly; whereas *R. rhizogenes* causes the hairy root disease, *R. tumefaciens* causes crown gall disease. These bacterial species are of particular interest in biotechnology as scientists can use their virulence process to stably transform plants by replacing native T-DNA genes of the pathogenic plasmid with other genes of interest or using a binary

* Corresponding author.
** Corresponding author.
E-mail addresses: robles@plen.ku.dk (A. Almeida), bak@plen.ku.dk (S. Bak).
[1] Current address: Carlsberg Research Laboratory, 1799 Copenhagen, Denmark.

https://doi.org/10.1016/j.plaphy.2023.107797
Received 17 March 2023; Received in revised form 11 May 2023; Accepted 23 May 2023
Available online 24 May 2023
0981-9428/© 2023 The Authors. Published by Elsevier Masson SAS. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

vector system; indeed *Rhizobium*-mediated transformation can be regarded as the preferred and most common method for plant genetic transformation in the laboratory.

The hairy root disease arises from the infection of a wounded plant cell by *Rhizobium rhizogenes*. This bacterium inserts a set of root oncogenic loci (rol) genes and open reading frames (ORFs) contained in the hairy root inducing (Ri) plasmid, that participate in the proliferation of neoplastic roots and root hairs (White and Nester, 1980; Veena and Taylor, 2007). The best characterized genes on the T-DNA, causing this phenotype, are *rol*A, *rol*B, *rol*C and *rol*D, which correspond to ORFs 10, 11, 12 and 15, respectively (White et al., 1985; Slightom et al., 1986). Additionally, *ORFs13* and *14* were later acknowledged as also playing a significant role in root induction in carrot disks and tobacco leaves. The enzymatic activity of proteins encoded by some of these *R. rhizogenes*-specific genes have been categorized (Estruch et al., 1991a, 1991b), but it is important to note that their functions/mechanisms *in planta* remain uncharacterized (Faiss et al., 1996).

Transformation with *R. rhizogenes* gives several advantages over *R. tumefaciens* transformation which researchers have been trying to exploit in order to obtain stable transgenic plants, root biology studies, gene functional analysis and specialized metabolite production (Veena and Taylor, 2007). *R. rhizogenes* induced hairy-root cultures show rapid growth *in vitro* and do not require supplementation of phytohormones and are relatively stable, this makes their containment and scale-up feasible. There are also indications that it is easier to regenerate transgenic plants from hairy root tumors than from *R. tumefaciens* induced

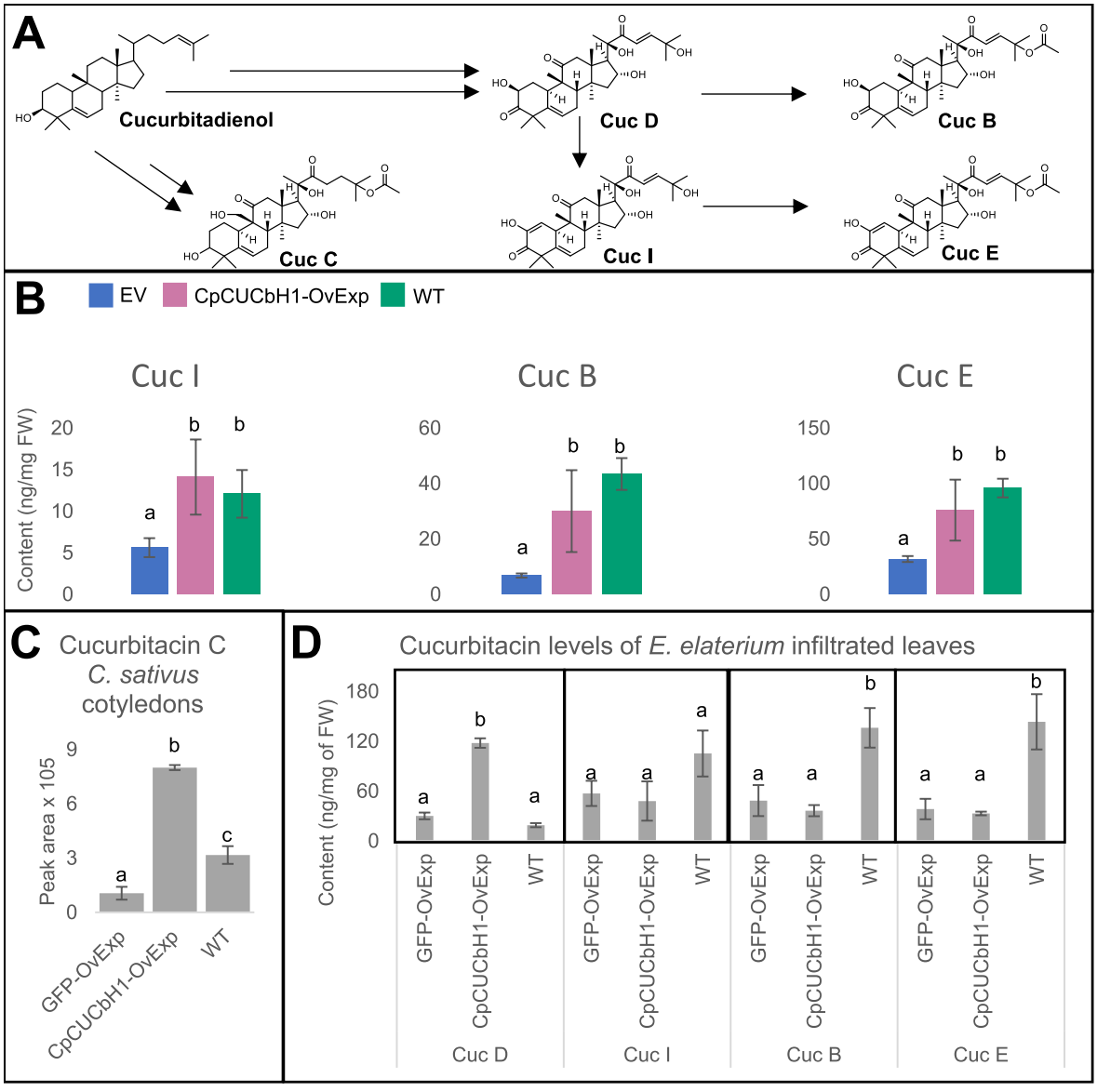


Fig. 1. Cucurbitacin production of *Cucurbita pepo* hairy roots, *Cucumis sativus* cotyledons and *Ecballium elaterium* leaves is affected by *Rhizobium* infection. (A) Schematic of the biosynthesis pathway for the major cucurbitacins (B, C, D, E and I), which were the focus of this study. (B) Comparison of Cucurbitacin B, E and I levels between *Cucurbita pepo* hairy roots overexpressing *CpCUCbH1* ($n = 3$), EV ($n = 5$) and untransformed *C. pepo* roots ($n = 6$). (C) Relative abundance of Cucurbitacin C in infiltrated *Cucumis sativus* cotyledons and wild-type cotyledons ($n = 3$) (Transient gene expression). (D) Concentration of Cucurbitacin B, D, E and I in *Ecballium elaterium* infiltrated leaves and uninfiltrated leaves ($n = 3$) (Transient gene expression)

Abbreviations: EV- empty vector control, CpCUCbH1- samples overexpressing transcription factor CpCUCbH1, GFP- Green Fluorescent Protein control, WT- Wild type (untransformed tissue).

Error bars are standard deviation. Data for EV and CpCUCbH1-overexpressing treatments is the same as reported in Almeida et al. (2022). Statistical significance was tested using one-way ANOVA ($p < 0.05$).

crown gall tumors thanks to the predisposition of rooting from the former tumors (Gelvin, 1990). In addition, hairy roots are considered to have equal or greater biosynthetic capacity for phytochemicals compared to the wild type plant (Kim et al., 2002).

For the above-mentioned reasons our lab has previously used hairy roots of *Cucurbita pepo* for metabolic engineering of cucurbitacins (Almeida et al., 2022; Dong et al., 2018). Cucurbitacins are triterpenoids pharmaceutically important for their anti-cancer bioactivities (Rios.J et al., 2012; Tan et al., 2008; Abdelwahab et al., 2011). Cucurbitacins arise from multiple oxidation reactions on the cucurbitadienol backbone. Cucurbitacin B (CucB), D (CucD), E (CucE) & I (CucI) are the most reported among species of the Cucurbitaceae (Fig. 1A) (David and Vallance, 1955; Rehm et al., 1957). Up to now, several steps in the biosynthesis of these four cucurbitacins remain uncharacterized. Nevertheless, in *Cucumis sativus*, *Cucumis melo* and *Citrullus lanatus* five genes encoding enzymes in the Cuc E pathway have been identified and these genes reside in a metabolic gene cluster (Shang et al., 2014). In addition, basic helix-loop-helix (bHLH) transcription factors (TFs) that can induce the expression of the previously reported Cucurbitacin-biosynthetic genes have been described (Shang et al., 2014; Zhou et al., 2016); and our recent work has shown that induction of cucurbitacin biosynthesis by orthologs of this transcription factor is conserved in *Cucurbita pepo* and other Cucurbitaceae species (Almeida et al., 2022).

In this paper we show an in-depth study of *C. pepo* hairy roots induced by *R. rhizogenes,* and how the infection changes the root transcriptome and that 11% of the differentially expressed genes are TFs. The findings described in this paper are important, as comparison of *R. rhizogenes*-transformed and non-transformed tissues at the transcriptomic and metabolomic level are not common in metabolic engineering studies of hairy roots (Pei et al., 2018; Cao et al., 2018), and we address how the unintended effects of transformation detailed here may significantly impact the outcome of such studies.

2. Materials and methods

2.1. Cloning of candidate genes

First the Spectrum Plant Total RNA Kit (Sigma-Aldrich) was used to isolate total RNA from *Cucurbita pepo* var. golden glory ten-day-old roots. Afterwards, the GoScriptTM Reverse transcriptase kit (Promega) was used to synthesize first-strand cDNA from total RNA of roots following the manufacture's instructions.

Primers used in cloning reactions are given in Table S4. Amplification was done as described in Almeida et al. (2022). Gel purified PCR products were cloned into either: i) the pJCV51 binary vector (VIB-U-Gent Center for Plant Systems Biology, Vector ID: 4_60) through gateway cloning using pDONR 207 (Invitrogen) ii) pJET 1.2/blunt vector from CloneJet PCR cloning kit (Thermo Fisher Scientific). Ligated vectors were transformed into *E. coli* and the retrieved constructs were sequenced by Sanger sequencing service (Macrogen, Europe).

2.2. Hairy root transformation

Hairy roots were produced as described in Almeida et al. (2022) by inoculating cotyledons of *Cucurbita pepo* var. golden glory grown in tissue culture with *Rizhobium rhizogenes* strain LBA9402 transformed with adequate constructs.

2.3. Transient expression experiments

Leaves of *Nicotiana benthamiana, Ecballium elaterium* and cotyledons of *Cucumis sativus* were transformed using *Rhizobium tumefaciens* (*Agrobacterium tumefaciens*) strain AGL1 containing constructs of pEAQ-HT-DEST expression vector (Sainsbury et al., 2009) harboring *eGFP* and *CpCUCbH1* as described in Almeida et al. (2022). More details of statistical analysis are given in the legend of Figs. 1 and 2.

2.4. Cucurbitacins analysis by LC-QToF-MS

Cucurbitacins B, D, E and I were determined as described in Almeida et al. (2022) with the LC-QToF method described in Dong et al. (2018). More details of statistical analysis are given in the legend of Fig. 1.

2.5. Gene expression analysis by qRT-PCR

Primers used in RT-qPCR reactions are given in Table S4, and reactions were carried out as described in Almeida et al. (2022). More details of statistical analysis are given in the legend of Figs. 2 and 5.

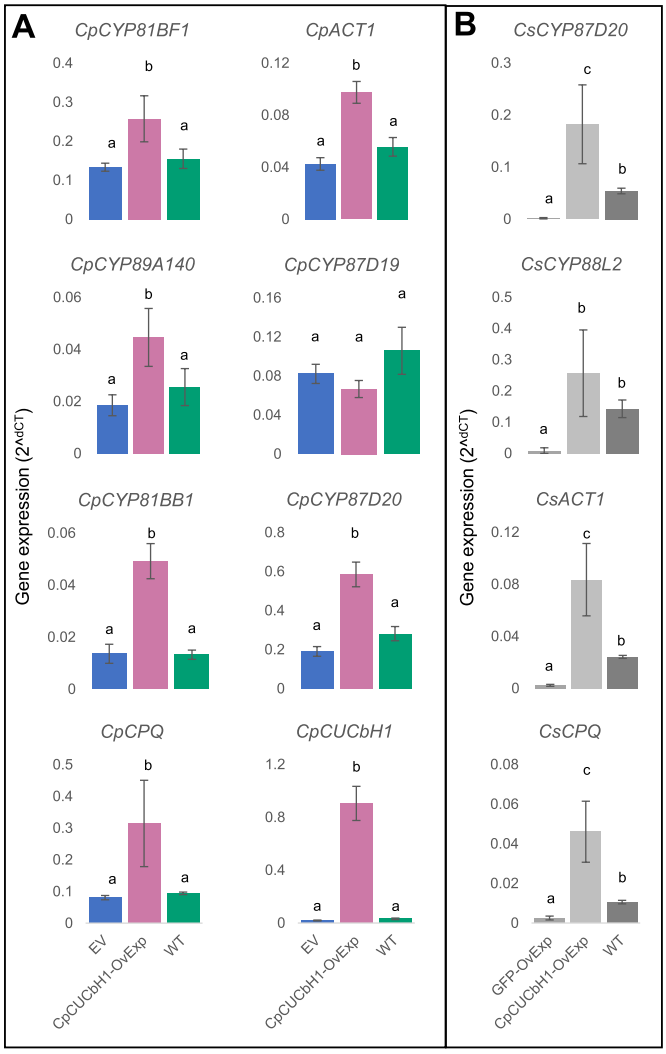


Fig. 2. Gene expression levels of cucurbitacin-biosynthetic genes in experiments of *C. pepo* hairy roots and *C. sativus* cotyledons. (A) Relative expression of selected cucurbitacin biosynthetic genes between *C. pepo* hairy root lines transformed with EV, overexpressing *CpCUCbH1* and untransformed roots (WT) ($n = 3$): CpCYP81BF1 (CsCYP81Q58 homolog), CpCYP89A140 (CsCYP89A140 homolog), CpCYP81BB1 (CsCYP81Q59 homolog), CpCPQ (CsCPQ homolog), CpACT1 (CsACT1 homolog), CpCYP87D19 (CsCYP87D19 homolog), CpCYP87D20 (CsCYP87D20 homolog), CpCUCbH1 (CsBl homolog). (B) Relative expression of cucurbitacin-biosynthetic genes from *Cucumis sativus* in cotyledons transiently expressing different constructs and untransformed (WT) cotyledons ($n = 3$): CsCPQ, CsACT1, CsCYP87D20, CsCYP88L2.
Data for EV and CpCUCbH1-overexpressing treatments is the same as reported in Almeida et al. (2022). Data points are given in Table S5. Statistical significance was tested using one-way ANOVA ($p < 0.05$).

2.6. Metabolic profile analysis

Total ion chromatograms from raw data were calibrated using an internal sodium formate standard and exported as.mzML file using DataAnalysis 4.1 (Bruker). MZmine2 (v.2.5.3) (Pluskal et al., 2010) was used for further processing of the raw chromatogram data. At first, a signal intensity noise cutoff of 1000 ion count was applied and only scans between 5 and 37 min retention time were considered. Chromatograms were built using the 'ADAP chromatogram builder' with m/z tolerance of 0.003 Da (5 ppm). The generated extracted ion chromatograms were deconvoluted into individual putative metabolites using the 'local minimum search' module. After isotopes were removed, the feature list was aligned using the 'joint aligner' tool using m/z tolerance of 0.006 Da (10 ppm) and retention time tolerance of 0.15 min with subsequent gap-filling using intensity, m/z, and retention time tolerance of 1%, 0.003 Da (5 ppm) and 0.15 min, respectively. The obtained feature list was used to conduct a principal component analysis within MzMine2.

2.7. Transcriptome assembly and analysis

Total RNA was isolated from *Cucurbita pepo* var. golden glory EV and CpCUCbH1-overexpressing HR lines as well as wild-type roots (ten days old) using Spectrum Plant Total RNA Kit (Sigma-Aldrich) and treated with RQ1 RNase-free DNase I (Promega). Quality and quantity of RNA were assayed using Bioanalyzer 2100 (Agilent).

Sequencing was performed at the facilities of BGI Hong Kong using the BGISEQ-500. First, 1 μg Total RNA was randomly fragmented on a Covaris instrument (Brighton, UK). Afterwards, first-strand cDNA was generated using random hexamer primed reverse transcription, followed by a second strand cDNA synthesis. Synthesized cDNA fragments were end repaired and 3′ adenylated. Adaptors were ligated to the 3' cDNA adenylated fragment ends. The DNA fragments with adaptors, were amplified and purified using the AMPure XP medium kit (Agencourt). Libraries were validated on Bioanalyzer 2100. Finally, the double-stranded PCR products were heat denatured and circularized by the splint oligonucleotide sequence yielding single strand circle DNAs (ssCir DNAs) which were formatted as the final library. ssCir DNA libraries were amplified with phi29 to make DNA nanoball (DNB) each containing more than 300 copies of one molecule. The DNBs were then loaded into the patterned nanoarray and paired-end 100-bases reads were generated through combinatorial Probe Anchor Synthesis (cPAS). Adapter sequences and low-quality bases (<Q20) were trimmed from the raw reads using SOAPnuke v1.5.5 software at BGI Hong Kong.

The filtered reads were *de novo* assembled using Trinity (Grabherr et al., 2011). Tgicl (Pertea et al., 2003) was used on cluster transcripts to remove abundance and get Unigenes. Gene ontology of Unigenes was annotated using Blast2GO (Conesa et al., 2005) and distributed in 3 categories: biological process, cellular component and molecular functions. The software TransDecoder was used to identity the candidate coding region among the Unigenes. The longest Open reading frame was selected and blast with SwissProt and Hmmscan database to search for the Pfam protein homology sequence for the prediction of CDS. The list of Unigenes that encode TF was predicted using InterProscan software to annotate the Unigenes with InterPro database, categorizing Unigenes according to the protein domain. The software package RSEM was used for accurate transcript quantification, this package uses Bowtie2 to align reads against reference transcript sequences (Langmead and Salzberg, 2012). The DESeq2 package (Love et al., 2014) was used to statistically analyse differential gene expression between samples. To detect genes which expression was correlating to that of *rolB, rolC* and ORF13a, pair-wise Pearson correlation coefficients were calculated (Usadel et al., 2009).

3. Results

3.1. Infection by Rhizobium species decreases cucurbitacin accumulation in Cucurbitaceae species

The current study was initiated to measure the differences in cucurbitacin accumulation and transcriptomic changes in Cucurbitaceae tissues upon overexpressing the transcription factor *CpCUCbH1* that was used in our previous work to increase the cucurbitacin accumulation (Almeida et al., 2022) and compare them to untransformed tissues. CpCUCbH1 is encoded by the Cp4.1LG05g03810.1 locus in *C. pepo* and is the functional ortholog of the CsBl transcription factor reported to increase cucurbitacin C levels in *Cucumis sativus* (Shang et al., 2014). An LC-MS analysis performed in our previous work (Almeida et al., 2022) showed that overexpression of *CpCUCbH*1 in *C. pepo* hairy roots increased the production of Cuc I&B approximately 5-fold and Cuc E by 3-fold (Fig. 1B) compared to an empty vector (EV) control. However, in this current paper when these previously obtained values of EV and CpCUCbH1-overexpression lines were compared to those of non-transformed 10-day-old roots (WT), we observed that infection with *R. rhizogenes* (EV lines) decreased the levels of Cucs B,E&I (3-fold, 3-fold and 2-fold, respectively) (Fig. 1B). In fact, the significant boost obtained from overexpressing *CpCUCbH1* only partially restored the cucurbitacin levels back to those observed in WT roots (Fig. 1B).

In our previous work we had shown that the induction of cucurbitacin biosynthesis by *CpCUCbH1* overexpression was partially conserved in species of the Cucurbitaceae by performing transient transformation experiments on *Cucumis sativus* cotyledons (common name: cucumber) and *Ecballium elaterium* (common name: Squirting cucumber) leaves using *R. tumefaciens* (Almeida et al., 2022). In this current paper the previous data was compared to that of non-infiltrated tissues. In these experiments, overexpression of Green Fluorescent Protein (GFP) was used to both monitor successful transient expression in infiltrated tissues, and as a control to compare cucurbitacin induction by overexpression of *CpCUCbH1*. In the case of *C. sativus* cotyledons, LC-MS analysis showed that *R. tumefaciens*-transformation for *GFP* overexpression lowered levels of cucurbitacin C by 3-fold in comparison to the non-transformed (WT) cotyledons (Fig. 1C); nevertheless, overexpression of *CUCbH1* raised cucurbitacin C levels more than 2-fold of those observed for WT cotyledons. As for the experiments with *E. elaterium*, overexpression of *CUCbH1* increased levels of cucurbitacin D almost 4-fold of that observed in non-transformed (WT) leaves and leaves overexpressing *GFP* (Fig. 1D), however levels of Cucs B&E were significantly higher in WT leaves compared to infiltrated leaves; Cuc I content did not show statistically significant difference between the treatments (Fig. 1D).

In summary, the LC-MS analysis shows that *Rhizobium*-mediated transformation lowers cucurbitacin levels in the tissues of the three Cucurbitaceae species tested when compared to WT tissues.

3.2. Differences in cucurbitacin accumulation between EV and WT roots cannot be explained by expression levels of cucurbitacin biosynthetic genes alone

To explain the lowered cucurbitacin accumulation between the transformed and WT tissues, we quantified the expression levels of known genes in the cucurbitacin-biosynthetic pathway (downstream of 2,3-oxidosqualene formation) using RT-qPCR. In the case of *C. pepo* roots, the expression levels of the oxidosqualene cyclase (*CpCPQ*) and the characterized Cytochromes P450 (P450s) as well as the acyl transferase (*CpACT1*) were all significantly increased in the transgenic hairy roots overexpressing the transcription factor *CpCUCbH1* (Fig. 2A); only expression levels of *CpCYP87D19* appeared the same in all roots. Interestingly and in contrast to the decreased levels of metabolites measured (Fig. 1B), there was no significant difference in expression levels of all genes analyzed between EV hairy roots and WT roots

(Fig. 2A).

In the case of the transient-expression experiments of *C. sativus* cotyledons, the overexpression of *CpCUCbH1* increased the expression of characterized genes involved in cucurbitacin C biosynthesis (i.e. *CsCYP87D20*, *CsACT1* and *CsCPQ*) significantly higher than the *GFP*-Ovexpression and WT cotyledons (Fig. 2B); with the exception of *CYP88L2* gene expression which was not significantly different in *CpCUCbH1*-overexpressing lines and WT cotyledons (Fig. 2B). It is important to note that gene expression for these four genes was lowest in the *C. sativus* cotyledons overexpressing *GFP*, which served as the control in these experiments.

Since no genetic resources exist for *E. elaterium*, no gene expression analysis could be conducted.

In summary, for the case of *C. pepo* the gene expression levels of cucurbitacin biosynthetic genes did not effectively explain the difference in cucurbitacin accumulation between the hairy roots and WT roots, as there was no significant difference between them.

3.3. Infection by Rhizobium rhizogenes decreases Cucurbitacin levels and changes the root metabolic profile and transcriptome

To explore more in-depth the effects of *R. rhizogenes* infection we performed LC-MS and RNA-seq analysis on *C. pepo* hairy roots and 10-day-old WT-roots. LC-MS analysis detected 551 putative metabolites from minute 5 to 37 where 390 were present in *C. pepo* hairy roots overexpressing *CpCUCbH1*, EV control and the WT roots (Fig. 3, Table S1). Visualization of the data covering LC-MS analysis in the space of the first two components of a principal component analysis (PCA) of peak area from these putative metabolites depicts the overall metabolite profile changes in *C. pepo* hairy roots overexpressing *CpCUCbH1* and EV control when compared to WT roots (Fig. 4A).

Similarly, visualization of the data covering differentially expressed genes in the space of the first two principal components after performing PCA on gene expression levels shows that EV and CpCUCbH1 HR lines cluster together and are separate from the WT roots (Fig. 4B). A total of 5688 and 4217 genes were significantly upregulated and down regulated, respectively, in WT roots compared to EV lines (Fig. S1A; Table S2). Gene ontology (GO) enrichment analysis of differentially expressed genes revealed that most of the DEGs fell into the Biological Processes category, followed by Cellular Component and Molecular Processes (Fig. S1B). Most interestingly was that out of the 9905 differentially expressed genes, 1099 of them (11%) were TFs. The distribution of TF expression levels can be observed in Fig. S1C which depicts a stark difference between WT roots and EV and CpCUCbH1-OvExp HR supporting the observations from the PCA of the metabolomic and transcriptomic data.

To detect genes that were correlating to the expression of *rolB, rolC* and *ORF13a*, pair-wise Pearson correlation coefficients (PCC) were calculated and a cut-off threshold of 0.98 was used. This returned a total of eight genes ($r > 0.98$) whose expression correlated with *rolB*, four with *rolC* and seven uniquely correlating to expression of *ORF13a* (Table S3). Since *R. rhizogenes* ORFs/rol genes are not present in WT roots, all the genes mentioned above were significantly upregulated in HR lines compared to WT roots (Fig. S2). We focused further work on those genes for which a functionally characterized potential homolog in *Arabidopsis thaliana* could be identified using blastx, and also a couple of genes that were starkly upregulated in WT roots compared to HR lines (Table S3). The observed differential gene expression profiles in the RNA-seq were confirmed with RT-qPCR (Fig. 5A–D). The genes with the highest fold changes between WT roots and HR lines were the ones that co-expressed with *ORF13a*. These included glutamate and glutamine synthases as well as two transcriptional activators homologous to AtWRKY22 and NAC-domain containing protein along with fructokinase 4, which was also in the genes with highest PCC to *rolC*. According to the annotation, the genes whose expression correlated with *rolB* and *rolC* appear to also have regulatory roles.

Taken together, these results suggested that *R. rhizogenes* infection changes the transcriptome and affects the metabolome in hairy roots making it significantly different from that of WT roots; these significant changes could explain the difference in accumulation of cucurbitacins that were observed in our studies.

4. Discussion

4.1. Rhizobium-mediated transformation affects the cucurbitacin content of Cucurbitaceae species

In previous studies delivering constructs with individual or multiple *ORFs/rol* genes in a range of plant species using *R. tumefaciens*, it was found that transformation of plants with *rolB* could significantly increase specialized metabolites in *Vitis amurensis*, *Rubia cordifolia* and *Artemisia cordifolia* (Shkryl et al., 2008; Dilshad et al., 2015). This has led to a premise in the scientific community that hairy root induction can result in higher yields of plant specialized metabolites. Indeed, we have previously shown that marked increases in cucurbitacin content can be observed when *C. pepo* hairy roots transformed with an EV are compared

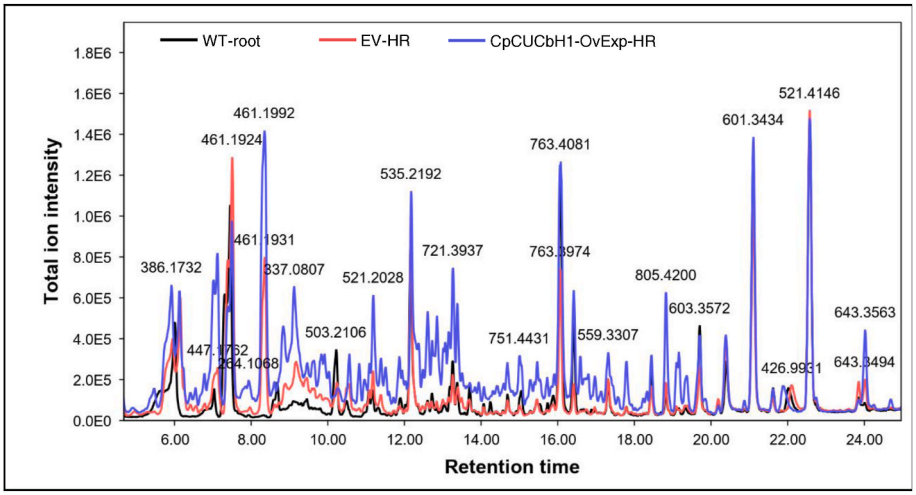


Fig. 3. *Rhizobium rhizogenes*-mediated transformation changes the root metabolome upon causing hairy root disease.
Representative UPLC-MS total ion chromatogram displaying metabolite profile differences between hairy root lines and wild type roots. Numbers on the peak represent the m/z ratio of most abundant ion in the particular peak.

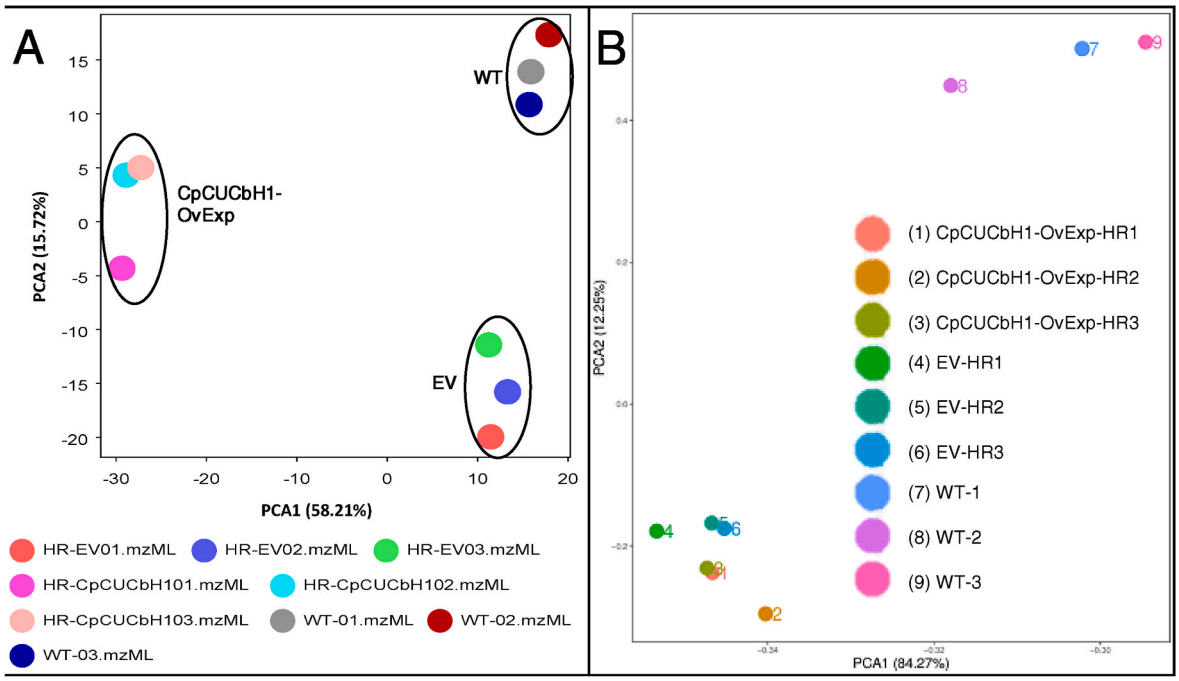


Fig. 4. *Rhizobium rhizogenes*-mediated transformation changes the plant transcriptome upon causing hairy root disease. (A) Distribution of hairy root and wild type roots in the space of the first two principal components after performing Principal component analysis on LC-MS analysis. (B) Distribution of hairy root and wild type roots in the space of the first two principal components after performing Principal component analysis on gene expression data.

to transgenic hairy roots transformed to overexpress specific genes influencing the specialized metabolite pathway *i. e.* Squalene Epoxidase 2 from *C. pepo* (CpSQE2) (Dong et al., 2018) or TFs like CpCUCbH1 (Almeida et al., 2022). However, from the few studies comparing production of specialized metabolites in hairy roots to untransformed WT roots, the premise that hairy root disease increases specialized metabolite content is inconclusive. Although, some studies have reported higher concentrations of specialized metabolites in hairy roots than in WT roots (Moyano et al., 2002; Kim et al., 2014), we had previously observed that accumulation of α-onocerin in non-transgenic hairy roots of *Ononis spinosa* was 15-fold lower than that of 45-day-old WT roots (Almeida et al., 2018). Similarly, in this study we observed reduction of cucurbitacin content upon induction of hairy roots in *C. pepo* when compared to WT roots. These observations are supported by other studies that have reported a decrease or a similar level of specialized metabolites in hairy roots compared to WT roots (Park and Facchini, 2000; Moyano et al., 2003). Our studies show that α-onocerin content in *Ononis spinosa* and cucurbitacin content in the Cucurbitaceae species can be added to the list of specialized metabolites which production is apparently decreased upon infection with *R. rhizogenes* and *R. tumefaciens*.

In the current manuscript we observed that infection with *R. rhizogenes* changes significantly the metabolic profile of *C. pepo* roots. Recent publications have indicated that a certain level of stress modulation is related to *rol*-genes, especially in reducing oxidative stress by the production of peroxidases, modulating NADPH-oxidase expression, suppression of reactive oxygen species or altering the expression of heat-shock proteins and cyclophilins (Bulgakov et al., 2013). The lack of MS/MS data in the metabolic profile data acquired in this study does not allow to probe if the significant change in metabolism of hairy roots and WT roots is caused by similar responses. The shift in metabolic profiles will need to be further explored in subsequent studies.

4.2. *Rhizobium rhizogenes*-mediated transformation significantly altered the root transcriptome

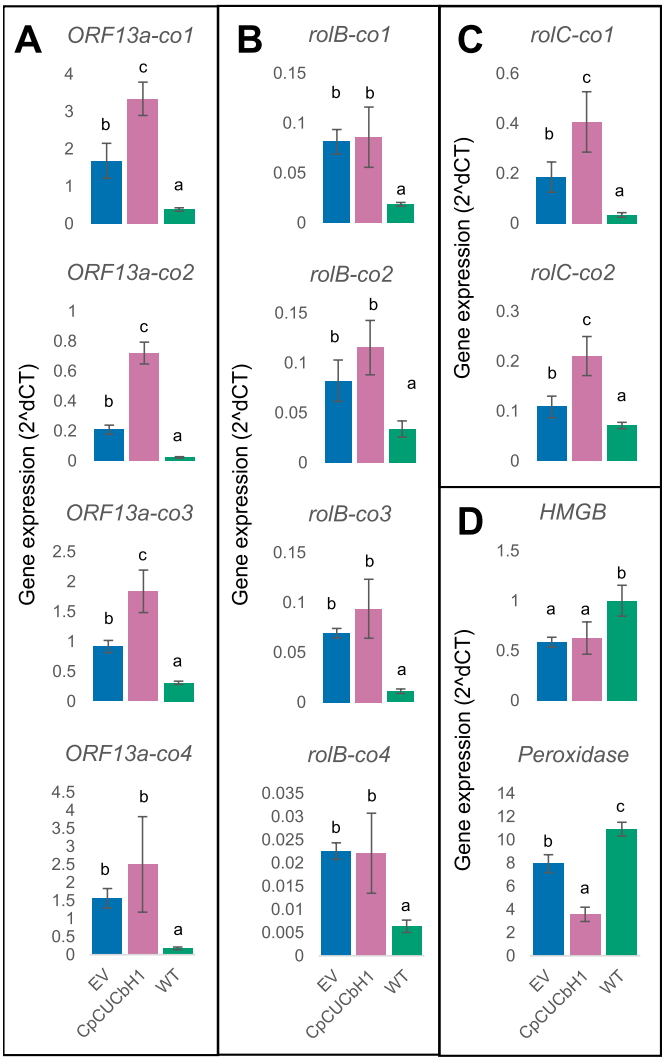
To the extent of our knowledge this is the first combined metabolomic and transcriptomic study that explores differential gene expression between hairy roots and WT roots and gives additional insights to begin establishing the gene regulatory network of hairy root disease.

The *R. rhizogenes rol*-genes A, B, C and D were initially the first loci determined as common denominators for hairy root formation (White et al., 1985). Alone or in combination they are certain to cause root induction post genome integration. Additionally, *R. rhizogenes ORFs13* and *14* were later acknowledged as also playing a significant role in root induction in carrot disks and tobacco leaves. Earlier studies attempting to elucidate the function of *rol* genes/*ORFs* in *R. rhizogenes*, have delivered constructs with individual or multiple *ORFs/rol* genes in a range of plant species using *R. tumefaciens*. However, transcriptomic level comparisons between hairy roots and untransformed tissues are lacking in the literature.

In this manuscript, a comparison of EV and *CpCUCbH1*-overexpressing hairy roots to WT roots through RNA-seq analysis suggested that *R. rhizogenes* infection changes the transcriptome of roots and that 11% of differentially expressed genes are TFs (Fig. S1C) (when all *rol genes/ORFs* are transformed). Our data suggests that *rolB* and *ORF13a* could have a great role in this transcriptional changes, as several *C. pepo* genes with highest PCC to *rolB* and *ORF13a* show high sequence identity to *A. thaliana* regulatory genes (Table S3). The genes with highest PCC for *ORF13a* showed the highest fold changes in EV and transgenic hairy roots compared to WT roots (Fig. 5A). *C. pepo ORF13a* encodes a protein with repeats of the DNA-binding SPXX motif, the presence or absence of these motifs can significantly affect the phenotype of plants transformed with this gene (Lemcke and Schmülling, 1998; Favero et al., 2021). Our observations in this study coupled to the presence of the SPXX motif repeats, suggests that ORF13a may indeed function as a transcription factor, thus also shedding light on the less characterized ORFs of *R. rhizogenes*. On the other hand, the protein encoded by *rolC* has been shown to have β-glucosidase activity (Estruch et al., 1991a) and probably influences transcription reprogramming indirectly. In terms of *rolC*, its individual delivery has led to a strong stimulatory effect on anthraquinone biosynthesis (Bulgakov et al., 2004) and this increase is not related to plant defense mechanisms, nor Ca^{2+} dependent NADPH

oxidase pathway (Bulgakov et al., 2003), nor in the octadecanoid or salicylic acid-mediated response (Bulgakov et al., 2004, 2013).

As for *rolB*, the genes with highest PCC mostly encoded TFs (Table S3). Previously, a tyrosine-phosphatase activity had been described for RolB and therefore it is assumed that RolB plays a major role in the signal transduction of auxin (Filippini et al., 1996). Previous work had shown that overexpression of *rolB* would increase expression of *MYB12* and heat-shock proteins (Hsp) of the DnaK subfamily while reducing that of *MYB11* and *MYB111* (Bulgakov et al., 2016, 2018). Our study confirms the upregulation of Hsp70 (Table S3), however we did not focus on it since the PCCs were below the established threshold. It is important to notice that the cited studies delivered *rolB* alone while our study delivered the complete T-DNA of *R. rhizogenes*. Similarly, earlier studies noticed delivering *rolB* together with *rolA* and *rolC* in *Rubia cordifolia* decreased the growth inhibition and anthraquinone induction effects of delivering *rolB* alone (Shkryl et al., 2008). In addition, different signal transduction pathways may be disturbed by the presence of complete T-DNA or individual *rol*-genes, inconsistencies in the gene expression and phenotypical effects of *rol* genes/ORFs may also be due to the random integration, copy number, soma-clonal variation and mutation of constructs. The uncontrollable effects from random gene integration will remain an obstacle until T-DNA delivery can be targeted.



(caption on next column)

Fig. 5. Several of the genes whose expression pattern correlate with that of rol genes/ORFs after *Rhizobium rhizogenes*-mediated transformation encode transcription factors.

Each graph shows the relative expression of a gene from *C. pepo* with high Pearson Correlation Coefficient to a specific *R. rhizogenes rol gene/ORF*. For practical reasons the graphs were titled with a code starting with the name of the *rol gene/ORF* concerned followed by a co (for correlation) and a number.

(A) Relative expression for genes whose expression correlated to that of *R. rhizogeous ORF13a* between *C. pepo* hairy roots and wild-type roots. ORF13a-co1, has a high PCC to *ORF13*, and based on a blast analysis is a putative homolog of *A. thaliana* AtAF1, which belongs to a large family of putative transcriptional activators with NAC domain. Similarly, ORF13a-co2 is homologous to *A. thaliana* WRKY22 a member of WRKY Transcription Factor Group II-e. It is involved in regulation of dark induced leaf senescence. ORF13a-co3 is homologous to glutamine synthase from *A. thaliana*. ORF13a-co4 is homologous to glutamate synthase from *A. thaliana*.

(B) Relative expression for genes whose expression correlated to that of *RolB* between *C. pepo* hairy roots and wild-type roots. RolB-co1 is homologous to SR45a from *A. thaliana* encoding a serine/arginine rich-like protein involved in the regulation of stress-responsive alternative splicing. RolB-co2 is homologous to a MADS box transcription factor from *A. thaliana*. RolB-co3 is homologous to *A. thaliana* ARF6 encoding a member of the auxin response factor family. RolB-co4 is homologous to a subunit of CCAAT-binding complex from *A. thaliana*. This binds to CCAAT box motif present in some plant promoter sequences.

© Relative expression for genes whose expression correlated to that of *RolC* between *C. pepo* hairy roots and wild-type roots. RolC-co1 is homologous to a member of the SHI family in *A. thaliana*. RolC-co2 is homologous to a fructokinase in *A. thaliana*.

(D) Relative expression for gene upregulated in *C. pepo* wild-type roots compared to hairy roots. HMGB gene is homologous to *A. thaliana* HMGB encoding a protein belonging to the subgroup of HMGB (high mobility group B) proteins that have a distinctive DNA-binding motif, the HMG-box domain. Peroxidase gene is homologous to a peroxidase superfamily protein in *A. thaliana*.

Statistical significance was tested using one-way ANOVA ($n = 3$), ($p < 0.05$). Data points are given in Table S5.

5. Conclusions

In the present study we show that Rhizobium-mediated transformation significantly change the metabolic profile and transcriptome in *C. pepo* hairy roots. RNA-seq analysis showed that 11% of differentially expressed genes were annotated as TFs. Pearson correlation coefficients calculated for few rol genes/ORFs revealed significant correlations between several TFs and *rol*B and ORF13a. The analysis in this manuscript reveal for the first time the underlying transcriptomic changes of hairy root disease and invites future research on the less characterized ORFs from *R. rhizogenes*. In addition, our results show that metabolic engineering in *C. pepo* hairy roots can effectively modify cucurbitacin content in the hairy root system, and provide the lesson to consider the major metabolomic and transcriptomic changes that rol genes/ORFs may cause for future metabolic engineering studies in hairy roots. We conclude that the potential of the hairy root disease to increase the production of specialized metabolites has to be empirically assessed for each individual plant and metabolic pathway.

Funding

S. B. and A. A. were supported by grants from the Novo Nordisk Foundation grant No. NNF17OC0027646 and Independent Research Fund Denmark grant No. 7017-00275B.

Author contributions

A.A. conceived the original research plan, performed most of the experiments, analyzed the data and drafted the article; P. D. C. assisted in the Pearson correlation analysis.; S. B. supervised the work and revised the manuscript. B. T. F., L. D., J. S-M. and H. L. complemented

the writing.

Declaration of competing interest

The authors acknowledge that this study received funding from the Novo Nordisk Foundation. The funder was not involved in the experimental design, in the collection and interpretation of data, the writing of this article, or the decision to submit it for publication.

Data availability

Data will be made available on request.

Acknowledgements

We would like to thank Nagamani Balagurusamy for his advice in the classification of *Rhizobium/Agrobacterium* species and informing that taxonomy in these genera continues to improve and these terms can currently be used synonymously.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.plaphy.2023.107797.

References

Abdelwahab, S.I., et al., 2011. Anti-inflammatory activities of cucurbitacin E isolated from Citrullus lanatus var. citroides: role of reactive nitrogen species and cyclooxygenase enzyme inhibition. Fitoterapia 82 (8), 1190–1197.
Almeida, A., et al., 2018. A single oxidosqualene cyclase Produces the Seco-Triterpenoid α-onocerin. Plant Physiol. 176 (2), 1469.
Almeida, A., et al., 2022. Metabolic engineering of cucurbitacins in Cucurbita pepo hairy roots. Front. Plant Sci. 13.
Bulgakov, V.P., et al., 2003. Effects of Ca2+ channel blockers and protein kinase/phosphatase inhibitors on growth and anthraquinone production in Rubia cordifolia callus cultures transformed by the rolB and rolC genes. Planta 217 (3), 349–355.
Bulgakov, V.P., et al., 2004. The rolB and rolC genes activate synthesis of anthraquinones in Rubia cordifolia cells by mechanism independent of octadecanoid signaling pathway. Plant Sci. 166 (4), 1069–1075.
Bulgakov, V.P., et al., 2013. Recent advances in the understanding of agrobacterium rhizogenes-derived genes and their effects on stress resistance and plant metabolism. In: Doran, P.M. (Ed.), Biotechnology of Hairy Root Systems. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–22.
Bulgakov, V.P., et al., 2016. The rolB gene activates secondary metabolism in Arabidopsis calli via selective activation of genes encoding MYB and bHLH transcription factors. Plant Physiol. Biochem. 102, 70–79.
Bulgakov, V.P., et al., 2018. The rolB plant oncogene affects multiple signaling protein modules related to hormone signaling and plant defense. Sci. Rep. 8 (1), 2285.
Cao, W., et al., 2018. Transcription factor SmWRKY1 positively promotes the biosynthesis of tanshinones in salvia miltiorrhiza. Front. Plant Sci. 9.
Conesa, A., et al., 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21 (18), 3674–3676.
David, A., Vallance, D.K., 1955. Bitter principles of cucurbitaceæ. J. Pharm. Pharmacol. 7 (1), 295–296.
Dilshad, E., et al., 2015. Genetic transformation of Artemisia carvifolia buch with rol genes enhances artemisinin accumulation. PLoS One 10 (10), e0140266.
Dong, L., et al., 2018. Co-expression of squalene epoxidases with triterpene cyclases boosts production of triterpenoids in plants and yeast. Metab. Eng. 49, 1–12.
Estruch, J.J., et al., 1991a. The plant oncogene rolC is responsible for the release of cytokinins from glucoside conjugates. EMBO J. 10 (10), 2889–2895.
Estruch, J.J., Schell, J., Spena, A., 1991b. The protein encoded by the rolB plant oncogene hydrolyses indole glucosides. EMBO J. 10 (11), 3125–3128.
Faiss, M., et al., 1996. Chemically induced expression of the rolC-encoded β-glucosidase in transgenic tobacco plants and analysis of cytokinin metabolism: rolC does not hydrolyze endogenous cytokinin glucosides in planta. Plant J. 10 (1), 33–46.
Favero, B.T., et al., 2021. Transgenic kalanchoë blossfeldiana, containing individual rol genes and open reading frames under 35S promoter, exhibit compact habit, reduced plant growth, and altered ethylene tolerance in flowers. Front. Plant Sci. 12, 840.
Filippini, F., et al., 1996. A plant oncogene as a phosphatase. Nature 379 (6565), 499–500.
Gelvin, S.B., 1990. Crown gall disease and hairy root disease 1: a sledgehammer and a tackhammer. Plant Physiol. 92 (2), 281–285.
Grabherr, M.G., et al., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29 (7), 644–652.
Kim, Y., Wyslouzil, B.E., Weathers, P.J., 2002. Secondary metabolism of hairy root cultures in bioreactors. Vitro Cell. & Deve. Biol. - Plant 38 (1), 1–10.
Kim, Y.-K., et al., 2014. Enhanced triterpene accumulation in panax ginseng hairy roots overexpressing mevalonate-5-pyrophosphate decarboxylase and farnesyl pyrophosphate synthase. ACS Synth. Biol. 3 (10), 773–779.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9 (4), 357–359.
Lapham, R.A., et al., 2021. Agrobacterium VirE2 protein modulates plant gene expression and mediates transformation from its location outside the nucleus. Front. Plant Sci. 12.
Lemcke, K., Schmülling, T., 1998. Gain of function assays identify non- rol genes from Agrobacterium rhizogenes TL-DNA that alter plant morphogenesis or hormone sensitivity. Plant J. 15 (3), 423–433.
Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15 (12), 550.
Moyano, E., et al., 2002. Alkaloid production in Duboisia hybrid hairy root cultures overexpressing the pmt gene. Phytochemistry 59 (7), 697–702.
Moyano, E., et al., 2003. Effect of pmt gene overexpression on tropane alkaloid production in transformed root cultures of Datura metel and Hyoscyamus muticus. J. Exp. Bot. 54 (381), 203–211.
Nishizawa-Yokoi, A., et al., 2021. Agrobacterium T-DNA integration in somatic cells does not require the activity of DNA polymerase θ. New Phytol. 229 (5), 2859–2872.
Park, S.U., Facchini, P.J., 2000. Agrobacterium rhizogenes-mediated transformation of opium poppy, Papaver somniferum L., and California poppy, Eschscholzia californica Cham., root cultures. J. Exp. Bot. 51 (347), 1005–1016.
Pei, T., et al., 2018. SmJAZ8 acts as a core repressor regulating JA-induced biosynthesis of salvianolic acids and tanshinones in Salvia miltiorrhiza hairy roots. J. Exp. Bot. 69 (7), 1663–1678.
Pertea, G., et al., 2003. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics 19 (5), 651–652.
Pluskal, T., et al., 2010. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. BMC Bioinf. 11 (1), 395.
Rehm, S., et al., 1957. Bitter principles of the cucurbitaceae. VII.—the distribution of bitter principles in this plant family. J. Sci. Food Agric. 8 (12), 679–686.
RiosJ, L., et al., 2012. Cucurbitacins as inducers of cell death and a rich source of potential anticancer compounds. Curr. Pharmaceut. Des. 18 (12), 1663–1676.
Sainsbury, F., Thuenemann, E.C., Lomonossoff, G.P., 2009. pEAQ: versatile expression vectors for easy and quick transient expression of heterologous proteins in plants. Plant Biotechnol. J. 7 (7), 682–693.
Shang, Y., et al., 2014. Biosynthesis, regulation, and domestication of bitterness in cucumber. Science 346 (6213), 1084.
Shkryl, Y.N., et al., 2008. Individual and combined effects of the rolA, B, and C genes on anthraquinone production in Rubia cordifolia transformed calli. Biotechnol. Bioeng. 100 (1), 118–125.
Slightom, J.L., et al., 1986. Nucleotide sequence analysis of TL-DNA of Agrobacterium rhizogenes agropine type plasmid. Identification of open reading frames. J. Biol. Chem. 261 (1), 108–121.
Tan, M.-J., et al., 2008. Antidiabetic activities of triterpenoids isolated from bitter melon associated with activation of the AMPK pathway. Chem. Biol. 15 (3), 263–273.
Unay, J., Perret, X., 2020. A minimal genetic passkey to unlock many legume doors to root nodulation by rhizobia. Genes 11. https://doi.org/10.3390/genes11050521.
Usadel, B., et al., 2009. Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. Plant Cell Environ. 32 (12), 1633–1651.
Veena, V., Taylor, C.G., 2007. Agrobacterium rhizogenes: recent developments and promising applications. Vitro Cell. & Deve. Biol. - Plant 43 (5), 383–403.
Velázquez, E., et al., 2005. The coexistence of symbiosis and pathogenicity-determining genes in Rhizobium rhizogenes strains enables them to induce nodules and tumors or hairy roots in plants. MPMI (Mol. Plant-Microbe Interact.) 18 (12), 1325–1332.
Wang, S., et al., 2020. Evolutionary timeline and genomic plasticity underlying the lifestyle diversity in rhizobiales. mSystems 5 (4), e00438, 20.
White, F.F., Nester, E.W., 1980. Hairy root: plasmid encodes virulence traits in Agrobacterium rhizogenes. J. Bacteriol. 141 (3), 1134–1141.
White, F.F., et al., 1985. Molecular and genetic analysis of the transferred DNA regions of the root-inducing plasmid of Agrobacterium rhizogenes. J. Bacteriol. 164 (1), 33–44.
Young, J.M., 2001. Implications of alternative classifications and horizontal gene transfer for bacterial taxonomy. Int. J. Syst. Evol. Microbiol. 51 (3), 945–953.
Young, J.M., 2001. Classification, naming, and plant pathogenic bacteria — what is to be done? In: De Boer, S.H. (Ed.), Plant Pathogenic Bacteria: Proceedings of the 10th International Conference on Plant Pathogenic Bacteria, Charlottetown, Prince Edward Island, Canada, July 23–27, 2000. Springer Netherlands, Dordrecht, pp. 30–37.
Young, J.M., et al., 2003. Classification and nomenclature of Agrobacterium and Rhizobium – a reply to Farrand et al. Int. J. Syst. Evol. Microbiol. 53 (5), 1689–1695, 2003.
Young, J.M., 2008. Agrobacterium—taxonomy of plant-pathogenic Rhizobium species. In: Tzfira, T., Citovsky, V. (Eds.), Agrobacterium: from Biology to Biotechnology. Springer New York, New York, NY, pp. 183–220.
Zhou, Y., et al., 2016. Convergence and divergence of bitterness biosynthesis and regulation in Cucurbitaceae. Nature Plants 2 (12), 16183.