



## OPEN ACCESS

## EDITED BY

Carlos Henrique Meneses,  
State University of Paraíba, Brazil

## REVIEWED BY

Katia Scortecci,  
Federal University of Rio Grande do Norte,  
Brazil  
Jean Luiz Simoes-Araujo,  
Brazilian Agricultural Research Corporation  
(EMBRAPA), Brazil  
David Johnston-Monje,  
University of the Valley, Colombia

## \*CORRESPONDENCE

Omar Gätjens-Boniche  
✉ ogatjens@itcr.ac.cr

RECEIVED 10 June 2023

ACCEPTED 18 October 2023

PUBLISHED 29 November 2023

## CITATION

Gätjens-Boniche O, Jiménez-Madrugal JP,  
Whetten RW, Valenzuela-Díaz S,  
Alemán-Gutiérrez A, Hanson PE and  
Pinto-Tomás AA (2023) Microbiome  
and plant cell transformation trigger  
insect gall induction in cassava.  
*Front. Plant Sci.* 14:1237966.  
doi: 10.3389/fpls.2023.1237966

## COPYRIGHT

© 2023 Gätjens-Boniche, Jiménez-Madrugal,  
Whetten, Valenzuela-Díaz, Alemán-Gutiérrez,  
Hanson and Pinto-Tomás. This is an open-  
access article distributed under the terms of  
the [Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that  
the original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Microbiome and plant cell transformation trigger insect gall induction in cassava

Omar Gätjens-Boniche <sup>1\*</sup>, Jose Pablo Jiménez-Madrugal <sup>1</sup>,  
Ross W. Whetten <sup>2</sup>, Sandro Valenzuela-Díaz<sup>3</sup>,  
Alvaro Alemán-Gutiérrez<sup>1,4</sup>, Paul E. Hanson <sup>5</sup> and  
Adrián A. Pinto-Tomás <sup>6</sup>

<sup>1</sup>Laboratorio de Biología Molecular, Escuela de Ciencias Naturales y Exactas, Campus Tecnológico Local San Carlos, Instituto Tecnológico de Costa Rica, Alajuela, Costa Rica, <sup>2</sup>Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, United States, <sup>3</sup>Human Microbiome Research Program, Faculty of Medicine, The Helsinki University, Helsinki, Finland, <sup>4</sup>Laboratorio de Genómica y Biodiversidad, Facultad de Ciencias, Universidad del Bío-Bío, Chillán, Chile, <sup>5</sup>Escuela de Biología, Universidad de Costa Rica, San Pedro, San José, Costa Rica, <sup>6</sup>Center for Research in Microscopic Structures and Department of Biochemistry, School of Medicine, University of Costa Rica, San José, Costa Rica

Several specialised insects can manipulate normal plant development to induce a highly organised structure known as a gall, which represents one of the most complex interactions between insects and plants. Thus far, the mechanism for insect-induced plant galls has remained elusive. To study the induction mechanism of insect galls, we selected the gall induced by *Iatrophobia brasiliensis* (Diptera: Cecidomyiidae) in cassava (Euphorbiaceae: *Manihot esculenta* Crantz) as our model. PCR-based molecular markers and deep metagenomic sequencing data were employed to analyse the gall microbiome and to test the hypothesis that gall cells are genetically transformed by insect vectored bacteria. A shotgun sequencing discrimination approach was implemented to selectively discriminate between foreign DNA and the reference host plant genome. Several known candidate insertion sequences were identified, the most significant being DNA sequences found in bacterial genes related to the transcription regulatory factor CadR, cadmium-transporting ATPase encoded by the *cadA* gene, nitrate transport permease protein (*nrtB* gene), and arsenical pump ATPase (*arsA* gene). In addition, a DNA fragment associated with ubiquitin-like gene *E2* was identified as a potential accessory genetic element involved in gall induction mechanism. Furthermore, our results suggest that the increased quality and rapid development of gall tissue are mostly driven by microbiome enrichment and the acquisition of critical endophytes. An initial gall-like structure was experimentally obtained in *M. esculenta* cultured tissues through inoculation assays using a *Rhodococcus* bacterial strain that originated from the inducing insect, which we related to the gall induction process. We provide evidence that the modification of the endophytic microbiome and the genetic transformation of plant cells in *M. esculenta* are

two essential requirements for insect-induced gall formation. Based on these findings and having observed the same potential DNA marker in galls from other plant species (ubiquitin-like gene *E2*), we speculate that bacterially mediated genetic transformation of plant cells may represent a more widespread gall induction mechanism found in nature.

#### KEYWORDS

*latrophobia brasiliensis*, *Manihot esculenta*, plant galls, metagenomics, induction mechanism, genetic transformation, endophytes

## Introduction

Insect galls are abnormal structures developed by the presence and stimuli of insects in the host plant. Insect-induced plant galls are specialised plant tissues with an organised arrangement of cells and predetermined growth. The size, structure, and metabolism of galls are under the control of gall-forming insects and host plant species (Rohfritsch and Shorthouse, 1982; Leitch, 1994; Raman, 2011).

Insect gall tissues exhibit biochemical and cytological modifications that provide them with a higher nutritional quality than the surrounding plant tissue, thus facilitating a continuous source of food and additional benefits to the inducing insect (Koyama et al., 2004; Nabity et al., 2013; Ferreira et al., 2017). Most galls contain highly specialised tissue known as nutritive tissue, characterised by high concentrations of sugar (Nogueira et al., 2018), lipids, proteins, nitrogen, and a variety of other compounds (Shorthouse and Rohfritsch, 1992; Huang et al., 2015; Isaias et al., 2018).

The ability to induce galls in plants has emerged several times among and within insect orders, with representatives of gall-inducing species currently known in Diptera, Hymenoptera, Hemiptera, Coleoptera, Lepidoptera, and Thysanoptera (Mani, 1992; Espírito-Santo and Fernandes, 2007). In each order, the ability to form such structures appears to have an independent origin (Nyman and Julkunen, 2000; Ronquist et al., 2015).

In addition to insects, plant galls can be induced by mites and nematodes (Favery et al., 2016; de Lillo et al., 2018; Harris and Pitzschke, 2020; Olmo et al., 2020; Desnitskiy et al., 2023). Some fungi and bacteria species and even some viruses can also induce primary gall-like growths or, simply, neoplasm formation with low levels of cell differentiation (Ananthkrishnan, 1998; Raman, 2011; Gätjens-Boniche, 2019; Harris and Pitzschke, 2020). Examples of gall-like growths induced by microorganisms include *Agrobacterium tumefaciens* (crown gall), *Rhodococcus fascians*, *Pseudomonas savastanoi*, *Xanthomonas citri*, *Pantoea agglomerans*, *Taphrina betulina* (witches broom), and *Ustilago esculenta* (Swarup et al., 1991; Jump and Woodward, 1994; Chalupowicz et al., 2009; You et al., 2011; Dolzblasz et al., 2018; Harris and Pitzschke, 2020). Of these, *A. tumefaciens* is the best studied, as its habit of genetically transforming plant cells has found extensive use in plant biotechnology (Kavipriya et al., 2019; Song et al., 2019; Lian et al., 2022). Because *A. tumefaciens*-mediated genetic transformation of

plant cells is the most understood mechanism of plant gall formation (Chou et al., 2022; Hopp et al., 2022; Azizi-Dargahlou and Pouresmaeil, 2023), this constitutes the best referenced system to propose alternative induction mechanisms in which complex gall formation caused by insects such as cynips (Hymenoptera) and cecidomyiids (Diptera) (Sinnott, 1960; Raman, 2011) is also the result of a based plant cell genetic transformation.

A large number of endosymbiotic bacteria in different insect groups, including gall-inducing insects, have been reported in several studies (Campbell et al., 2015; El-Sayed and Ibrahim, 2015; Gutzwiller et al., 2015; Michell and Nyman, 2021; Yang et al., 2021; Coolen et al., 2022; Yang et al., 2022). Insect-associated microorganisms could be important mediators of interactions between insects and plants (Hammer and Bowers, 2015; Sugio et al., 2015; Wielkopolan and Jakubowska, 2021; Coolen et al., 2022). Symbiotic relationships between inducing insects and microorganisms have been hypothesised to be involved in plant gall development (Hansen and Moran, 2014; Tooker and Helms, 2014; Gätjens-Boniche, 2019; Klimov et al., 2022). Delivery of bacteria under natural or artificial conditions by insect vectors has been reported in many insect-plant interactions (Zeidan and Czosnek, 1994; Galambos et al., 2021; Wielkopolan and Jakubowska, 2021; Ratcliffe et al., 2022). Insect-vectored bacteria in plants have been described in well-known systems such as the Huanglongbing (HLB) disease of citrus caused by the phytopathogenic bacterium *Candidatus Liberibacter asiaticus* (CLAs), which is transmitted by the psyllid *Diaphorina citri*. Likewise, acquisition and effective delivery of *A. tumefaciens* by the whitefly *Bemisia tabaci* was demonstrated by Zeidan and Czosnek (1994).

It has been hypothesised that phytohormone elicitor molecules delivered by the insect inducer (Rohfritsch and Shorthouse, 1982; Tooker and Helms, 2014; Ponce et al., 2021) or indirectly by an associated microorganism (Giron et al., 2013; Bartlett and Connor, 2014; Giron et al., 2016), and effector proteins secreted by the gall-inducing insect (Zhao et al., 2015; Cambier et al., 2019; Zhao et al., 2019; Korgaonkar et al., 2021) may be the main triggering stimuli responsible for the gall induction process.

Cassava (*Manihot esculenta* Crantz) is a widely cultivated crop in Africa, Asia, and Central and South America that provides an important food source for millions of people worldwide. Cassava plants can be grown yearlong in the tropics. Moreover, it can be

easily propagated in greenhouse and *in vitro* conditions. Owing to its versatility and potential, plant breeding programmes and international consortiums have invested significant resources in understanding cassava genetics. This has led to the creation of genetic linkage maps and chromosome-scale genome assembly (Wang et al., 2014; ICGMC (International Cassava Genetic Map Consortium), 2015; Lyons et al., 2022). Cassava plants are subject to gall formation, particularly cylindrical galls induced by the Cecidomyiidae, *Iatrophobia brasiliensis* (Montaldo, 1977; Rivera Hernández, 2011). However, detailed studies are lacking on the induction and formation processes of this gall.

Here, we tested the hypothesis proposed by Gätjens-Boniche (2019), who postulated that characteristic hyperplasia in the initial phase of gall induction can be triggered by the insertion of exogenous genetic elements into the genome of plant cells through an endosymbiotic bacterium originating from the inducing insect. To test this hypothesis, we employed a combination of genetic marker analyses, metagenomic analyses, and experimental gall induction in cassava plants. Taken together, our findings provide support for the role of microorganisms and genetic transformation in gall induction mechanisms.

## Materials and methods

### Galling insect model

We chose the cylindrical gall induced in Euphorbiaceae *M. esculenta* Crantz (cassava) (Figure 1A) by the Cecidomyiidae *I. brasiliensis* (Montaldo, 1977; Rivera Hernández, 2011) as our biological model (Figure 1B). The cassava genome has been well documented (Wang et al., 2014; ICGMC (International Cassava Genetic Map Consortium), 2015; Lyons et al., 2022), allowing the identification of potential alterations or insertions from exogenous sources.

Additionally, six prosoplasmic gall morphotypes (greater structural complexity) and two kataplasmic galls (low structural complexity) were selected to validate the general induction hypothesis (Figure 1D). For details, see Supplementary Text.

Samples were obtained from several plant specimens collected in the Guanacaste Conservation Area, Guanacaste, Costa Rica, as well as around the Santa Clara community, San Carlos, Alajuela, Costa Rica. A voucher specimen of each host plant species and galls was deposited at the Cecidiarium (specialised gall herbarium), established at the Instituto Tecnológico de Costa Rica.

### Microdissection of salivary glands from *Iatrophobia brasiliensis* larvae

Larval salivary glands from *I. brasiliensis* were extracted by microdissection in a biosafety flow-hood (High Ten, Model 3BH-24), under sterile conditions (Figure S1). Larvae were surface decontaminated according to the protocol established by Zahner et al. (2008) with some modifications. Consecutive rinses

were done for 2 min in 3% sodium hypochlorite and then 70% ethanol, followed by a rinse in sterile nuclease-free water. All dissected salivary glands were collected and employed in the DNA extraction. All life stages were surface cleaned.

### Isolation and characterisation of endosymbiotic insect bacteria

Two colony-forming units (CFUs) were preliminarily isolated from the larval head of *I. brasiliensis* under sterile conditions. The dissected larvae were in instars one or two. The larvae were extracted live from the inner chamber of the gall, and only active larvae were collected. Different parts of the larva, such as the head and segments of the digestive system, were placed in YEB 1× culture medium (Piñol et al., 1996) to obtain a primary bacterial culture. Serial dilutions were made from the original culture to obtain single colonies (CFUs). The cells were recovered and maintained in YEB 1× solid and liquid suspension to 4–8°C and in glycerol (20% w/v) at –80°C. Each of the CFUs was named isolated symbiotic bacteria (ISB).

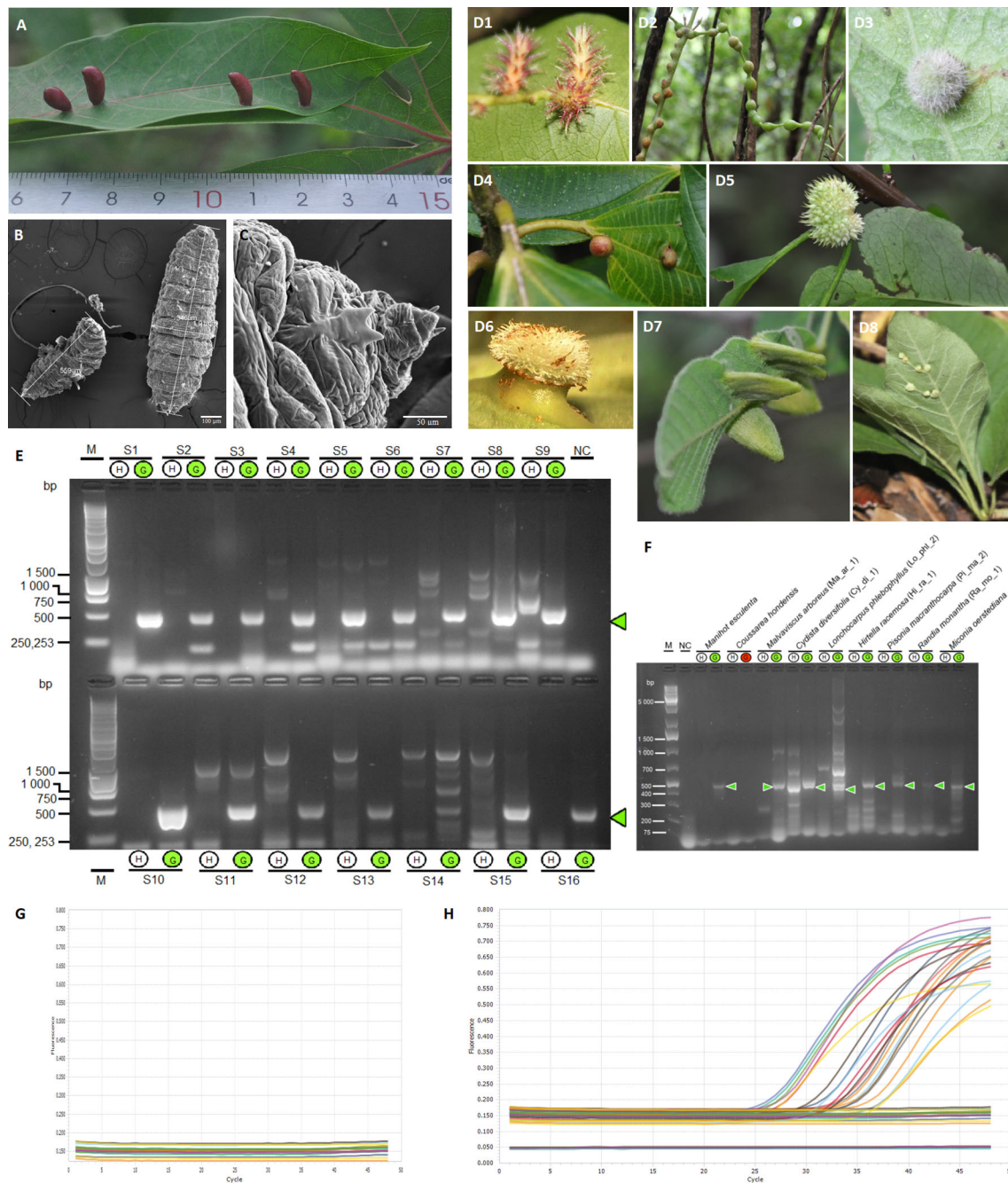
ISB 2 formed short rods and macroscopically coccoid-like elements and produced round, entire, convex, orange/pinkish colonies with smooth matte surfaces on YEB 1× Bacto agar medium after 4–5 days of incubation at 26°C. The bacterium showed features of a pinkish, flat, circular, slightly irregular border, variable length with curved shapes in some of them, and a mucous texture; it was a Gram-positive bacillus (Figure S2).

Molecular characterisation of bacteria was done through 16S sequencing using the procedures described by Rainey et al. (1996) and Hanshew et al. (2013), with some modifications. Taxonomic identification of the endosymbiotic bacteria was done using Kraken2 (Wood et al., 2019), the same for the endophytic bacteria, as described later. Details of the analytical pipeline are described below. Bacterial strains were identified by 16S gene sequencing before each inoculation assay to induce gall formation in the selected cassava tissues. We prepared templates for PCR either by total DNA extraction or by lysis of CFU bacteria growing in a solid culture medium. Bacterial cells were diluted in Elution Buffer (10 mM Tris, pH 8.0) (Qiagen, Hilden, Germany). Samples were lysed by vortexing and heating at 95°C for 5 min. The filtrate was then centrifuged at 8,000 g for 5 min.

### Isolation of culturable microorganisms associated with plant gall tissue

We isolated CFUs that potentially represent endophytic bacterial strains and one fungal growth from surface-sterilised internal sections of *M. esculenta* gall tissue sections. Seven were selected for high-throughput sequencing (HTS). Each of these CFUs was named an isolate of endophytic bacteria (IEB).

Surface tissue was sterilised following the methodology described for DNA purification. CFUs grown directly from gall slice explants were carefully collected by a sterile bacteriological



**FIGURE 1**

Identification of a potential universal marker in different gall systems. **(A)** Gall induced by the Cecidomyiidae *latrophobia brasiliensis* in *Manihot esculenta* (Montaldo, 1977). **(B)** Scanning electron micrograph of *I. brasiliensis* larva. **(C)** Scanning electron micrograph of the ventral area near the head of *I. brasiliensis* larva. **(D1)** Gall induced in *Hirtella racemosa* Lam. (Chrysobalanaceae), morphotype Hi\_ra\_1, by an unidentified Cecidomyiidae. **(D2)** Gall induced in *Cydista diversifolia* (Kunth) Miers (Bignoniaceae), morphotype Cy\_di\_1 by an unidentified Cecidomyiidae. **(D3)** Gall induced in *Malvaviscus arboreus* Dill. ex Cav. (Malvaceae), morphotype Ma\_ar\_1, by an unidentified Cecidomyiidae. **(D4)** Gall induced in *Miconia oerstediana* (Melastomataceae) by an unidentified Cecidomyiidae. **(D5)** Gall in *Pisonia macranthocarpa* (Donn. Sm.) Donn. Sm. (Nyctaginaceae), morphotype Pi\_ma\_4, induced by an unknown insect species. **(D6)** Gall in *Coussarea hondensis* (Standl.) C.M. Taylor & W.C. Burger (Rubiaceae) induced by an unknown insect. **(D7)** Gall induced in *Lonchocarpus phlebophyllus* Standl & Steyerem. (Fabaceae), morphotype Lo\_ph\_1, induced by a Psyllidae. **(D8)** Gall induced in *Randia monantha* Benth. (Rubiaceae), morphotype Ra\_mo\_1, induced by an unknown insect species. **(D2–D8)** Gätjens-Boniche et al. (2021). **(E)** Agarose gel electrophoresis of DNA fragments (specific gall fragment marker, SGF) amplified by PCR, comparing healthy leaf tissue DNA samples (**H**) and gall tissue DNA samples (**G**). Lane M, molecular weight marker (1 kb ladder); line NC, negative control (reagents only); lines S1–S16, samples of healthy leaf and gall tissues growing in the same plant organ (pair-compared). **(F)** Gel electrophoresis of PCR products using primers for the specific gall fragment marker (SGF) in gall morphotypes of different host plant species. Lane M, molecular weight marker (Gene Ruler 1 KB Plus); line NC, negative control (reagents only); lines 3–20, samples of healthy leaf and gall tissues of different plants growing in the same plant organ (pair-compared). The green circle indicates a positive sample for the specific gall fragment amplification (**E**, **F**). **(G, H)** Real-time PCR by Taq Man Probe showing the detection of the specific gall fragment from gall DNA samples of *Manihot esculenta* (amplification plot **H**) and from healthy leaf samples (amplification plot **G**). Each trace shows the  $\Delta Rn$  (normalised net fluorescence signal of the PCR product) plotted against the number of PCR cycles.

loop in a sterile biosafety chamber. Primary cultured bacteria were diluted in liquid medium YEB 1× (Piñol et al., 1996) until CFUs were obtained on solid medium. This process was repeated two to three times. Bacteria in cultured media were grown for 3–4 days at 26°C. Subsequently, bacterial cells were recovered and maintained in YEB 1× solid medium and suspended in liquid medium YEB 1× at 4–8°C and in glycerol (20% w/v) at –80°C.

## DNA purification

### Isolation of total genomic DNA from plant tissues

All tissues were sampled from plants morphologically identifiable as *M. esculenta*. Healthy leaf tissue from *M. esculenta* was checked in a stereoscope to avoid visible microgalls or some other type of foliar damage or disturbance. Prior to DNA extraction, the galls were dissected by carefully splitting them in half without damaging the larva present in the internal chamber, which was removed from the tissue. Healthy leaves and galls were washed with alkaline liquid soap and Triton X-100 detergent (dissolved in sterile water). The plant material was vortexed three times with 70% ethanol (v/v) and sterile water. Subsequently, the leaves and galls were surface sterilised according to the methodology suggested by Meyer and Hoy (2008) and Mueller et al. (2004), with some modifications. Plant materials were immersed for less than 30 s in sodium hypochlorite sequential solutions (3% NaOCl, 6% NaOCl, and 3% NaOCl) and rinsed with sterile water. The sterilisation process was conducted in a biosafety flow-hood (High Ten, Model 3BH-24). After cleaning and sterilisation, all samples were immediately stored at –80°C until DNA extraction.

Genomic DNA extraction was performed according to the methodology established by Dellaporta et al. (1983), with modifications and following column precipitation steps as described by the Dneasy Power Plant Pro Kit protocol (Mobio/QIAGEN, Carlsbad, CA, USA). The phenolic separation solution (PSS) in the Dneasy Power Plant Pro Kit was used during maceration to avoid DNA methylation. Total genomic DNA was extracted from both healthy leaf and gall tissues (different morphotypes). In the case of galls, DNA samples were collected from individual and pooled galls growing on the same leaf.

### Genomic DNA isolation from salivary glands of inducing insects, *Iatrophobia brasiliensis*

Larval samples for DNA extraction were collected by dissecting galls from *M. esculenta*. Pools of 10 larvae of *I. brasiliensis* were decontaminated by being exposed to 3% NaOCl for 3–4 min, followed by successive 70% alcohol and water rinses under a biosafety flow-hood (High Ten, Model 3BH-24). Genomic DNA purification from the salivary glands of the insect larvae was performed using the Power Soil DNA Isolation Kit (Mobio/QIAGEN, Carlsbad, CA, USA) with the additional steps, mainly K proteinase (100 mg/mL) treatment and incubation at 65°C for 1 h. Furthermore, salivary gland DNA samples were extracted three times with phenol-chloroform isoamyl alcohol (PCI).

### DNA isolation from endophyte and putative insect endosymbiotic bacteria

Total genomic DNA from endosymbiotic bacteria isolated from *I. brasiliensis* insects was extracted using a Dneasy Blood and Tissue Kit (Qiagen, Hilden, Germany). The same DNA extraction protocol was used for endophytic bacteria strains grown from surface-sterilised internal sections of *M. esculenta* gall tissue sections. Plasmid DNA from the isolated bacteria strains was carried out by alkaline lysis, as described by Li et al. (1995) with some modifications. The integrity and yield of bacterial genomic DNA and wild-type plasmids were checked by 0.8% agarose gel electrophoresis.

For all genomic DNA samples (plant, insect, and bacteria), quality was assessed by 260/280 and 260/230 ratios measured on a NanoDrop 8000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA).

### Modified RAPD methodology for discovering gall-associated molecular markers

Samples of healthy leaf and gall tissues growing in the same plant organ were compared to detect differentially amplified DNA fragments in the DNA of the gall and not present in the same healthy tissue. The theoretical approach of this methodology is shown in Figure S3A. Assays were made under standard RAPD methodology conditions with the commercial random primers OPC-06, OPI-04, OPA-03, OPD-18, OPD-03, OPE-06, OPA-17, and OPB-04 (Operon Technologies, Alameda, California, USA). In these analyses, a modified methodological approach was also carried out by simultaneously using a combination of two non-random primers of conserved sequences from *A. tumefaciens* genes along with a non-standard RAPD thermal profile. Primers from conserved *A. tumefaciens* genes were used because the genetic transformation of plant cells mediated by this bacterium is the best-studied genetic transformation in a plant system. Decamer primers derived from conserved sequences of the Isopentyl Transferase Gene (*ipt*) and from the *iaaM* gene (Tryptophan 2-monooxygenase) harboured in the transfer DNA of the Ti plasmid of *Agrobacterium* species generated the highest number of differentially amplified fragments in previous RAPD assays. Primers were designed from the alignment of conserved regions in different species that harbour these genes using DNA Star Lasergene 99 (Madison, Wisconsin, USA) and BioEdit version 4.8.10.1 (Hall, 1999). Nucleotide sequences for these genes were obtained from the National Center for Biotechnology Information ([www.ncbi.nlm.nih.gov/pubmed/](http://www.ncbi.nlm.nih.gov/pubmed/)). Accessions for the *iaaM* genes were M91609, Z18270, X77327, U04358, and L33867, and for the *ipt* genes, they were X77327, X53945, X17428, M91610, and Z46375. RAPD reactions were performed on 200-μL sterile, pyrogen-safe, thin-walled plastic tubes for PCR RNase-DNase (Axygen, CA, USA), using 0.5 units of Dream Taq (Fermentas Life Sciences, Lithuania), PCR 1× (750 mM Tris-HCl [pH 8.8], 0.2 mM dNTPs, 1.5 mM MgCl<sub>2</sub>, 0.5 μM of each primer, and 10 ng of

DNA), adjusted to a final volume of 25  $\mu$ L with nuclease-free sterile water (Promega Corporation, Madison, Wisconsin, USA). The primer set used was ipt forward, 5'-CGGTGAACGA-3' and iaaM reverse, 5'-TCCAATTTCT-3'. DNA was initially denatured for 3.5 min at 95°C, followed by 15 cycles of 95°C for 30 s, 34.5°C for 30 s, and 72°C for 2 min. This was followed by 35 cycles of 30 s denaturation at 95°C, 30 s annealing at 46°C, and 2 min elongation at 72°C, with a final elongation step of 72°C for 7 min. Reactions were carried out in a thermocycler PTC-200 DNA Engine (MJ-Research, Waltham, Massachusetts, USA). Samples of healthy and gall tissues growing in the same plant were compared in pairs. Each primer set and RAPD condition was carefully tested more than three times using 20 to 30 samples. Reactions were performed in a flow-hood (High Ten, Model 3BH-24).

RAPD products were analysed by 1%–1.5% agarose gel electrophoresis with 0.5  $\times$  TBE and 1 $\times$  Gel Red at 75 V (Electrophoresis chamber and Power Pac 300 Bio-Rad, Hercules, California, USA). Amplicons were separated in a MultiNA automated system for DNA and RNA microchip analysis (MultiNA-Shimadzu, Tokyo, Japan). PCR products were cleaned using Promega Wizard SV gel and the PCR clean-up system, as per the manufacturer's directions (Promega, Madison, WI).

The DNA gall fragments differentially amplified from the insect gall tissue of cassava by modified RAPD methodology were sequenced at the Centro de Investigaciones en Biología Celular y Molecular, Universidad de Costa Rica, San José, Costa Rica. Sequencing reactions were performed using dideoxynucleotide chain termination with the BigDye<sup>TM</sup> Terminator Kit (Applied Biosystems, USA) and 5 pmol of each ipt forward and iaaM reverse sequencing primers and analysed with an ABI Prism® 3700 Automated Sequencer (Applied Biosystems, USA). Sequenced fragments were aligned using DNA Star Lasergene 99 (Madison, Wisconsin, USA) and BioEdit version 4.8.10.1 (Hall, 1999).

## PCR amplification and sequencing for the specific gall marker

The DNA gall fragments differentially amplified from gall samples were used as templates to design a potential gall molecular marker based on our RAPD results. Primers were designed using Primer3 v.4.1.0 (Koressaar and Remm, 2007; Untergrasser et al., 2012). The primer sequences were 5'-CTT GAC ATG TTC TGG AGC GG-3' for the forward primer (Primer\_Gall-Forward) and 5'-AAC GAG CGT GGT ACT GTG AT-3' for the reverse primer (Primer\_Gall-Reverse) (Invitrogen, Carlsbad, CA, USA). The expected amplicon size was 471 bp. Primers were tested with DNA samples from both healthy and gall tissues from *M. esculenta* plants and other gall morphotypes, as well as with DNA extracted from insect salivary glands. Subsequently, the target gene was amplified in the isolated wild-type plasmids from two putative insect endosymbiotic bacteria of the genus *Rhodococcus* and *Pseudomonas* and in wild-type plasmids of all endophytic bacteria isolated from the cassava gall tissue. PCR was carried out with 1 $\times$  PCR buffer (750 mM Tris-HCl [pH 8.8],

200 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 0.1% Tween 20), 1.5 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 2 U Taq Polymerase (Thermo Fisher Scientific, Wilmington, DE, USA), 0.5  $\mu$ M forward and reverse primers (Invitrogen, Carlsbad, CA, USA), 10 ng of sample DNA, and H<sub>2</sub>O (Promega Corporation, Madison, Wisconsin, USA) for a final reaction volume of 25  $\mu$ L. Amplification was performed in a thermal cycler PTC-200 DNA Engine (MJ-Research, Waltham, Massachusetts, USA) using the following cycling conditions: initial denaturation at 95°C for 3.5 min.; 40 cycles of 95°C for 30 s, 62°C for 45 s, and 72°C for 1 min; and final extension at 72°C for 5 min. The transition temperature between each step was 1°C/s. The assays were repeated multiple times for 170 galls and healthy tissue samples. Three samples with varying amounts of salivary glands were used in total, as were two bacterial plasmid samples. PCR reactions for these samples were repeated at least six times.

The PCR products were separated on a 1.5% agarose gel with 1  $\times$  TAE and 1 $\times$  Gel Red at 75 V (Electrophoresis chamber and Power Pac 300 Bio-Rad, Hercules, California, USA). The expected amplicons were excised from the gel and purified using the QIAquick Gel Extraction Kit, following the manufacturer's directions (QIAGEN, Hilden, Germany). Samples were sequenced using the Sanger method through a Macrogen service provider (Macrogen Inc., Seoul, Korea). Samples were purified and prepared for sequencing according to the methodology established by Macrogen. Sequencing of nonspecific amplified fragments for PCR was attempted; however, good quality and long base sequences could not be obtained, except for two of the samples. Sequenced fragments were aligned using DNA Star Lasergene 99 (Madison, Wisconsin, USA) and BioEdit version 4.8.10.1 (Hall, 1999).

All amplified fragments were aligned to the SGF consensus sequence using BioEdit version 4.8.10.1 (Hall, 1999) and Jalview version 2-a, 2.11.1.5 (Waterhouse et al., 2009). For salivary gland amplified fragments, an overlap of bases was frequent in several of the sequenced PCR fragments, which could show variants of the target DNA sequence in this insect tissue. A similar trend was observed in one of the colonies from endosymbiont *Rhodococcus* isolated from the inducer insect when their wild-type plasmids were analysed and linked to the DNA amplification profile, which could indicate polymorphic variants of their wild-type plasmids.

Both DNA fragments amplified from gall samples (RAPD modified technique) and the SGF were analysed using BLAST [National Center for Biotechnology Information (NCBI) ([www.ncbi.nlm.nih.gov/pubmed/](http://www.ncbi.nlm.nih.gov/pubmed/))], Integrated Microbial Genomes & Microbiomes (IMG/M) system (<https://img.jgi.doe.gov/>).

## Detection of the specific gall fragment by real-time qPCR

A real-time PCR marker was designed and tested in healthy and gall tissue using a TaqMan probe with homology to the consensus of differentially amplified DNA from galls. TaqMan-based qPCR was carried out in a 20- $\mu$ L reaction mixture containing Go Taq Master Mix 2X (Promega, USA) and 1 $\mu$ L of 20X TaqMan<sup>®</sup> Gene Expression Primer/Probe Mix (Applied Biosystem, CA, USA).

Primers and the TaqMan probe were designed and synthesised from the specific gall fragment target with 0.5  $\mu\text{M}$  of each primer (AI7ZYBP\_Forward: 5'-TGTTGCTGCACAGAGTTCT-3' and AI7ZYBP\_Reverse: 5'-GGCTTGAGTGCTTCGATTTCG-3'), 0.25  $\mu\text{M}$  of the MGB probe, labelled with FAM reporter dye at the 5' and non-fluorescence Quencher TAMRA at the 3' end (AI7ZYBP\_M\_TCTGCCACCGACCCT\_NFQ), and 2.5  $\mu\text{L}$  of 5 ng/ $\mu\text{L}$  of extracted DNA. qPCR cycling conditions included initial denaturation at 95°C for 3 min and 35 cycles of denaturation at 95°C for 30 s and primer annealing and extension steps together at 60°C for 60 s. Three replicates of the negative template control consisting of nuclease-free water for molecular biology grade reactions were included in each amplification assay. Assays were performed on a LightCycler 96 Real-Time PCR System (Roche Diagnostics, Risch-Rotkreuz, Switzerland). We tested over 30 healthy tissue and gall samples and obtained an amplification average of 90% for the PCR marker or amplification signal in the case of qPCR. The assays were repeated more than six times for the same healthy leaf and gall samples.

## Bioinformatic analysis

### Library preparation and sequencing

Short-read libraries were prepared using Illumina's DNA TruSeq Nano Library preparation kit, following the manufacturer's instructions. For the leaf and gall genomic DNA samples, 150-bp single-ended (SE) libraries were prepared and sequenced on the Illumina HiSeq2500 platform using the SBS sequencing kit version 4. For the bacterial isolates' genomic DNA samples, 250-bp paired-ended (PE) libraries were prepared and sequenced in the Illumina MiSeq platform using Reagent kit version 2. Both library preparation and next-generation sequencing were performed by the NC State University Genomic Sciences Laboratory (Raleigh, NC, USA). Raw data can be retrieved from the Short Read Archive (SRA) under the Bioproject accession number: [PRJNA905450](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA905450).

No sequencing was performed for any endosymbiotic or endophytic bacteria wild-type plasmids due to the stability and integrity of the plasmid DNA during the fragmentation procedure.

### Host discriminant genomic analysis

The aim of this analysis was to identify potential genes or insertion sequences from bacteria in the cassava genomic DNA samples (genotype Valencia); these sequences were believed to induce gall formation in otherwise healthy plants. To do so, a stepwise filtering strategy was implemented. First, sequence read quality was assessed with FastQC (Andrews, 2010); when needed, index, adapters, and low-quality sequences were eliminated. After quality control, reads from both healthy and gall tissues were mapped against the cassava reference genome CV AM560-2 (Phytosome genome ID: 520; Bredeson et al., 2016) using BBmap, part of the BBTools suite (Bushnell, 2015), with the semiperfect mode option. Unmapped reads from healthy leaf tissue and leaf

galls were compared against each other, and the shared reads were filtered using BBduk, part of the BBTools suite. The remaining gall reads, unmapped to the reference genome and unmatched to healthy leaves according to defined parameters, were *de novo* assembled into contigs using SPAdes (Prjibelski et al., 2020). A MegaBlast of the selected contigs was performed, setting the minimum length to 200 bp and the percentage of homology to  $\geq 95\%$ . The resulting contigs, the product of gall exclusive reads, were then mapped against the cassava reference genome using the Burrows–Wheeler Aligner (BWA-MEM) (Li and Durbin, 2009). Alignments with low mapping quality ( $\geq 20$ ) were filtered using SAMtools (Li et al., 2009). Contigs with matching sequences in either their 5', 3', or both ends but non-matching sequences in their core, based on the CIGAR string, were selected as potential foreign DNA insertion sites. Reads were mapped back to the gall hybrid/fusion contigs containing potential foreign DNA insertion sites to assess coverage. Contigs with less than three cover reads were discarded, except for three of them that showed significant annotations with known gene sequences.

We also analysed the putative gall insertion sequence marker (named gall fragment) experimentally in plasmids isolated from all seven isolated endophytic bacteria and the two bacteria isolated from the larval insect head of *I. brasiliensis* and *in silico* by bioinformatic tools in the assembled genomes of the same sequenced bacteria species. In particular, for the *in silico* analysis, isPCR and the primer sequences designed for the gall fragment marker did not show any amplification from the cassava reference genome, nor did it produce amplification products for the insect endosymbiotic bacteria genomes.

### Metagenomic sequence analysis and taxonomic profile assignment

Shotgun metagenomic sequencing approach was used to analyse the microbiome associated with cassava galls and healthy leaf tissues. Taxonomic profiles were performed using Kraken2 (Wood et al., 2019) and its standard database, which includes bacteria, archaea, viruses, and eukaryotic genomes (<https://benlangmead.github.io/aws-indexes/k2>). Analyses were applied to reads data (QC filtered reads) from the sequenced gall and healthy tissue samples with an abundance filter of 10K reads.

Taxonomic profiles were also carried out to identify unique reads in the gall tissue following the host discriminant genomic analysis (HDGA) methodological approach to determine their taxonomic origin. Reads were mapped against the cassava genome, retaining only unmapped reads. We then clustered the sequences to obtain unique reads by sample only. This finding was further confirmed by contrasting the resulting taxonomic profiles of healthy samples with gall samples.

After a detailed analysis of the taxonomic profiles was carried out to identify the likely contaminants introduced during processing or due to any potential contaminant event during the final stage of DNA purification of cassava healthy leaves and gall samples, five bacteria taxa were removed as possible external contamination in all sequenced samples.

## Bacteria endophytic condition determination by synteny analysis

We applied synteny bioinformatic analysis to confirm whether isolated bacteria from both gall tissue and the inducing insect were components of the endophytic microbiome in this structure. To carry out this analysis, exclusive gall reads and bacteria samples were *de novo* assembled using SPAdes (Bankevich et al., 2012) in their metagenome mode for gall reads and isolate mode for bacteria samples. To confirm synteny, contigs from both sides were compared using BLAST+ (Camacho et al., 2009).

## Genome annotation and functional analyses

To analyse potential differences in terms of function in bacterial groups (endophytic vs. endosymbiont), an enrichment analysis was performed using the Gene Ontology (GO) terms included in the TopGO R package (Alexa, 2022) and a cluster of gene ontology approach (COG) (Galperin et al., 2019). To carry out the analysis, bacterial genomes previously assembled were annotated using prokka (Seemann, 2014). To obtain their corresponding GO and COG terms, the proteins were further annotated using eggnoG (Huerta-Cepas et al., 2019). To obtain the raw counts of genes, original assembly reads were mapped back to the annotated genes using BWA MEM (Li, 2013) and filtered with SAMtools (Li et al., 2009), including flags  $0 \times 8$  for non-paired reads and  $0 \times 2$  for properly paired reads. Then, raw counts were obtained using the HTSeq-count from the HTSeq framework (Anders et al., 2015). Finally, differential abundance analysis was performed on shared genes using DESeq2 (Love et al., 2014), allowing their differentiation in abundance post-analysis of their GO terms.

## Ab initio predictions

*Ab initio* prediction was performed on the differentially amplified fragments from *M. esculenta* gall tissues (consensus DNA sequence) and on amplified fragments from purified plasmids of the endophytes IEB 1-2, IEB 3-1, and IEB 5-1, as well as purified plasmids from ISB 1 and ISB 2 isolated from the larval head of the inducing insect. These fragments were amplified with specific gall primers as previously described in the PCR methodology section. *Ab initio* prediction was also carried out on the hybrid sequence fasta files. Gene predictions and their functions were performed using the pipelines of Rapid Prokaryotic Genome Annotation (Prokka) (Seemann, 2014) on the web server Galaxy ([www.usegalaxy.org](http://www.usegalaxy.org)). The pipeline of Prokka uses Prodigal and UniProt for the prediction of coding regions and functional similarity, respectively.

## Identification of gene segments

The identification of conserved genetic regions in all target sequences obtained by PCR and in the possible hybrid sequences of interest, obtained by high-throughput sequencing, was performed using the PipMaker program (<http://pipmaker.bx.psu.edu/cgi-bin/pipmaker?advanced>) to produce local alignments of the genetic sequences using BlastZ (Schwartz et al., 2003), and dot plots of the gap-free segments of the alignments were generated as unbroken

diagonal lines, indicating a high degree of identity across the genetic sequences (Schwartz et al., 2000).

## Inoculation assays for primary gall induction

Two bacteria grown in solid and liquid YEB 1× culture medium (Piñol et al., 1996) were used to inoculate emerging leaves from apical buds and young leaves of *M. esculenta* (Cassava). Bacterial strains isolated from the larval head of *I. brasiliensis*, subsequently identified as *Rhodococcus* (related to *Rhodococcus* sp. P-2 and *Rhodococcus erythropolis* species, according to the taxonomic profile), were cultured in YEB 1× solid medium and cultured at  $26 \pm 1^\circ\text{C}$  in a shaker to 110 rpm (Hotech® model 721-2T, Hotech Instruments Corp., New Taipei City, Taiwan) for 3–4 days in the dark. Subsequently, a single clone was used to inoculate 5 mL of YEB 1× liquid medium. After culturing under shaking conditions at  $26 \pm 1^\circ\text{C}$  for 2–3 days in the dark, a small bacterial suspension was transferred to 10 mL of YEB 1× liquid medium and transferred once again under a sterile biosafety flow-hood to YEB 1× solid medium and cultured for 2–3 days under the above conditions. Then, a single clone was used to inoculate 100 mL of YEB 1× liquid medium to the same preconditions. The endophytic bacterium *Pantoea ananatis*, one of the endophytic bacteria isolated from gall tissue, was used as a control bacterium. The management and culture of this bacterium were carried out under the same conditions described above.

*In vitro* micropropagated *M. esculenta* plants were used for bacterial inoculation assays under sterile conditions. Apical bud and leaf sections of 0.5–1.0 cm were used for inoculation. Explants were maintained first in solid MS medium for different periods of time. Prior to plant material co-cultivation with bacterial cultures, all explants were treated with slight mechanical abrasion using wet filter paper impregnated with ground glass. Just before inoculation, bacterial biomass was measured with a Lambda 25 spectrophotometer (Perkin Elmer–Applied Biosystems, Foster City, California, USA) and set to a  $\text{OD}_{600}$  value of 0.5–1.0. Then, *in vitro* plant material was submerged into bacteria cultured in YEB 1× liquid medium and shaken in the dark at  $28 \pm 1^\circ\text{C}$ , for 16–24 h. Inoculated explants were placed in sterile glass Erlenmeyer flasks and washed three times with sterile water. Excess water was removed, and explants were grown separately on MS solid (agar 5.7 g/L) or liquid medium without plant growth regulators. Half of the cultures were supplemented with  $3 \text{ mg L}^{-1}$  Kathon (only in solid medium), a bacteriostatic reagent used to inhibit bacterial growth in the cultured medium, under white, fluorescent light with an irradiance of  $27 \mu\text{mol m}^{-2} \text{ s}^{-1}$  at  $28 \pm 1^\circ\text{C}$  and 110 rpm in a shaker for the liquid medium (Hotech® model 721-2T, Hotech Instruments Corp., New Taipei City, Taiwan). Both types of explants not co-cultured with the bacteria but treated with mechanical abrasion were employed as controls. Thirty to fifty explants were used in each experiment, which was repeated five times in solid medium and two times in liquid medium. Kathon reagent was used only in the first assay carried out in a solid medium.



Inoculation of plants grown in soil under controlled greenhouse conditions was also attempted. The experimental conditions and bacterial growth were carried out using the same methodology. Groups of 50 plants per treatment grown in pots with soil were used to manually inoculate one to two new emerging leaves. Before bacterial inoculation, the leaf surface was carefully cleaned with 70% ethanol. Then, the emerging leaves were treated with a slight mechanical abrasion using a sterilised YEB 1× liquid medium containing ground glass, followed by rubbing the leaf surface by hand using clean latex gloves. After this, sterilised cotton impregnated with liquid medium containing the bacterial growth was placed and fixed on the prepared leaf surface for 1–2 days. The inoculated leaf area was covered with sterile gauze and sterile insulating plastic. This inoculation test was repeated three times.

A randomised block design was followed in this assay. However, because of the viability decline observed in cassava leaves under the established experimental conditions, only presence–absence results for gall-like structure and abnormal tissue development were displayed graphically. Infostat software version 2020p (<http://www.infostat.com.ar>) was employed to graph the results obtained from the experiments (Di Rienzo et al., 2020). Daily evaluations were performed in all treatments for 6 weeks. Data are shown as the mean and standard error ( $\pm$ SE) of tissue response due to inoculation with the bacteria. Taxonomic identification of the alleged endosymbiotic bacteria strain from the inducing insect (*Rhodococcus*, isolation ISB 2), used in these gall induction assays, was determined before each test by 16S gene sequencing and then confirmed by a taxonomic profile using genomic bioinformatics tools, as also performed for endophytic bacteria.

## Sample preparation for scanning electron microscopy and transmission electron microscopy

Samples of healthy leaves and galls of different sizes from *M. esculenta* plants, as well as the larvae of the inducing insect, were prepared as described by Sánchez et al. (2006), with the modification of fragmentation in liquid nitrogen and dehydration by passage through a gradient of acetone solutions. Sample preparation was performed at the Center for Research in Microscopic Structures, University of Costa Rica, San José, Costa Rica. Samples were observed using a scanning electron microscope (Hitachi S-3700, Tokyo, Japan) with an acceleration voltage of 15 kV.

For transmission electron microscopy (TEM) analyses, samples were fixed with a modified Karnovsky solution, then post-fixed with 2% osmium tetroxide for 1 h, submitted to contrast, and dehydrated in a gradient of acetone solutions (i.e., 30%, 50%, 70%, 90%, and 100%). After washing with 100% acetone, the samples were submitted to pre-infiltration with 1:1 Spurr resin/100% acetone for 5 h with shaking. The subsequent filtration of the samples was performed with pure resin for 12 h. The samples were then shaped and polymerised at 70°C for 3 days. The blocks were then cut into 70-nm-thick sections on an ultramicrotome (Leica Power Tome PC, Leica Microsystems GmbH, Wetzlar, Germany) with a diamond blade (45°), mounted on uncoated 200-mesh

copper grids, and stained with uranyl acetate and lead citrate. TEM observations were performed using a Hitachi model HT-7700 microscope (Tokyo, Japan) operating at 100 kV.

## Results and discussion

### Characterising a potential gall molecular marker

To assess our hypothesis, DNA samples from healthy leaf and gall tissues from the same plant were purified (Figures 1A, D1–D8). Samples were collected from healthy leaves of *M. esculenta* and galls induced by Cecidomyiidae *I. brasiliensis* (Figures 1B, C), and these were carefully cleaned and sterilised to guarantee total epidermis disinfection, including the inner gall chamber.

Possible exogenous DNA, or a DNA insertion sequence present only in gall cells but not in the healthy plant tissue of *M. esculenta* plants, was initially explored using specific PCR primers as a potential gall marker. The primer pair was designed based on the consensus DNA sequence resulting from the alignment of differentially amplified fragments obtained through a previously carried out modified RAPD assay.

A high number of RAPD assays were performed using different random primers; however, decamer primers derived from conserved sequences of the *ipt* and *iAAM* genes harboured in the transfer DNA of the Ti plasmid of *Agrobacterium* species and other related bacteria generated the highest number of differentially amplified fragments from gall samples. The quality of purified DNA, as well as the modified RAPD amplification conditions and thermal conditions, allowed us to obtain high reproducibility and reliability in the DNA profiles at different concentrations of analysed DNA. Analytical detections of the RAPD amplicons by gel electrophoresis and the Microchip Electrophoresis System for DNA/RNA (MultiNA) showed differentially amplified fragments of several sizes in all gall samples tested, often from 100 to 4,500 base pairs (bp) (Figures S3B, C). Using these approaches, we isolated and obtained the nucleotide sequences of four samples from the most common differentially amplified fragments derived from gall genomic DNA. Three of these sequences were identical, showing the same nucleotide sequence as the fragment of approximately 500 bp (Supplementary Data 1). The consensus DNA sequence obtained was used as a template to design and test specific PCR primers. For further information, see Supplementary Text.

The gall DNA fragment differentially amplified from gall samples (which we call a specific gall fragment, SGF) was isolated and sequenced. The expected PCR fragment was specifically amplified only in gall samples from *M. esculenta* (Figure 1E), total DNA from the inducing-insect salivary glands, and in the isolated wild-type plasmids from two putative insect endosymbiotic bacteria of the genera *Rhodococcus* (ISB 2 bacterial isolate) and *Pseudomonas* (ISB 1 bacterial isolate) (Figures 2A, S4). Moreover, PCR products with sizes between 350 and 600 bp were also amplified from wild-type plasmids of all endophytic bacteria isolated from cassava gall tissue (Figures 2A, S4). Additionally, specific gall PCR fragments showing a size similar to that expected



from cassava galls were selectively amplified from gall morphotype samples of eight different host plant species (Figure 1F). Many of these fragments were sequenced and then aligned to the consensus sequence of the differentially amplified fragment from the cassava gall. The aligned sequences showed high identity with the reference DNA fragment in all *M. esculenta* gall samples (99%–100%) (Figure 2C) and with the two IEB isolates 1-2 isolated from inner plant gall tissue (98.9%) and ISB 2 isolated from the larval insect head (98.4%) (Figure 2E). Furthermore, samples of other galls from *Cydista diversifolia* (97.4%) and *Hirtella racemosa* (99%) also showed high identity (Figure 2D), but a less conspicuous homology (between 48% and 53%) was detected from galls on *Lonchocarpus phlebophyllus*, *Malvaviscus arboreus*, and *Miconia oerstediana* (Figure 2D). This, in turn, suggests a diversification of the coding genetic element under different selection pressures for functional adaptation. Moreover, a repeatedly amplified DNA fragment from the salivary glands of a similar size to that expected showed an identity mean value comparable to the previous reference (49.3%), which could suggest that other bacteria in the salivary glands or an insect homologous gene might harbour a similar DNA sequence (Figure 2E).

A real-time PCR marker by the TaqMan probe was designed and tested following a similar approach used for end-point PCR assays with corresponding modifications. The results showed amplification signals only in the gall tissue (Figures 1G, H) and in plasmids purified from the isolated colonies of the insect endosymbiotic bacteria *Pseudomonas* sp. and *Rhodococcus* sp. (results not shown).

The differentially amplified gall DNA fragment from gall samples (SGF) did not show statistically significant similarity with any other reported gene according to the Basic Local Alignment Search Tool [BLAST, NCBI Genbank database, and Integrated Microbial Genomes & Microbiomes (IMG/M) system, <https://img.jgi.doe.gov>] and annotation analysis. Nevertheless, several results showed partial pairings with low to medium and often discontinuous length coverage with some ubiquitin-like genes, more specifically, the ubiquitin-like gene *E2*, a component of the ubiquitin-proteasome system (UPS), which has frequently been reported in different fungus species. This fragment showed partial homology to the Ubiquitin-conjugating enzyme E2 Q2 of the fungus *Fulvia fulva*, which was among the most significant (E-value = 6.37e-29, identity = 71.8%, 266 bp of length, accession number CP090172) (Figure 2B). However, we predicted that it was different enough to represent a new ubiquitin-like regulatory genetic element associated with the manipulation of the ubiquitin-proteasome system, used indirectly by the inducing insect via the bacteria to manipulate and redirect plant development during gall formation.

The ubiquitin gene family encodes peptides involved in protein-protein signalling and destination as a component of the basic cellular regulation machinery. These types of proteins regulate gene expression at the transcriptional (Adams and Spoel, 2018) and post-translational levels (Xu and Xue, 2019; Liu et al., 2020). The ubiquitin gene family has been reported as an essential part of molecular cell manipulation mechanisms in different pathogen and endosymbiotic-host interactions (Janjusevic et al., 2006; Vierstra, 2009; Park et al., 2012; Singer et al., 2013; Banfield, 2015; Kud et al., 2019). Ubiquitin-like

genes from bacterial secreted effector molecules with structural and/or functional similarity to UPS pathway components mimic and modify the host UPS system (Ramachandran et al., 2021), allowing the hijacking of the cellular machinery, as has been previously reported in the crown gall system induced by *A. tumefaciens* (Magori and Citovsky, 2012; Lacroix and Citovsky, 2015). The UPS system is considered the major protein turnover pathway found across all domains of life and is especially important in regulating almost all plant development signalling pathways, including hormone-mediated plant growth and development, as well as plant responses to stress (Santner and Estelle, 2010; Sadanandom et al., 2013; Shu and Yang, 2017; Adams and Spoel, 2018; Xu and Xue, 2019). Although less is known about UPS manipulation in galling insects, there is emerging evidence that plant development can be manipulated through the UPS system by effector molecules in the salivary glands of the gall inducer (Zhao et al., 2015).

## Genetic insertion events inferred from bioinformatic analysis provide evidence for plant cell transformation

Through a discrimination approach using shotgun metagenomic sequencing, we showed the presence of potential foreign-exclusive DNA within plant gall tissue. We named this methodological approach host discriminant genomic analysis (HDGA) (Figure S5). HTS data from healthy leaves and gall tissue (16× sequencing depth) were processed to separate, assemble, and analyse the DNA sequence reads different from the referenced host plant genome of *M. esculenta*. On average, 4.6% of the raw reads from both healthy leaf and gall tissues did not map to the reference genome. Furthermore, of those unmapped reads, 12.4% were unique to the gall tissue samples. *De novo* assembly of these specific reads produced 17,148 contigs. Eight of these assembled contigs produced fragments with sizes between 2,000 and 2,501 bases, 306 between 1,000 and 2,000 bases, 2,161 contigs of 500–1,000 bases, 14,664 fragments between 200 and 500 bases, and 8 fragments smaller than 200 bases (Supplementary Data 2). Using this approach, reads that did not map to the cassava reference nor were they shared between healthy and gall tissue, represent potential foreign DNA from endophytic organisms unique to galls, such as bacteria or fungi, or possible foreign DNA integrated into the genome of gall cells (see Figure S5 for the general pipeline-flow diagram approach). The resulting assemblies were compared against the reference *M. esculenta* genome to identify possible hybrid/fusion fragments, which must harbour homologous sequences with the host plant, along with external sequences without any homology to the host reference genome. When the assembled contigs were filtered once again following this approach, 407 contigs were retained, 59 of which had 300–571 bases, and 348 had 229–300 bases (Supplementary Data 3). Our analysis found 130 of the reported hybrid/fusion contigs with coverage from 10 to 30 reads. A total of 124 showed coverage ranging from six to nine. Only 74 displayed five to four reads, and 63 had three reads. Moreover, seven contigs showed the largest number of reads, with more than 95 (Supplementary Data 3). These hybrid sequences represent potential insertion regions for foreign DNA integrated into the

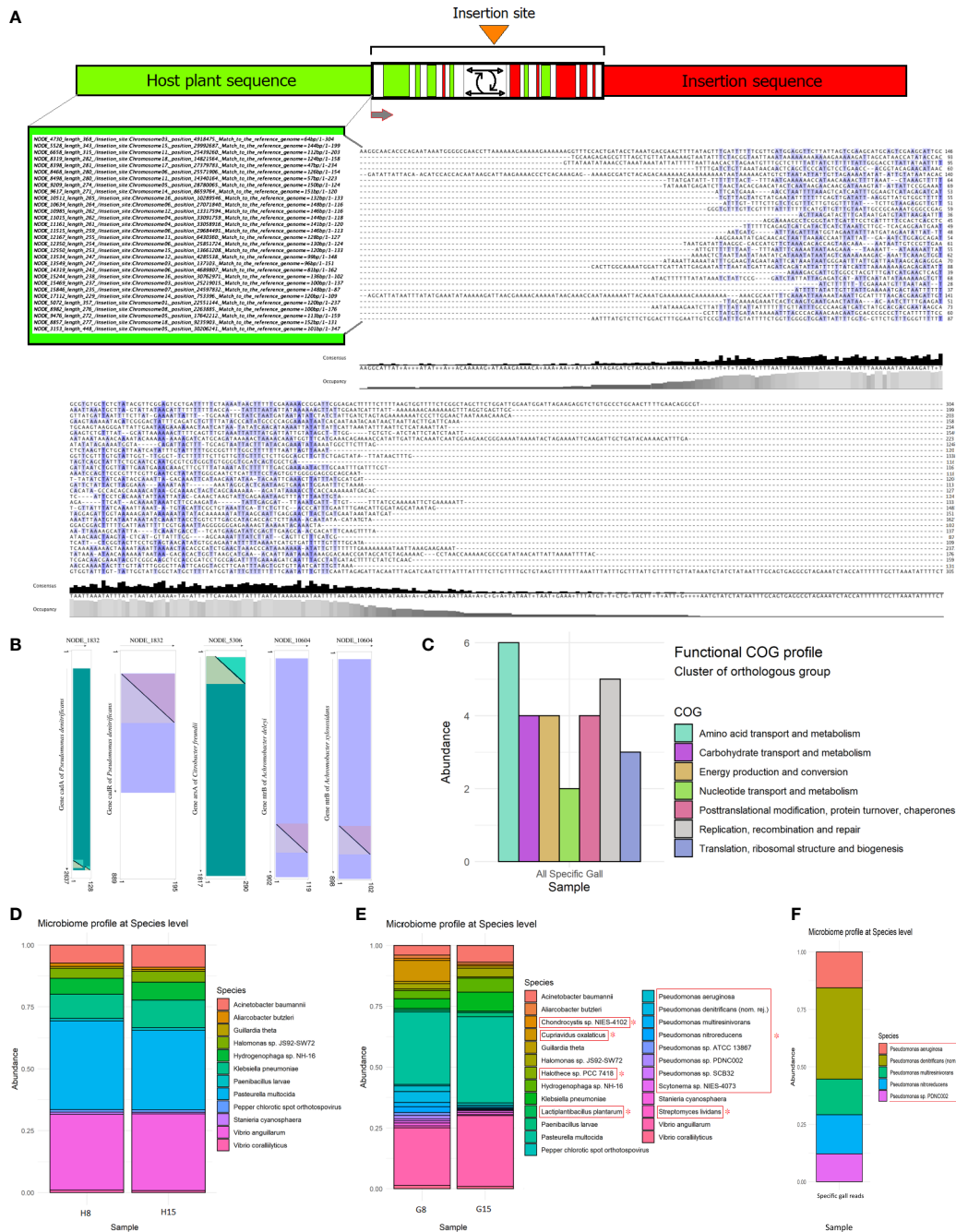


FIGURE 3

Bioinformatic analysis expands the catalogue of gall-specific sequences. **(A)** Theoretical diagram of the insertion regions according to the methodological approach applied. Position of the insertion sites into each host chromosome in the *Manihot esculenta* reference genome is shown on the left in a selected subgroup of hybrid contigs, mostly in forward orientation. Only the alignment framework of unmatched sequences to the cassava reference genome from hybrid/fusion contigs is shown on the right. Grey bar plots show the occupancy within each sequence position, and the black bar shows the base consensus within each sequence position in the resulting alignment scheme. **(B)** Dot plot representation of the aligned and annotated hybrid contigs region of unmatched sequences to the cassava reference genome, showing candidate insertion sequences in several of the hybrid contigs assembled, harbouring known DNA sequences revealing partial significant identity matches and covering with reported genes. Overlapping regions between the contigs and the annotated genes are shown as coloured triangles in each of the represented axes. Purple colour in the referenced gene represents forward strain orientation, and the blue-green colour (viridian) represents reverse strain orientation. **(C)** COG function classification histogram. Count of genes belonging to the COG categories related to exclusive gall reads involved in essential metabolic pathways and biological functions. **(D, E)** Microbiome profile of healthy plant tissue and gall tissue samples. Relative abundance of microorganism taxa identified in the microbiome of healthy leaves and gall samples of cassava. Taxonomic profiles were carried out to a 10K filter using raw reads generated by the shotgun sequencing approach. Each bar represents the organism taxon detected in one sample. The profile showed a similar abundance between both healthy samples **(D)**, but a different relative abundance of microorganisms between gall samples. **(E)** Asterisks indicate enriched or exclusive microorganism species present only in gall tissue according to the taxonomic profile carried out, comparing sequenced healthy leaf samples with gall samples. The most common core endophyte taxa between leaf and gall tissues are also shown. **(F)** Taxonomic identity profile (10K) associated with some of the selected gall-specific reads. Gall-specific reads were bioinformatically filtered from the sequenced gall samples, which mismatched with the cassava reference genome and filtered against shared reads from healthy tissue samples.

genome of gall cells in the sample of mixed galls used for library preparation. Thirty hybrid contigs showing the highest alignment parameters when compared with the cassava reference genome, mainly mapped to the forward strand of this reference genome, are shown in Figure 3A. The estimated insertion position within the assigned chromosomes in the host *M. esculenta* genome is also

shown, evidencing multiple possible insertion sites in the genome of plant cells (see Supplementary Data 3 for all candidate hybrid assemblies).

Structural variations, such as repeated sequences, including retrotransposons (Bartlett et al., 2014), inversions, in tandem sequences (Gang et al., 2019), deletions, duplications, and other

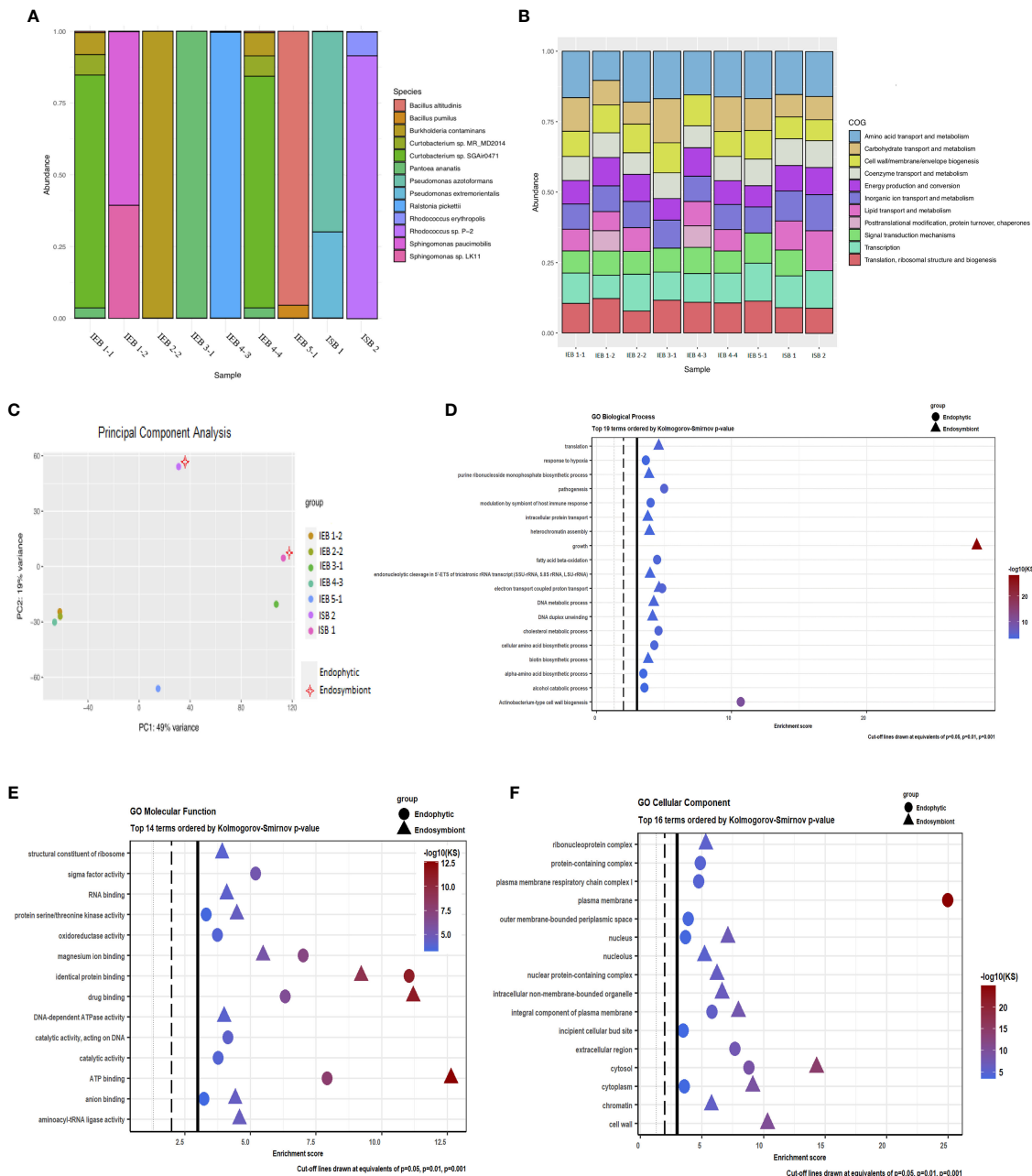


FIGURE 4

Microbiome taxonomic profile and functional analysis of gall endophytes and putative endosymbionts of the inducing insect. (A) Relative abundance of bacterial taxa identified in each of the samples of colony-forming units (CFUs) obtained from gall tissue of *Manihot esculenta* and from the larva head of the inducing insect *latrophobia brasiliensis*. The figure displays the most abundant taxa individually, with the remainder grouped together. Each bar represents the bacterial taxa detected in one sample. Each CFU from the insect head was called an isolated symbiotic bacteria (ISB). Each of the CFUs isolated and grown from internal sections of sterilised gall epidermis tissue was called an isolated endophytic bacteria (IEB).

(B) Distribution of COG functional categories for CFU isolates from gall tissue and from the larval head of the inducing insect. (C) Genetic identity comparison of sequenced bacterial isolates classified up to genus by principal component analysis. (D–F) Functional analysis showing the top Gene Ontology enriched pathways of endophytic and alleged endosymbiont bacteria genomes. A high score indicates a high degree of enrichment.

complex rearrangements, including vector backbone or chromosome sequences carried over together with insertion sequences, may occur around or near the insertion point regions bordering the exogenous DNA and in the host plant DNA. A similar phenomenon has been reported in the flanking regions of the T-DNA insertion fragment harboured in the Ti plasmid of *A. tumefaciens* (Krizkova and Hrouda, 1998; Brunaud et al., 2002; Bartlett et al., 2014; Kleinboelting et al., 2015). They could also occur due to the activity of transposable elements (Wicker et al., 2016), such as *Mariner*-like elements in the genomes of seven species of *Rhus* gall aphids (Ahmad et al., 2021). The integration sites within plant genomes seem to be largely randomly distributed under non-selective conditions (Kim and Gelvin, 2007). Small areas of microhomology in insertion sites between T-DNA and neighbouring plant genomic DNA have also been reported (Brunaud et al., 2002; Kleinboelting et al., 2015). The analyses of microhomology indicate that this type of sequence could be, most of the time, a prerequisite for integration events (Gorbunova and Levy, 1999). Furthermore, a depth analysis of the host plant flanking sequences revealed a high proportion of the characterised T-DNAs inserted into or close to repetitive elements in transgenic barley lines without causing negative effects on transgene expression (Bartlett et al., 2014).

Because each insertion process is a single event that potentially generates a structurally hypervariable region in the DNA around the specific insertion site, finding a consensus sequence or motif around these sites is a challenge (Figure 3A). Those regions containing hypervariable sequences around the integration sites could certainly limit the assembly of reads towards the exogenous DNA sequence, thus restricting the length in hybrid/fusion contigs reported in this survey. Likewise, the extension of read assemblies towards the plant genome orientation in the hybrid contigs is also restricted because all common reads between the sequenced genomes of healthy plants and galls were filtered out, primarily with the reference cassava genome and then by pairing against themselves. The alignment quality with regard to the reference cassava genome, filtering stringency, and the read coverage obtained in our analysis provide evidence in support of the hypothesis that genetic material is inserted into the genomes of plant gall cells (Figure 3A, Supplementary Data 3). Technical artifacts, such as sequencing errors, formation of chimeric DNA during library preparation, and nonspecific assemblies, are unlikely to account for all our observations of putative hybrid/fusion fragments, nor is the presence of hypothetical orthologues of genes from an endophytic microorganism in the plant genome.

BLAST and annotation analysis of the unmatched sequences in some hybrid contigs not associated with the *M. esculenta* reference genome did not show a high identity frequency with reported DNA sequences, while others showed low identity and coverage. However, following our methodological approach, we located known candidate insertion sequences in several assembled hybrid contigs. Some of these candidates harbour DNA sequences revealing partially significant identity matches and cover with reported genes (Figure 3B, Supplementary Data 4). Among the outstanding associated genes, transcription regulatory factor CadR and CadA (Cd<sup>2+</sup> transporting

ATPase enzyme) were found in the hybrid contig NODE\_1832\_length\_571\_cov\_2.279352\_0 (E-value = 8E-89, identity = 96%, 194 bp of length, accession number CP043626 and E-value = 6E-33, identity = 81%, 127 bp of length, accession CP043626, respectively) (Supplementary Data 4, Figure 3B), located at the distal end to the possible insertion site in the region not associated with the cassava reference genome, which showed identity with *Pseudomonas* species, such as *Pseudomonas nitroreducens* strain HBP1, *Pseudomonas denitrificans* strain BG1, and *Pseudomonas multiresinovorans* strain populi, which will be reported later as exclusive or enriched components of the gall microbiome (Figures 3D, E). Likewise, significant putative insertion sequences associated with the *arsA* gene (arsenical pump ATPase), NODE\_5306\_length\_349\_cov\_1.058824\_1 (E-value = 9E-112, identity = 100%, 215 bp of length, accession number CP070545), GTP-binding protein encoded by the *obg* gene, NODE\_9526\_length\_272\_cov\_0.441026\_16 (E-value = 1E-12, identity = 76%, 107 bp of length, accession AP018162), nitrate transport permease protein encoded by the *nrtB* gene, NODE\_10604\_length\_264\_cov\_0.786096\_0 (E-value = 4E-46, identity = 92%, 118 bp of length, accession CP065997), assimilatory nitrite reductase enzyme encoded by the *nasE* gene involved in biological nitrate assimilation, NODE\_10604\_length\_264\_cov\_0.786096\_0 (E-value = 2E-44, identity = 91%, 118 bp of length, accession LT976871), and bicarbonate transport system permease protein encoded by the *cmpB* gene, NODE\_10604\_length\_264\_cov\_0.786096\_0 (E-value = 3E-35, identity = 85%, 118 bp of length, accession LR594671) were also found (Figure 3B, Supplementary Data 4). Interestingly, despite its similarity to reported genes in bacteria, all of them are essential for rapid plant growth, maintenance, and survival under unfavourable conditions and are also associated with gene regulation mechanisms (Supplementary Data 4).

Additionally, the association among the 17,148 gall-specific contigs as potential components of an integrated DNA fragment in the gall cells could be inferred based on their annotated function. However, for many hypothetical proteins, a function could not be assigned (Supplementary Data 5). The presence of different transcriptional, post-translational, and cell cycle regulatory factors, as well as exogenous polymerases, transposases (transposase A), or the integration host factor subunit alpha of bacteriophage lambda (which plays a crucial role in the insertion process of lambda DNA into the *Escherichia coli* chromosome), among others, within the gall-specific contigs, provide indirect evidence and are particularly revealing.

Functional Cluster of Orthologous Group (COG) profile analysis, using gall exclusive contigs, showed that the COG categories of amino acid transport and metabolism, carbohydrate transport and metabolism, energy production and conversion, nucleotide transport and metabolism, posttranslational modification, the translation of the ribosomal structure and biogenesis, and DNA replication, recombination, and repair were the main functional categories represented. Although a relatively small number of contigs were used due to bioinformatic processing, these results could indicate that functions related to growth, transport of plant metabolites to inner gall tissues, replication, and expression of nucleic acids, in addition to gene regulation at

different levels, might be promoted or increased by the exogenous DNA detected in gall cells (Figure 3C).

## Metagenomes reveal an enriched microbial community in galls

We used HTS to analyse the metagenomes of two gall samples, each consisting of a pool of galls and two healthy tissue samples. The average percentage of assigned reads to any species-level taxon for healthy tissue samples was 4.64%, whereas 4.77% was the average for gall samples, according to reference databases. We showed only the significant microorganisms found at a filter resolution of 10K. The resulting taxonomic profile showed a common microbiome between gall and healthy tissue samples, including a core community dominated by 12 taxa, 9 of which were bacteria, namely, *Acinetobacter baumannii*, *Aliarcobacter butzleri*, *Halomonas* sp. JS92–SW72, *Hydrogenophaga* sp. NH–16, *Klebsiella pneumoniae*, *Paenibacillus larvae*, *Pasteurella multocida*, *Vibrio anguillarum*, and *Vibrio coralliilyticus*, in addition to two cyanobacteria, *Guillardia theta* and *Stanieria cyanosphaera*, and the Pepper chlorotic spot orthotospovirus (Figures 3D, E). Furthermore, an exclusive or enriched microbial community was detected in samples of gall tissues. The taxonomic composition included a community dominated by eight species belonging to the genus *Pseudomonas*, two species of cyanobacteria (*Chondrocystis* sp. and *Halotheca* sp. PCC 7418), and three other bacteria species, *Cupriavidus oxalaticus*, *Lactiplantibacillus plantarum*, and *Streptomyces lividans* (filamentous bacterium) (Figure 3E). A brief overview of the relevant functional characteristics reported for these microorganisms and their possible role in gall development and maintenance can be found in the [Supplementary Text section](#).

From the sequencing data exclusive to gall tissue, up to 69.2% of the reads were unassigned to any taxon. These may correspond to microorganisms not included in the available databases at the time of our analysis, or even indicate the possibility of new species as part of this unique microbial community. Nevertheless, a significant amount of exclusive gall reads was associated with several of the same *Pseudomonas* species reported as exclusive or enriched by metagenomic analysis in gall tissue (Figures 3E, F).

## Isolation and genomic analysis of putative gall-inducing symbionts and gall endophytes

Two isolates (CFUs) from the original culture grown from the larval head of the inducing insect, *I. brasiliensis*, as well as seven endophytic bacteria isolates selected from cassava gall tissue, were sequenced using high-throughput platforms. Taxonomic profile analysis applied to all isolates revealed that four corresponded to a single species (IEB 2-2 = *Burkholderia contaminans*, IEB 3-1 = *P. ananatis*, IEB 4-3 = *Ralstonia pickettii*, and IEB 5-1 = *Bacillus altitudinis*). Moreover, three isolates showed association with two bacteria species of the same genus (IEB 1-2 = *Sphingomonas* sp. LK11 and *Sphingomonas paucimobilis*), and the two isolates from

the original culture grown from the inducer insect, isolate ISB 1, which was *Pseudomonas azotoformans* and *Pseudomonas extremorientalis*, and isolate ISB 2, which was *Rhodococcus* sp. P-2 (which could not be classified to the species level, major component) and *R. erythropolis*. Likewise, isolates IEB 1-1 and IEB 4-4 were assigned to the same bacteria species: *Burkholderia contaminans*, *Curtobacterium* sp. MR\_MD2014, *Curtobacterium* sp. SGAir0471, and *P. ananatis* (Figure 4A). However, none of the bacteria isolated from either the gall or the inducing insect larva were precisely detected in the resulting taxonomic profiles belonging to the sequenced samples of healthy plants and gall tissues (Figures 3D, E). Nevertheless, through synteny comparison analysis using specific gall contigs as target sequences compared to each of the sequenced bacterial genomes, the endophytic condition was determined only for the bacteria species isolated from gall tissue, not for those from the inducing insect (Figure S6). Thus, these endophytes might be a marginal component of the gall microbiome, but many of them could also be exclusive components of the microbial community within gall tissues. None of the gall-specific fragment consensus sequences (obtained experimentally) shared homology or aligned with any of the bacterial genomes. The isolation, identification, and subsequent characterisation of microbiome components are difficult tasks due to their low abundance in tissue. This is especially true when using whole-host genome sequencing techniques. However, this barrier may often be overcome using different culture media for growth.

Interestingly, Yang et al. (2021) reported the same bacterial genera sequenced in this study (both gall and inducing insect), except for *R. pickettii*, in their bacterial community of *Lithosaphonecrus arcoverticus* (Hymenoptera: Cynipidae) and in their gall (galled twigs) in *Lithocarpus glaber* (Fagaceae). *Pseudomonas* and *Sphingomonas* have also been identified as members of a common core bacterial community in willow-galling sawflies (Michell and Nyman, 2021). Moreover, *R. erythropolis*, *P. ananatis*, and *Pseudomonas* spp. were recently identified from the bacterial communities by Yang et al. (2022) as predominant species in chestnut tree galls induced by *Dryocosmus kuriphilus*. These bacteria were also identified as components of the microbial communities of the inducing insect and in *Torymus sinensis*, a host-specific parasitoid of *D. kuriphilus*.

Microbial genome comparison by COG functional categories is shown in the COG profile for each isolated CFU (Figure 4B). Also, genetic identity comparison by principal component analysis of sequenced bacterial isolates is shown in Figure 4C. Gene Ontology enrichment analysis of endophytic bacteria genomes showed that the most relevant enriched terms were those related to *Actinobacterium*-type cell wall biogenesis in the category of biological processes and plasma membrane in the category of cellular components. Of the 14 most significant molecular functions identified, sigma factor activity, magnesium ion binding, identical protein binding, drug binding, and ATP binding were the most significantly enriched. Moreover, growth factors for biological processes, in addition to the nucleus, intracellular non-membrane-bounded organelle, nuclear protein-containing complex, integral components of the plasma membrane, cytosol, cytoplasm, and cell wall of GO cellular component, were significantly increased in putative larval endosymbiotic bacteria.

Similarly, magnesium ion binding, identical protein binding, drug binding, and ATP binding were the most significant molecular functions in the enrichment values for these bacteria (Figures 4D–F). Significant discrete GO terms from endophytic bacteria and possible insect endosymbiotic bacteria are shown in Supplementary Data 6. Thus, functional analyses carried out on the endophytic bacteria and on the presumed endosymbiotic bacteria isolated and sequenced from the gall tissue and insect head, respectively, as well as functions associated with the exclusive or enriched microbial community identified in cassava galls by metagenomic analysis (Figure 3E), suggest that the microbial community may play an important role in gall induction, growth, and maintenance. The inferred metabolic pathways and biological functions show that several endophytic bacterial species have a range of different potential functions, including biodegradation of phenolic and potentially harmful metabolic compounds, nutrient supplementation, synthesis of plant hormones, and secondary metabolite degradation. Moreover, this endophytic bacterial community has the potential to synthesise essential amino acids and vitamins, and some of them could be involved in nitrogen and phosphate metabolism, suggesting that these bacterial species could contribute significantly to the nutritional quality of gall tissue. Hence, from an ecological, evolutionary, and functional point of view, our data show that some components of this microbiome can play important roles, both in the host plant itself and in galls. Furthermore, we argue that an unrevealed, induced microbial community that seems to be evident in our findings might have a critical impact on insect gall induction and maintenance despite their low abundance in the gall tissue.

## Artificial gall induction by potential bacterial symbionts

An initial gall-like structure was obtained using the *Rhodococcus* bacterial strain, which was related to the gall induction process (Figures 5A, I–P). Putative insect endosymbiotic bacterial lines isolated from the inducing insect (larva head), initially classified as *Rhodococcus* sp. by sequencing the 16S gene and then identified as the genus *Rhodococcus* based on the taxonomic profile (related to *Rhodococcus* sp. P-2 and *R. erythropolis*), were used in these assays. The endophytic bacterium *P. ananatis* isolated from gall tissue was used as a control bacterium.

Plant tissue culture inoculation assays on *M. esculenta* were carried out under controlled laboratory conditions. Most of the gall-like structures were formed in leaves and leaf sections of *M. esculenta* rather than in micro-stakes with apical shoots (apical buds with small leaf primordia), where their formation was the lowest in the first 2 weeks. However, a greater increase in tissue necrosis was observed after 2 weeks of culture (Figure 5A). Moreover, since the culture medium used was not optimised for the propagation and maintenance of leaves, it was not possible to maintain the structure for more than 4–5 weeks. Necrosis and early fall of the inoculated leaves did not allow long-term monitoring of

*in vitro* materials (with a few exceptions), an effect also observed in inoculated greenhouse plants.

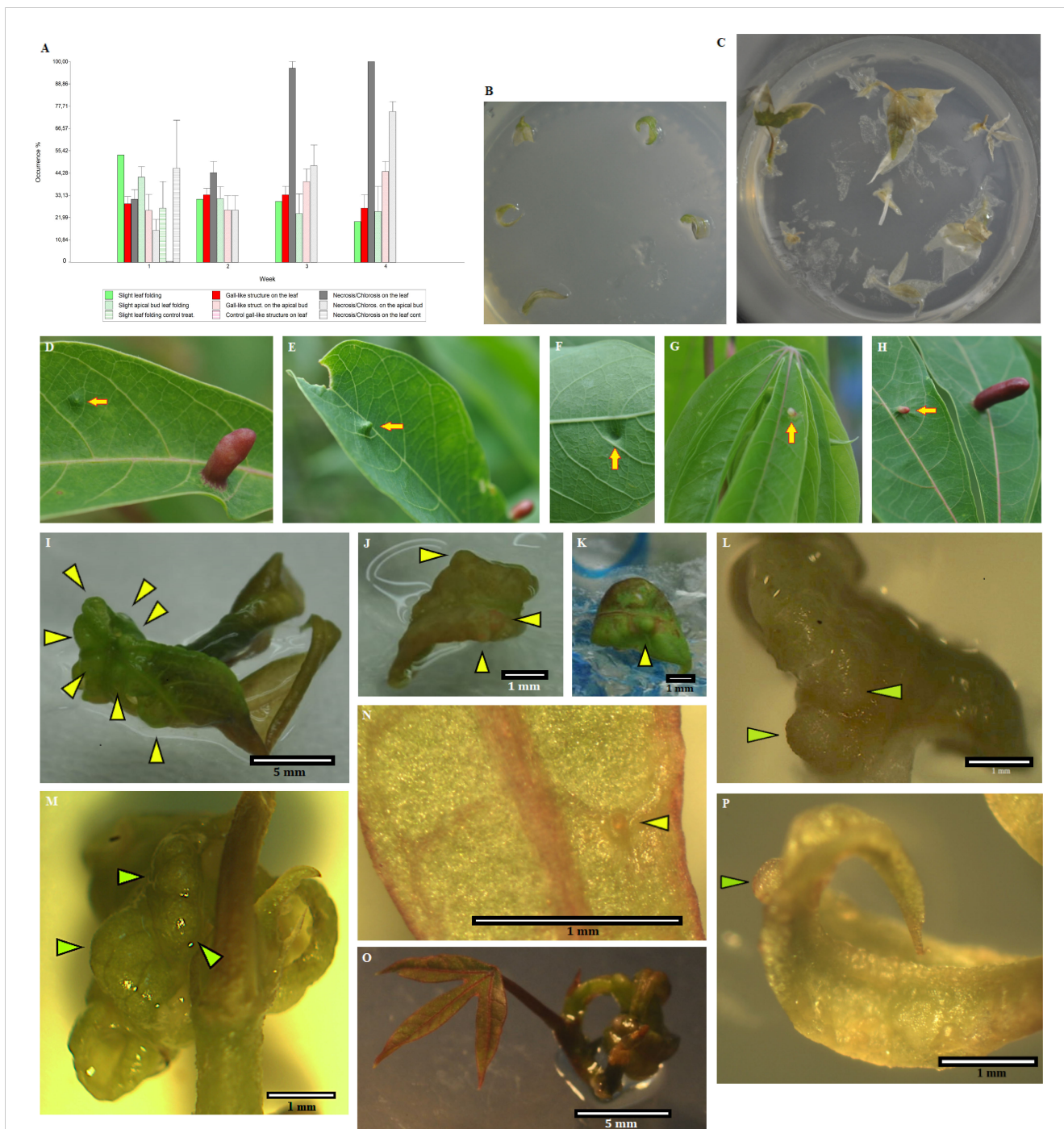
The inoculation of plants grown in soil under greenhouse conditions was also attempted. However, except for the rare event of neoplasm formation over the inoculated leaf primordia, no gall-like structures were clearly observed (results not shown). This could be the consequence of an increased level of control by the plant in its growth process or unaccounted for environmental factors. Moreover, strong tissue chlorosis was observed after the first week, when the bacteria *P. ananatis* was used in the assay (control inoculation bacterium), observing broad bacterial growth on the tissues without tissue folding or gall-like structures (Figure 5C). Likewise, no significant tissue folding or gall-like structure was observed in the control material treated only with mechanical abrasion using the ground glass (Figure 5B). The different success rates of inducing gall-like structures in the *in vitro* assays and under controlled greenhouse conditions indicate that specific inoculation conditions have yet to be optimised.

After inoculation and incubation with cassava explants, bacterial colonies on plant tissue were frequently observed and were usually associated with invaginations and folds of the leaf blade. However, these leaf deformations and invaginations were not frequently associated with macroscopically visible bacterial growth (Figures 5I, K, L, N, P). Furthermore, a few showed differentiated neoplastic tissue that visually differed from the surrounding tissue (Figures 5I, N, P). An increased tissue reaction was observed when a high bacterial density grew over the plant tissues. This increased reaction induced a higher rate of deformation and invagination of the leaf blade, which usually triggered an earlier necrosis condition in the explants (Figure 5A, images not shown), probably generated by the hypersensitive reaction of the plant tissue. Initial gall formation induced under natural field conditions on young leaf primordia and on medium-mature young leaves is shown in Figures 5D–H.

The ability to induce abnormal gall-like growth has also been reported in *Rhodococcus fascians* (Stes et al., 2011; Dolzblasz et al., 2018; Harris and Pitzschke, 2020). The development of plant growth is closely related to a linear virulence plasmid harbouring an array of cytokinin genes encoded by the fasciation (*fas*) operon in most pathogenic isolates (Francis et al., 2012; Creason et al., 2014; Radhika et al., 2015; Jameson et al., 2019). Furthermore, despite the absence of authentic leafy galls in Pistachio Bushy Top Syndrome (PBTS), synergistic coinfection has been reported between *Rhodococcus corynebacterioides* and *R. fascians* by Vereecke et al. (2020).

Bacterial strains belonging to the genus *Rhodococcus* (Figure 4A), isolated from the inducing insect (ISB 2 isolate) and used in gall induction assays, were not detected with certainty at the species level in the sequenced samples of galls and healthy plant tissues when metagenomic analysis was applied (Figures 3D–F). Therefore, we suggest three possible explanations for our findings. The first scenario proposes that the bacterial lines used in our gall induction assays were not conclusively detected because the bacteria had a lower relative abundance within the gall tissue. Thus, the target bacterial DNA would not have a representative fraction in the gall samples purified





for sequencing. An additional explanation could be that the putative bacterium is involved in the gall induction process, but it is not a component of the gall cell endophytic microbiome, according to our results, thus performing its action externally to the plant cell.

Furthermore, the bacterium could exert its action at the initial stage of gall induction. Third, another taxonomically related bacterium sharing a similar molecular mechanism might also induce the formation of this structure in its initial stage.

## Detection of the specific gall marker fragment in sequenced plant material and sequenced bacterial isolates

The non-detection of the differentially amplified gall fragment (specific gall fragment experimentally obtained) in the sequenced *M. esculenta* gall genome and specifically in the sequenced *Rhodococcus* spp., which we argue is related to the gall induction process, could be explained by the sequencing of large low-copy-number wild-type plasmids harbouring large duplications, making it nearly impossible to correctly determine a plasmid genome sequence using a short-read sequencing platform, such as Illumina HiSeq and MiSeq (Smalla et al., 2015; Orlek et al., 2017). Moreover, large wild-type plasmids are difficult to reconstruct from whole-genome sequencing data; this arduous task usually requires a hybrid assembly approach that combines the long reads with the accuracy of short-read sequencing (Berbers et al., 2020). Consequently, localising genes in specific plasmids may be difficult (Orlek et al., 2017). Hence, based on our findings, no direct evidence that the specific gall fragment could be part of an insertion sequence was obtained, and we therefore argue otherwise that this specific gall fragment would be an accessory genetic component of the bacteria transformation machinery, harbouring a transformation plasmid, but this would not be part of the insertion sequence integrated into the genome of the plant cell, analogous to the configuration of the Ti plasmid in *A. tumefaciens* strains (Suzuki et al., 2015; Shao et al., 2018; Chou et al., 2022). Therefore, its specific detection in the DNA extracted from gall tissue by PCR-based methodologies could be the consequence of the transferred endophytic form of *Rhodococcus* spp. by the inducing insect, or another bacterium sharing the same genetic element, in which genomic and plasmid DNA are co-precipitated along with host plant DNA. Moreover, *in silico* PCR analysis of the specific gall fragment did not produce any amplicon from the cassava reference genome (results not shown), thus demonstrating that this DNA fragment is not a component of the host plant genome. However, further research on this specific topic should be carried out to provide more evidence.

## Conclusion

We provide evidence suggesting an insect-induced gall formation mechanism mediated by genetic transformation events in host plant cells. Our data allude to a potential mobile genetic element harboured in *Rhodococcus* spp. bacteria isolated from the inducing insect that could be involved in the induction mechanism. Moreover, gall induction and growth could also be associated with a change in the microbiome composition in plant tissue. It is possible that the inducing insect injects components of the gall-specific endophytic community during female oviposition and larval feeding. Genetic transformation and microbiome modification of gall tissue in *M. esculenta* could be a more widely distributed induction mechanism in nature. This was supported by the detection of a potential accessory genetic component of the transformation machinery (ubiquitin-like gene *E2*) or with a

similar identity in other insect galls in other plant species, and even detected in wild-type plasmids purified from endophytic bacteria species isolated from the same cassava gall tissue. While our results provide greater insight into plant–bacteria–insect interactions, further research is needed to fully understand the complex mechanisms of gall induction and formation. Moreover, the insertion of genetic elements from a putative bacterium could function as a switch in the molecular interaction between the inducing insect and the host plant.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

## Ethics statement

The manuscript presents research on animals that do not require ethical approval for their study.

## Author contributions

OG-B: conceived the survey; substantial contribution to the concept and design of the bioinformatic approach; performed most of the molecular biology experiments; contribution to data analysis and interpretation; contribution to the data collection, registration, and processing; obtained and provided financial support; and wrote the article. JJ-M: substantial contribution to the concept, design, and analysis of the bioinformatic approach; participated in writing the manuscript. RW: contribution to the design and analysis of the bioinformatic approach; and critical revision of the manuscript. PH: contribution to the concept and design of the study; contribution to critical revision; and addition of intellectual content. SV-D: contribution to bioinformatic analysis and interpretation; contribution to critical revision and adding intellectual content; and contribution to manuscript preparation. AA-G: contribution to data analysis and interpretation; and contribution to manuscript preparation. AP-T: contribution to critical revision; adding intellectual content; and contribution to manuscript preparation. All authors contributed to the article and approved the submitted version.

## Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article. This research received no external funding. The postgraduate system at University of Costa Rica covered the Article Processing Fee as a Ph.D. student support.

## Acknowledgments

We are grateful to Dr. Pedro León-Azofeifa, Dr. Federico Albertassi, and Dr. Jorge Lobo for their constructive comments on this research, and we would like to kindly thank Dr. Cecilia Díaz for the discussion of ideas, critical revision, and adding intellectual content. We would like to thank Dr. Julin Weing for the larva head microdissection. We appreciate the support and collaboration provided by the staff of the Biotechnology Laboratory (Escuela de Agronomía, Campus San Carlos del Instituto Tecnológico de Costa Rica), as well as Miguel Pacheco for Gram-staining bacteria. We thank the biologist Alonso Quesada at the Herbario Museo Nacional, Costa Rica, and Dr. Mario Blanco Coto at the Herbario Escuela de Biología, UCR, for their help with the taxonomic identification of some host plant materials used in this work. We also appreciate the technical cooperation of Luis Benavides. Finally, we appreciate the initial logistic and material support given by Dr. Pedro León-Azofeifa in his laboratory at Centro de Investigaciones en Biología Celular y Molecular, Universidad de Costa Rica.

## References

- Adams, E., and Spoel, S. H. (2018). The ubiquitin-proteasome system as a transcriptional regulator of plant immunity. *J. Exp. Bot.* 69 (19), 4529–4537. doi: 10.1093/jxb/ery216
- Ahmad, A., Wallau, G. L., and Ren, Z. (2021). Characterization of Mariner transposons in seven species of *Rhus* gall aphids. *Sci. Rep.* 11, 16349. doi: 10.1038/s41598-021-95843-5
- Alexa, A. (2022). *Rahnenfuhrer J. topGO: Enrichment analysis for gene ontology (2.46.0) [Computer software]. Bioconductor version: release (3.14).*
- Ananthkrishnan, T. (1998). Insect gall systems: patterns, processes and adaptive diversity. *Curr. Sci.* 75 (7), 672–676.
- Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31 (2), 166–169. doi: 10.1093/bioinformatics/btu638
- Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data.*
- Azizi-Dargahlou, S., and Pouresmaeil, M. (2023). *Agrobacterium tumefaciens*-mediated plant Transformation: a review. *Mol. Biotechnol.* 1–18. doi: 10.1007/s12033-023-00788-x
- Banfield, M. J. (2015). Perturbation of host ubiquitin systems by plant pathogen/pest effector proteins. *Cell Microbiol.* 17, 18–25. doi: 10.1111/cmi.12385
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19 (5), 455–477. doi: 10.1089/cmb.2012.0021
- Bartlett, J. G., Smedley, M. A., and Harwood, W. A. (2014). Analysis of t-DNA/host-plant DNA junction sequences in single-copy transgenic barley lines. *Biology* 3 (1), 39–55. doi: 10.3390/biology3010039
- Bartlett, L., and Connor, E. F. (2014). Exogenous phytohormones and the induction of plant galls by insects. *Arthropod-Plant Interact.* 8 (4), 339–348. doi: 10.1007/s11829-014-9309-0
- Berbers, B., Saltykova, A., Garcia-Graells, C., Philipp, P., Arella, F., Marchal, K., et al. (2020). Combining short and long read sequencing to characterize antimicrobial resistance genes on plasmids applied to an unauthorized genetically modified *Bacillus*. *Sci. Rep.* 10, 4310. doi: 10.1038/s41598-020-61158-0
- Bredeson, J. V., Lyons, J. B., Prochnik, S. E., Wu, G. A., Ha, C. M., Edsinger-Gonzales, E., et al. (2016). Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nat. Biotechnol.* 34, 562–570. doi: 10.1038/nbt.3535
- Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., et al. (2002). T-DNA integration into the *Arabidopsis* genome depends on sequences of pre-insertion sites. *EMBO Rep.* 3, 1152–1157. doi: 10.1093/embo-reports/kvf237
- Bushnell, B. (2015). *BBMap short-read aligner, and other bioinformatics tools.* (Berkeley: University of California).
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinf.* 15 (10), 421. doi: 10.1186/1471-2105-10-421
- Cambier, S., Ginis, O., Moreau, S. J. M., Gayral, P., Hearn, J., Stone, G. N., et al. (2019). Gall wasp transcriptomes unravel potential effectors involved in molecular dialogues with oak and rose. *Front. Physiol.* 10. doi: 10.3389/fphys.2019.00926
- Campbell, M. A., Van Leuven, J. T., Meister, R. C., Carey, K. M., Simon, C., and McCutcheon, J. P. (2015). Genome expansion via lineage splitting and genome reduction in the cicada endosymbiont *Hodgkinia*. *Proc. Natl. Acad. Sci. United States America* 112 (33), 10192–10199. doi: 10.1073/pnas.1421386112
- Chalupowicz, L., Barash, I., Panijel, M., Sessa, G., and Manulis-Sasson, S. (2009). Regulatory interactions between quorum-sensing, auxin, cytokinin, and the hrp regulon in relation to gall formation and epiphytic fitness of *Pantoea agglomerans* pv. *gypsophilae*. *Mol. Plant-Microbe Interact.* 22, 849–856. doi: 10.1094/MPMI-22-7-0849
- Chou, L., Lin, Y. C., Haryono, M., Santos, M. N., Cho, S. T., Weisberg, A. J., et al. (2022). Modular evolution of secretion systems and virulence plasmids in a bacterial species complex. *BMC Biol.* 20, 16. doi: 10.1186/s12915-021-01221-y
- Coolen, S., Magda, R. D., and Welte, C. U. (2022). The secret life of insect-associated microbes and how they shape insect-plant interactions. *FEMS Microbiol. Ecol.* 98 (9), fiac083. doi: 10.1093/femsec/fiac083
- Creason, A. L., Vandeputte, O. M., Savory, E. A., Davis, E. W. II, Putnam, M. L., Hu, E., et al. (2014). Analysis of genome sequences from plant pathogenic *Rhodococcus* reveals genetic novelties in virulence loci. *PLoS One* 9, e101996. doi: 10.1371/journal.pone.0101996
- de Lillo, E., Pozzebon, A., Valenzano, D., and Duso, C. (2018). An intimate relationship between eriophyid mites and their host plants - A review. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01786
- Dellaporta, S. L., Wood, J., and Hicks, J. B. (1983). A plant DNA miniprep: preparation II. *Plant Mol. Biology Rep.* 1 (14), 19–21. doi: 10.1007/BF02712670
- Desnitskiy, A. G., Chetverikov, P. E., Ivanova, L. A., Kuzmin, I. V., Ozman-Sullivan, S. K., and Sukhareva, S. I. (2023). Molecular aspects of gall formation induced by mites and insects. *Life* 13 (6), 1347. doi: 10.3390/life13061347
- Di Rienzo, J. A., Casanoves, F., Balzarini, M. G., Gonzalez, L., Tablada, M., and Robledo, C. W. (2020). *InfoStat versión* (FCA, Universidad Nacional de Córdoba, Argentina: Centro de Transferencia InfoStat).
- Dolzblasz, A., Banasiak, A., and Vereecke, D. (2018). Neovascularization during leafy gall formation on *Arabidopsis thaliana* upon *Rhodococcus fascians* infection. *Planta* 247 (1), 215–228. doi: 10.1007/s00425-017-2778-5
- El-Sayed, W. S., and Ibrahim, R. A. (2015). Diversity and phylogenetic analysis of endosymbiotic bacteria of the date palm root borer *Oryctes agamemnon* (Coleoptera: Scarabaeidae). *BMC Microbiol.* 15 (1), 88. doi: 10.1186/s12866-015-0422-8

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2023.1237966/full#supplementary-material>

- Espirito-Santo, M. M., and Fernandes, G. W. (2007). How many species of gall-inducing insects are there on earth, and where are they? *Ann. Entomological Soc. America* 100 (2), 95–99. doi: 10.1603/0013-8746(2007)100[95:HMSOGI]2.0.CO;2
- Favery, B., Quentin, M., Jaubert-Possamai, S., and Abad, P. (2016). Gall-forming root-knot nematodes hijack key plant cellular functions to induce multinucleate and hypertrophied feeding cells. *J. Insect Physiol.* 84, 60–69. doi: 10.1016/j.jinsphys.2015.07.013
- Ferreira, B. G., Álvarez, R., Avritzer, S. C., and Isaias, R. M. S. (2017). Revisiting the histological patterns of storage tissues: beyond the limits of gall-inducing taxa. *Botany* 95 (2), 173–184. doi: 10.1139/cjb-2016-018
- Francis, I., De Keyser, A., De Backer, P., Simón-Mateo, S., Kalkus, K., Pertry, I., et al. (2012). pFiD188, the linear virulence plasmid of *Rhodococcus fascians* strain D188. *Mol. Plant-Microbe Interact.* 25, 637–647. doi: 10.1094/MPMI-08-11-0215
- Galambos, N., Compant, S., Wäckers, F., Sessitsch, A., Anfora, G., Mazzoni, V., et al. (2021). Beneficial insects deliver plant growth-promoting bacterial endophytes between tomato plants. *Microorganisms* 9 (6), 1294. doi: 10.3390/microorganisms9061294
- Galperin, M. Y., Kristensen, D. M., Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2019). Microbial Genome analysis: COG approach. *Briefings Bioinf.* 20 (4), 1063–1070. doi: 10.1093/bib/bbx117
- Gang, H., Liu, G., Zhang, M., Zhao, Y., Jiang, J., and Chen, S. (2019). Comprehensive characterization of T-DNA integration induced chromosomal rearrangement in a birch T-DNA mutant. *BMC Genomics* 20, 311. doi: 10.1186/s12864-019-5636-y
- Gätjens-Boniche, O. (2019). The mechanism of plant gall induction by insects: revealing clues, facts, and consequences in a cross-kingdom complex interaction. *Rev. Biología Trop.* 67 (6), 1359–1382. doi: 10.15517/rbt.v67i6.33984
- Gätjens-Boniche, O., Sánchez-Valverde, M., Trejos-Araya, C., Espinoza-Obando, R., Pinto-Tomás, A., and Hanson, P. E. (2021). Plant galls recorded from Guanacaste Conservation Area-Costa Rica as an integrated concept of a biological database. *Biota Neotropica* 21 (3). doi: 10.1590/1676-0611-bn-2020-1153
- Giron, D., Frago, E., Glevarec, G., Pieterse, C. M. J., and Dicke, M. (2013). Cytokinins as key regulators in plant-microbe-insect interactions: connecting plant growth and defense. *Funct. Ecol.* 27, 599–609. doi: 10.1111/1365-2435.12042
- Giron, D., Huguet, E., Stone, G. N., and Body, M. (2016). Insect-induced effects on plants and possible effectors used by galling and leaf-mining insects to manipulate their host-plant. *J. Insect Physiol.* 84, 70–89. doi: 10.1016/j.jinsphys.2015.12.009
- Orbunova, V. V., and Levy, A. A. (1999). How plants make ends meet: DNA double-strand break repair. *Trends Plant Sci.* 4, 263–269. doi: 10.1016/S1360-1385(99)01430-2
- Gutzwiller, F., Dedeine, F., Käiser, W., and Giron, D. (2015). Correlation between the green-island phenotype and *Wolbachia* infections during the evolutionary diversification of Gracillariidae leaf-mining moths. *Ecol. Evol.* 5 (18), 4049–4062. doi: 10.1002/ece3.1580
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.
- Hammer, T. J., and Bowers, M. D. (2015). Gut microbes may facilitate insect herbivory of chemically defended plants. *Oecologia* 179 (1), 1–14. doi: 10.1007/s00442-015-3327-1
- Hansen, A. K., and Moran, N. A. (2014). The impact of microbial symbionts on host plant utilization by herbivorous insects. *Mol. Ecol.* 23 (6), 1473–1496. doi: 10.1111/mec.12421
- Hanshew, A. S., Mason, C. J., Raffa, K. F., and Currie, C. R. (2013). Minimization of chloroplast contamination in 16S rRNA gene pyrosequencing of insect herbivore bacterial communities. *J. Microbiological Methods* 95 (2), 149–155. doi: 10.1016/j.mimet.2013.08.007
- Harris, M. O., and Pitzschke, A. (2020). Plants make galls to accommodate foreigners: some are friends, most are foes. *New Phytol.* 225, 1852–1872. doi: 10.1111/nph.16340
- Hopp, H. E., Spangenberg, G., and Herrera-Estrella, L. (2022). Editorial: plant transformation. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.876671
- Huang, M. Y., Huang, W. D., Chou, H. M., Chen, C. C., Chen, P. J., Chang, Y. T., et al. (2015). Structural, biochemical, and physiological characterization of photosynthesis in leaf-derived cup-shaped galls on *Litsea acuminata*. *BMC Plant Biol.* 15 (1), 61. doi: 10.1186/s12870-015-0446-0
- Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47 (Database issue), D309–D314. doi: 10.1093/nar/gky1085
- ICGMC (International Cassava Genetic Map Consortium) (2015). High-resolution linkage map and chromosome-scale genome assembly for cassava (*Manihot esculenta* Crantz) from 10 populations. *G3 (Bethesda)* 5 (1), 133–144. doi: 10.1534/g3.114.015008
- Isaias, R. M. S., Ferreira, B. G., Alvarenga, D. R., Barbosa, L. R., Salminen, J. P., and Steinbauer, M. J. (2018). Functional compartmentalisation of nutrients and phenolics in the tissues of galls induced by *Leptocybe invasa* (Hymenoptera: Eulophidae) on *Eucalyptus camaldulensis* (Myrtaceae). *Austral. Entomology* 57 (2), 238–246. doi: 10.1111/aen.12336
- Jameson, P. E., Dhandapani, P., Song, J., Zatloukal, M., Strnad, M., Remus-Emsermann, M. N. P., et al. (2019). The cytokinin complex associated with *Rhodococcus fascians*: which compounds are critical for virulence? *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00674
- Janjusevic, R., Abramovitch, R. B., Martin, G. B., and Stebbins, C. E. (2006). A bacterial inhibitor of host programmed cell death defenses is an E3 ubiquitin ligase. *Science* 311, 222–226. doi: 10.1126/science.1120131
- Jump, B. A., and Woodward, S. (1994). Histology of witches' brooms on *Betula pubescens*. *Eur. J. For. Pathol.* 24, 229–237. doi: 10.1111/j.1439-0329.1994.tb00989.x
- Kavipriya, C., Yuvaraja, A., Senthil, K., and Menaka, C. (2019). Genetic transformation methods for crop improvement- A brief review. *Agric. Rev.* 40 (4), 281–288. doi: 10.18805/ag.R-1904
- Kim, S., and Gelvin, S. (2007). Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *Plant J.* 51, 779–791. doi: 10.1111/j.1365-313X.2007.03183.x
- Kleinboelting, N., Huep, G., Appelhagen, I., Viehoveer, P., Li, Y., and Weisshaar, B. (2015). The structural features of thousands of t-DNA insertion sites are consistent with a double-strand break repair-based insertion mechanism. *Mol. Plant* 8 (11), 1651–1664. doi: 10.1016/j.molp.2015.08.011
- Klimov, P. B., Chetverikov, P. E., Dodueva, I. E., Vishnyakov, A. E., Bolton, S. J., Paponova, S. S., et al. (2022). Symbiotic bacteria of the gall-inducing mite *Fragariocoptes setiger* (Eriophyoidea) and phylogenomic resolution of the eriophyoid position among Acari. *Sci. Rep.* 12 (1), 3811. doi: 10.1038/s41598-022-07535-3
- Koressaar, T., and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23 (10), 1289–1291. doi: 10.1093/bioinformatics/btm091
- Korgaonkar, A., Han, C., Lemire, A. L., Siwanowicz, I., Bannoun, D., Kopec, R. E., et al. (2021). A novel family of secreted insect proteins linked to plant gall development. *Curr. Biol.* 31, 1–14. doi: 10.1016/j.cub.2021.01.104
- Koyama, Y., Yao, I., and Akimoto, S. I. (2004). Aphid galls accumulate high concentrations of amino acids: a support for the nutrition hypothesis for gall formation. *Entomol. Exp. Appl.* 113 (1), 35–44. doi: 10.1111/j.0013-8703.2004.00207.x
- Krzkova, L., and Hroudá, M. (1998). Direct repeats of T-DNA integrated in tobacco chromosome: characterization of junction regions. *Plant J.* 16, 673–680. doi: 10.1046/j.1365-313x.1998.00330.x
- Kud, J., Wang, W., Gross, R., Fan, Y., Huang, L., Yuan, Y., et al. (2019). The potato cyst nematode effector RHA1B is a ubiquitin ligase and uses two distinct mechanisms to suppress plant immune signaling. *PLoS Pathog.* 15, e1007720. doi: 10.1371/journal.ppat.1007720
- Lacroix, B., and Citovsky, V. (2015). Nopaline-type Ti plasmid of *Agrobacterium* encodes a VirF-like functional F-box protein. *Sci. Rep.* 5, 16610. doi: 10.1038/srep16610
- Leitch, I. J. (1994). "Induction and development of the bean gall caused by *Pontania proxima*," in *Plant Galls: Organisms, Interactions, Populations. The Systematics Association (Systematics Association Special Vol. 49)*. Ed. M. Williams (London, United Kingdom: Clarendon Press, Oxford), 283–283.
- Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [Preprint] arXiv:1303.3997*.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25 (16), 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, X. Q., Stahl, R., and Brown, G. (1995). Rapid micropreps and minipreps of Ti plasmid and binary vector from *Agrobacterium tumefaciens*. *Transgenic Res.* 4, 349–351. doi: 10.1007/BF01972532
- Lian, Z., Nguyen, C. D., Liu, L., Wang, G., Chen, J., Wang, S., et al. (2022). Application of developmental regulators to improve in planta or *in vitro* transformation in plants. *Plant Biotechnol. J.* 20 (8), 1622–1635. doi: 10.1111/pbi.13837
- Liu, W., Tang, X., Qi, X., Fu, X., Ghimire, S., Ma, R., et al. (2020). The ubiquitin conjugating enzyme: an important ubiquitin transfer platform in ubiquitin-proteasome system. *Int. J. Mol. Sci.* 21 (8), 2894. doi: 10.3390/ijms21082894
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8
- Lyons, J. B., Bredeson, J. V., Mansfeld, B. N., Bauchet, G. J., Berry, J., Boyher, A., et al. (2022). Current status and impending progress for cassava structural genomics. *Plant Mol. Biol.* 109, 177–191. doi: 10.1007/s11103-020-01104-w
- Magori, S., and Citovsky, V. (2012). The role of the ubiquitin-proteasome system in *Agrobacterium tumefaciens*-mediated genetic transformation of plants. *Plant Physiol.* 160 (1), 65–71. doi: 10.1104/pp.112.200949
- Mani, M. S. (1992). "Introduction to cecidology," in *Biology of Insect-Induced Galls*. Eds. J. D. Shorthouse and O. Rohfritsch (New York, USA: Oxford University Press), 1–7.
- Meyer, J. M., and Hoy, M. A. (2008). Removal of fungal contaminants and their DNA from the surface of *Diaphorina citri* (Hemiptera: Psyllidae) prior to a molecular survey of endosymbionts. *Fla. Entomol.* 91 (4), 702–705.
- Michell, C. T., and Nyman, T. (2021). Microbiomes of willow-galling sawflies: effects of host plant, gall type, and phylogeny on community structure and function. *Genome* 64 (6), 615–626. doi: 10.1139/gen-2020-0018

- Montaldo, A. (1977). *Cultivo de Raíces y Tubérculos Tropicales* (Lima, Perú: Instituto Interamericano de Ciencias Agrícolas de la OEA).
- Mueller, G. M., Bills, G. F., Foster, M. S., and Burlington, M. A. (2004). *Biodiversity of Fungi: Inventory and Monitoring Methods* (Elsevier Academic Press). doi: 10.1016/B978-0-12-509551-8.X5000-4
- Nabity, P. D., Haus, M. J., Berenbaum, M. R., and Delucia, E. H. (2013). Leaf-galling *Phylloxera* on grapes reprograms host metabolism and morphology. *Proc. Natl. Acad. Sci.* 110 (41), 16663–16668. doi: 10.1073/pnas.1220219110
- Nogueira, R. M., Costa, E. C., Silva, J. S., and Isaias, R. M. (2018). Structural and histochemical profile of *Lopesia* sp. Rübsaamen 1908 pinnula galls on *Mimosa tenuiflora* (Willd.) Poir. in a Caatinga environment. *Hoehnea* 45 (2), 314–322. doi: 10.1590/2236-8906-80/201
- Nyman, T., and Julkunen, T. (2000). Manipulation of phenolic chemistry of willows by gall-inducing sawflies. *Proc. Natl. Acad. Sci. United States America* 97, 13184–13187. doi: 10.1073/pnas.230294097
- Olmo, R., Cabrera, J., Diaz-Manzano, F. E., Ruiz-Ferrer, V., Barcala, M., Ishida, T., et al. (2020). Root-knot nematodes induce gall formation by recruiting developmental pathways of post-embryonic organogenesis and regeneration to promote transient pluripotency. *New Phytol.* 227 (1), 200–215. doi: 10.1111/nph.16521
- Orlek, A., Stoesser, N., Anjum, M. F., Doumith, M., Ellington, M. J., Peto, T., et al. (2017). Plasmid classification in an era of whole-genome sequencing: application in studies of antibiotic resistance epidemiology. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.00182
- Park, C., Chen, S., Shirsekar, G., Zhou, B., Khang, C. H., Songkumarn, P., et al. (2012). The *Magnaporthe oryzae* effector AvrPiz-t targets the RING E3 ubiquitin ligase APIP6 to suppress pathogen-associated molecular pattern-triggered immunity in rice. *Plant Cell* 24 (11), 4748–4762. doi: 10.1105/tpc.112.105429
- Piñol, M., Palazón, J., Cusidó, R., and Serrano, M. (1996). Effects of Ri T-DNA *Agrobacterium* rhizogenes on growth and Hyoscyamine production in *Datura stramonium* roots cultures. *Botánica Acta* 109, 133–138. doi: 10.1111/j.1438-8677.1996.tb00553.x
- Ponce, G. E., Fuse, M., Chan, A., and Connor, E. F. (2021). The localization of phytohormones within the gall-inducing insect *Eurosta solidaginis* (Diptera: Tephritidae). *Arthropod-Plant Interact.* 15 (3), 375–385. doi: 10.1007/s11829-021-09817-5
- Prijibelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020). Using SPAdes de novo assembler. *Curr. Protoc. Bioinf.* 70 (1), e102. doi: 10.1002/cpbi.102
- Radhika, V., Ueda, N., Tsuboi, Y., Kojima, M., Kikuchi, J., Kudo, T., et al. (2015). Methylated cytokinins from the phytopathogen *Rhodococcus fascians* mimic plant hormone activity. *Plant Physiol.* 169, 1118–1126. doi: 10.1104/pp.15
- Rainey, F. A., Ward-Rainey, N., Kroppenstedt, R. M., and Stackebrandt, E. (1996). The genus *Nocardopsis* represents a phylogenetically coherent taxon and a distinct actinomycete lineage: proposal of *Nocardopsiaceae* fam. nov. *Int. J. Syst. Bacteriol.* 46, 1088–1092. doi: 10.1099/00207713-46-4-1088
- Ramachandran, P., Joshi, B. J., Maupin-Furlow, J. A., and Uthandi, S. (2021). Bacterial effectors mimicking ubiquitin-proteasome pathway tweak plant immunity. *Microbiological Res.* 250, 126810. doi: 10.1016/j.micres.2021.126810
- Raman, A. (2011). Morphogenesis of insect-induced plant galls: fact and questions. *Flora-Morphology, Distribution. Funct. Ecol. Plants* 206 (6), 517–533. doi: 10.1016/j.flora.2010.08.004
- Ratcliffe, N. A., Furtado Pacheco, J. P., Dyson, P., Castro, H. C., Gonzalez, M. S., Azambuja, P., et al. (2022). Overview of paratransgenesis as a strategy to control pathogen transmission by insect vectors. *Parasites Vectors* 15 (1), 112. doi: 10.1186/s13071-021-05132-3
- Rivera Hernández, J. F. (2011). La mosca de las agallas (*Jatrophobia brasiliensis*) en el cultivo de yuca. *Rev. la Universidad la Salle* 56, 277–288.
- Rohfritsch, O., and Shorthouse, J. D. (1982). "Insect galls," in *Molecular Biology of Plant Tumors*. Eds. G. Kahl and J. S. Schell (New York, USA: Academic Press), 131–152.
- Ronquist, F., Nieves-Aldrey, J. L., Buffington, M. L., Liu, Z., Liljebald, J., and Nylander, J. A. (2015). Phylogeny, evolution and classification of gall wasps: the plot thickens. *PLoS One* 10 (5), e0123301. doi: 10.1371/journal.pone.0123301
- Sadanandom, A., Bailey, M., Ewan, R., Lee, J., and Nelis, S. (2013). The ubiquitin-proteasome system: central modifier of plant signaling. *New Phytol.* 196, 13–28. doi: 10.1111/j.1469-8137.2012.04266.x
- Sánchez, E., Quesada, T., and Espinoza, A. M. (2006). Ultrastructure of the wild rice *Oryza grandiglumis* (Gramineae) in Costa Rica. *Rev. Biol. Trop.* 54, 377–385. doi: 10.15517/rbt.v54i2.13878
- Santner, A., and Estelle, M. (2010). The ubiquitin-proteasome system regulates plant hormone signaling. *Plant J.* 61, 1029–1040. doi: 10.1111/j.1365-3113.2010.04112.x
- Schwartz, S., Elmitski, L., Li, M., Weirauch, M., Riemer, C., Smit, A., et al. (2003). MultiPipMaker and supporting tools: alignments and analysis of multiple genomic DNA sequences. *Nucleic Acids Res.* 31 (13), 3518–3524. doi: 10.1093/nar/gkg579
- Schwartz, S., Zhang, Z., Frazer, K. A., Smit, A., Riemer, C., Bouck, J., et al. (2000). PipMaker—a web server for aligning two genomic DNA sequences. *Genome Res.* 10 (4), 577–586. doi: 10.1101/gr.10.4.577
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30 (14), 2068–2069. doi: 10.1093/bioinformatics/btu153
- Shao, S., Zhang, X., van Heusden, G. P. H., and Hooykaas, P. J. J. (2018). Complete sequence of the tumor-inducing plasmid pTiChry5 from the hypervirulent *Agrobacterium tumefaciens* strain Chry5. *Plasmid* 96–97, 1–6. doi: 10.1016/j.plasmid.2018.02.001
- Shorthouse, J. D., and Rohfritsch, O. (1992). *Biology of the Insect-Induced Galls* (New York, USA: Oxford University Press).
- Shu, K., and Yang, W. (2017). E3 ubiquitin ligases: ubiquitous actors in plant development and abiotic stress responses. *Plant Cell Physiol.* 58, 1461–1476. doi: 10.1093/pcp/pcx071
- Singer, A. U., Schulze, S., Skarina, T., Xu, X., Cui, H., Eschen-Lippold, L., et al. (2013). A pathogen type III effector with a novel E3 ubiquitin ligase architecture. *PLoS Pathog.* 9 (1), e1003121. doi: 10.1371/journal.ppat.1003121
- Sinnott, E. W. (1960). *Plant Morphogenesis* (New York: McGraw-Hill).
- Smalla, K., Jechalke, S., and Top, E. M. (2015). Plasmid detection, characterization, and ecology. *Microbiol. Spectr.* 3 (1). doi: 10.1128/microbiolspec.PLAS-0038-2014
- Song, G. Q., Prieto, H., and Orbovic, V. (2019). *Agrobacterium*-mediated transformation of tree fruit crops: Methods, progress, and challenges. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00226
- Stes, E., Vandeputte, O. M., El Jaziri, M., Holsters, M., and Vereecke, D. (2011). A successful bacterial coup d'état: how *Rhodococcus fascians* redirects plant development. *Annu. Rev. Phytopathol.* 49, 69–86. doi: 10.1146/annurev-phyto-072910-095217
- Sugio, A., Dubreuil, G., Giron, D., and Simon, G. C. (2015). Plant-insect interactions under bacterial influence: ecological implications and underlying mechanisms. *J. Exp. Bot.* 66 (2), 467–478. doi: 10.1093/jxb/eru435
- Suzuki, K., Moriguchi, K., and Yamamoto, S. (2015). Horizontal DNA transfer from bacteria to eukaryotes and a lesson from experimental transfers. *Res. Microbiol.* 166 (10), 753–763. doi: 10.1016/j.resmic.2015.08.001
- Swarup, S., De Feyter, R., Brlansky, R. H., and Gabriel, D. W. (1991). A pathogenicity locus from *Xanthomonas citri* enables strains from several pathogens of *X. campestris* to elicit cankerlike lesions on citrus. *Phytopathology* 81, 802–809. doi: 10.1094/Phyto-81-802
- Tooker, J. F., and Helms, A. M. (2014). Phytohormone dynamics associated with gall insects, and their potential role in the evolution of the gall-inducing habit. *J. Chem. Ecol.* 40 (7), 742–753. doi: 10.1007/s10886-014-0457-6
- Untergrasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., et al. (2012). Primer3-new capabilities and interfaces. *Nucleic Acids Res.* 40 (15), e115. doi: 10.1093/nar/gks596
- Vereecke, D., Zhang, Y., Francis, I. M., Lambert, P. Q., Venneman, J., Stamler, R. A., et al. (2020). Functional genomics insights into the pathogenicity, habitat fitness, and mechanisms modifying plant development of *Rhodococcus* sp. PBTS1 and PBTS2. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.00014
- Vierstra, R. D. (2009). The ubiquitin-26S proteasome system at the nexus of plant biology. *Nat. Rev. Mol. Cell Biol.* 10, 385–397. doi: 10.1038/nrm2688
- Wang, W., Feng, B., Xiao, J., Xia, Z., Zhou, X., Li, P., et al. (2014). Cassava genome from a wild ancestor to cultivated varieties. *Nat. Commun.* 5, 5110. doi: 10.1038/ncomms6110
- Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033
- Wicker, T., Yu, Y., Haberer, G., Mayer, K. F., Marri, P. R., Rounsley, S., et al. (2016). DNA transposon activity is associated with increased mutation rates in genes of rice and other grasses. *Nat. Commun.* 7, 12790. doi: 10.1038/ncomms12790
- Wielkopolan, B., and Jakubowska, M. (2021). Beetles as plant pathogen vectors. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.748093
- Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biol.* 20, 257. doi: 10.1186/s13059-019-1891-0
- Xu, F.-Q., and Xue, H.-W. (2019). The ubiquitin-proteasome system in plant responses to environments. *Plant Cell Environ.* 42, 2931–2944. doi: 10.1111/pce.13633
- Yang, X. M., Hui, Y., Zhao, L. Q., Zhu, D. H., Zeng, Y., and Yang, X. H. (2021). Comparison of auxin and cytokinins concentrations, and the structure of bacterial community between host twigs and *Lithosiphonectrus arcoverticus* galls. *Insects* 12 (11), 982. doi: 10.3390/insects12110982
- Yang, X., Hui, Y., Zhu, D., Zeng, Y., Zhao, L., Yang, X., et al. (2022). The diversity of bacteria associated with the invasive gall wasp *Dryocosmus kuriphilus*, its galls and a specialist parasitoid on chestnuts. *Insects* 13 (1), 86. doi: 10.3390/insects13010086
- You, W., Liu, Q., Zou, K., Yu, X., Cui, H., and Ye, Z. (2011). Morphological and molecular differences in two strains of *Ustilago esculenta*. *Curr. Microbiol.* 62 (1), 44–54. doi: 10.1007/s00284-010-9673-7
- Zahner, V., Lucarotti, C. J., and McIntosh, D. (2008). Application of 16S rDNA-DGGE and plate culture to characterization of bacterial communities associated with the sawfly, *Acantholyda erythrocephala* (Hymenoptera, Pamphiliidae). *Curr. Microbiol.* 57 (6), 564–569. doi: 10.1007/s00284-008-9243-4
- Zeidan, M., and Czosnek, H. (1994). Acquisition and transmission of *Agrobacterium* by the whitefly *Bemisia tabaci*. *Mol. Plant-Microbe Interact.* 7 (6), 792–798. doi: 10.1094/mpmi-7-0792
- Zhao, C., Escalante, L. N., Chen, H., Benatti, T. R., Qu, J., Chellappa, S., et al. (2015). A massive expansion of effector genes underlies gall-formation in the wheat pest *Mayetiola destructor*. *Curr. Biol.* 25 (5), 613–620. doi: 10.1016/j.cub.2014.12.057
- Zhao, C., Rispe, C., and Nabity, P. D. (2019). Secretory RING finger proteins function as effectors in a grapevine galling insect. *BMC Genomics* 20, 923. doi: 10.1186/s12864-019-6313-x