Supporting Information

Endara et al. 10.1073/pnas.1707727114

SI Text

Detailed Methods.

Study species. *Inga* is a genus of trees in the legume family (subfamily Mimosoideae) and is found in lowland moist forests through the New World. There are ∼300 described species (70). At Los Amigos and elsewhere, the genus *Inga* constitutes one of the most diverse and abundant genera in any western Amazonian forest. For example, in 25 ha of forest in Amazonian Ecuador, there are >40 Inga species representing 6% of stems (13). At Los Amigos, we collected data on 33 species of *Inga*. We focused our study on understory saplings, a key stage in the life cycle of a tree (71).

Inga is associated with multiple insect herbivore taxa, including Coleoptera, Orthoptera, phloem-feeding Coreidae (Heteroptera), Diptera, Tenthredinoidea (Hymenoptera), Phasmatodea, and Lepidoptera. Lepidoptera cause the most damage to *Inga* leaves (72), and we therefore focus on them here.

To record host associations of lepidopteran herbivores, we visually searched a minimum of 10 young leaf flushes per tree species and collected only those larvae that were found feeding. Insects were collected by hand from the leaves between 2010 and 2011 for a period of 10 mo. All caterpillars were assigned to morphospecies in the field. Because no identification keys to the caterpillars of this region exist, all herbivores were subsequently assigned to molecular operational taxonomic units (MOTUs) on the basis of DNA barcode sequences for the mitochondrial gene, cytochrome oxidase I (*COI*) (see below). We recorded a total of 1,567 individuals in 174 MOTUs from 19 families of Lepidoptera (Fig. S2).

Plant defensive traits. We focused our study on the defenses of expanding leaves because during this ephemeral stage they receive more than 75% of the damage accrued during the lifetime of a leaf (14, 22, 73). Therefore, the defensive traits most relevant for insect herbivores when selecting hosts would be those of young leaves.

Leaf defensive traits were collected from young leaves on 0.5- to 4-m-tall saplings in the shaded understory between 2007 and 2011. We measured multiple traits that capture the entire plant's defensive profile. We recorded the presence or absence of several classes of phenolic compounds, saponins, and metabolites containing primary or secondary amines that have been shown to decrease the growth and survival of herbivores (73, 74). Details on chemical procedures are reported in ref. 3.

We assessed the length and density of trichomes (number of hairs per 2 cm^2) in a minimum of 10 individuals per species. Young leaves are also defended against herbivory by expanding leaves rapidly and delaying the development of the chloroplast (23). Leaf expansion rate was determined as the percent increase in area per day for ∼13 individuals per species. Chlorophyll content of leaves between 30% and 80% of full expansion was estimated using three values from a Minolta SPAD 502DL meter (Spectrum Technologies). For calibration between SPAD units and chlorophyll content, a portable Spectronic 20 (Milton Roy) was first calibrated in the laboratory using expanding leaves. These were extracted with 90% acetone/10% water (vol/vol) containing Na_2CO_3, and centrifuged at 10,000 × *g* at 5 °C. Absorbances were obtained using a narrow-bandpass spectrometer at 647 and 664 nm. Chlorophyll content was determined using the equations of ref. 75. The same leaves were extracted in 95% ethanol containing Na_2CO_3, centrifuged at 10,000 × *g* at 5 °C, and transmittance measured at 663 and 725 nm. Regression analysis gave the following equation for the portable spectrophotometer:

$$\text{Chl } a + \text{Chl } b\left(\text{mg}\cdot\text{m}^{-2}\right) = 152.4 * (\text{A663} - \text{A725})(\text{volume in mL})$$
$$127 \text{ mL} * \left(\text{area in mm}^2\right)128,$$

where Chl*a* + Chl*b* is the total content of chlorophyll *a* and *b*, and A663 and A725 are the absorbance readings for wavelengths of 663 and 725 nm, respectively.

In the field, the SPAD meter was calibrated using expanding leaves. The SPAD meter was used according to the manufacturer's directions. For determination with the Spectronic 20, chlorophyll was extracted in 95% ethanol containing a small amount of $NaHCO_3$, Na_2CO_3, or Na_2HPO_4 and centrifuged at 25 °C in a minicentrifuge (SC1006-R) at 2,000 × *g*. The relationship was nonlinear and the equation to convert SPAD units to chlorophyll in milligrams of chlorophyll *a* and *b* per square meter is as follows:

$$\text{Chl} = \alpha * \text{SPADi}\beta,$$

where Chl is the total content of chlorophyll (*a* and *b*) of the sample *i*, SPAD is the unitless reading from the SPAD 502DL meter, and α (0.0417) and β (0.9524) are the fitted parameters.

We measured the timing and synchrony in leaf production. Because these are two measures of food availability for insect herbivores, they therefore could play important roles in structuring herbivore assemblages (21, 22). To measure these traits, we monitored between 30 and 70 individuals per species for monthly leaf production. To estimate the timing of leaf production, or mean angle, we converted months to angles, from 0° = January to 360° = December at intervals of 30°. The mean angle for a species indicates the average date of peak flushing activity among the individuals. We evaluated the significance of the mean angle using the Rayleigh test (56). We estimated synchrony in leaf production by calculating the coefficient of variation (CV) of plant individuals per species flushing each month.

Inga leaves have extrafloral nectaries that produce nectar and attract protective ants only during the short period of leaf expansion. We determined the identity and the abundance of ants visiting these nectaries (number of ants per nectary) in ∼30 individuals per species. Ants were identified to genus, and in some cases to species based on morphology.

Herbivore phylogenies. Phylogenetic analyses for the most abundant lepidopteran clades: the superfamily Gelechioidea, and the families Riodinidae and Erebidae were inferred using one to three individuals per MOTU (for MOTU assignment, see below), two nuclear loci: elongation factor 1α (*EF-1α*), wingless (*Wg*), and one mitochondrial locus, *COI*. DNA was extracted from legs or, for very small larvae, larger body parts. The remaining parts were preserved as vouchers. We extracted total genomic DNA in 50 μL of extraction buffer containing 5% Chelex 100 resin (Bio-Rad) as described in ref. 76. For *COI*, PCR amplification and DNA sequencing for most of our samples were generated at the Canadian Center for Barcoding using standard barcoding protocols (77, 78). For the nuclear loci and the remaining samples for *COI*, we performed PCR amplification with 1 μL of DNA extract, 0.2 μM of each primer, and 10 μL of Multiplex PCR kit (Qiagen) in a 20-μL reaction volume (Table S5). We used the same pair of primers for both amplification and sequencing. Primer sequences and annealing temperatures are in Table S6. PCR products were purified using a shrimp alkaline phosphatase protocol. Sequencing was performed using ABI BigDye chemistry (Perkin-Elmer Biosystems) on ABI 3730xl capillary sequencer. We sequenced all products in both directions. The sequences were assembled into contigs and manually edited using the program Sequencher version 5.1 (Gene

Codes). The resulting sequences were subsequently aligned using the program MUSCLE (79).

MOTU assignment used COI sequence data and the software package jMOTU (80), with a similarity cutoff of 15 bp (~2.3%). MOTUs identified using Automatic Barcode Gap Discovery (ABGD) (81) were identical. MOTUs were allocated to taxonomic families by BLASTing each consensus sequence against the NCBI BLAST web interface, with a minimum accepted similarity for family assignment of 90%.

To recognize misidentified taxa and/or confirm the correct placement of MOTUs into families by BLAST, we carried out a number of trials with varying taxon composition (including all of the MOTUs regardless of their taxonomic family) using maximum-likelihood (ML) methods. The analyses used a GTR + Γ model of substitution, and the data were partitioned by genes. ML analyses were implemented using RAxML (62) at the CIPRES Web portal (82), and support nodes were evaluated with 1,000 bootstrap replicates of the data. This preliminary phylogeny provided molecular support for the placement of MOTUs into the families Erebidae and Riodinidae and the superfamily Gelechioidea. Arctiinae and Erebinae were recovered as subfamilies of Erebidae, with Noctuidae as the sister group. These results are in agreement with the most recent published phylogeny for Erebidae (38). For Riodinidae, our preliminary phylogeny conforms to the most recent classification (37), with Riodinidae being monophyletic and with a sister group relationship between Riodinidae and Lycaenidae. The genera found on *Inga* include *Melese*, *Hypocrita*, *Pelochyta*, and *Areva* in the Arctiinae; *Coenipeta*, *Helia*, and *Letis* for Erebinae; *Nymphidium*, *Sarota*, and *Synargis* for Riodinidae; and *Iaspis*, *Ostrinotes*, *Techlopsis*, *Theritas*, *Strephonota*, and *Symbiopsis* for Lycaenidae. The superfamily Gelechioidea was recovered as monophyletic, which is in congruence with published phylogenies for this group (41). Nevertheless, several MOTUs allocated to this group by BLASTing were removed from the eventual analyses because their placement in this preliminary phylogeny was doubtful, being placed within other, distantly related (nongelechiid clades). These probably belong to the family Tortricidae and the superfamily Pyraloidea, groups that at a larval stage look identical to Gelechioidea. Because the Gelechioidea are poorly known, these MOTUs could not be placed to genus by BLAST.

Phylogenetic relationships for MOTUs correctly allocated to Gelechioidea, Riodinidae, and Erebidae were inferred using multilocus coalescent-based Bayesian species tree in *BEAST 2.2.0 (57), with substitution and codon partition models for each marker set according to the suggestions of PartitionFinder 1.1.0 (58) (Table S7). Tree species priors and sequences used for tree rooting for each clade are specified in Table S7. Alternative models were assessed using Bayes factors, following the guidelines from ref. 83. Parameters were estimated from three independent runs of 100 million generations combined using LogCombiner 1.8 (59) with a burn-in of 10 million generations and sampling every 10,000 generations in each run. BEAST model convergence was confirmed by examination of parameter estimate distributions in Tracer 1.6 (84).

Plant phylogeny. Phylogenetic relationships between *Inga* host species were inferred using seven chloroplast regions (*rpoCI*, *psbA-trnH*, *rps16*, *trnL-F*, *trnD-T*, *ndhF-rpl32*, *rpl32-trnL*) and the nuclear ribosomal internal transcribed spacer regions (ITS). DNA extractions used a modified CTAB protocol (85) or DNAeasy plant mini kits (Qiagen). PCR and sequencing protocols for chloroplast regions are given by ref. 3 and for ITS by refs. 60 and 61. Sequences were assembled using Sequencher, version 4.5 (Gene Codes Corporation), and aligned manually, which was unproblematic given low sequence divergence. Sequences were aligned using MAFFT, version 7.0 (86), and phylogenies were estimated in an ML framework using RAxML, with separate models for ITS and cpDNA (62). Phylogenies were subsequently time-calibrated using penalized likelihood (63), where the crown age was constrained to 6 My (following refs. 60 and 64).

From the resulting tree, we extracted pairwise distances between *Inga* species using the function "cophenetic" in the APE package (87) from the statistical programming language R, version 3.2.5 (R Development Core Team 2016). This phylogenetic distance matrix was used in all of the subsequent ecological analyses that involved the phylogeny of *Inga*.

Statistical Analyses.

Relationship between plant traits and phylogenetic signal. Associations between defense traits were investigated by using matrix correlation analyses and phylogenetic generalized linear models (65). To estimate the relationship between continuous and noncontinuous traits (e.g., of noncontinuous traits: presence or absence of chemical compounds and ant visitor community to extrafloral nectaries), we calculated the distances between pairs of *Inga* species for each trait and examined their correlations using partial Mantel tests, controlling for phylogenetic correlations. The distance matrices for biotic (number of ants), developmental, physical, and chemical defenses, as well as synchrony in leaf production were calculated using the Manhattan dissimilarity index. For the ant visitor community, the Bray–Curtis index was used. Because the timing in leaf production is a circular variable (mean angle), we used the angular separation method from the package circular (66) to calculate the distance matrix for this trait.

We also performed a phylogenetic PCA (PPCA) on continuous trait data to derive evolutionary independent axes of defense variation (67). Phylogenetic signal was evaluated on the significant axes of defense variation from the PPCA and on the principal coordinates of the chemistry and ant species distance matrices by using Blomberg's $\bar{K}$ (68). If there is no phylogenetic signal, $\bar{K}$ would be close to zero, whereas values approaching 1 would indicate that the trait value matches expectations under a Brownian model of evolution.

In a previous study and in this study, we showed chemistry and ant visitation to leaves to be divergent among close relatives in *Inga* (3). In contrast, developmental defenses show phylogenetic signal in the previous study but not in this study. Two potential reasons for these differences are that (*i*) we include more species of *Inga* and (*ii*) the present phylogeny resolution (the ITS marker is new, and this resolves several clades).

Constraints on host plant selection. To examine whether differences in total herbivore community structure were related to differences in phylogenetic relationships and/or defensive traits between pairs of *Inga* hosts, we used Mantel and partial Mantel tests with 9,999 permutations. All of the feeding records that were limited to a single individual in a particular host were not included in these analyses. Overlap in feeding records among hosts was estimated using the Bray–Curtis dissimilarity index with relative abundance raw data. The resulting matrix was then compared with a phylogenetic distance matrix for *Inga* hosts and with a defense distance matrix conditioned on the phylogenetic pairwise distances between species. The defense distance matrix between *Inga* species was generated by averaging the distance matrices for the different defense traits such that they were all weighted equally. These analyses were performed in the vegan package (88).

To quantify the extent to which host phylogeny and/or host defenses structure herbivore community and to determine which defense trait is more important, we used redundancy analyses (RDAs). The herbivore community similarity matrix was used as a response variable in a partial distance-based redundancy analysis (dbRDA) together with each one of the measured defensive traits and the principal coordinates of the phylogenetic distance matrix

as explanatory variables. The distance matrix was square-root transformed to avoid negative eigenvalues in the analyses (89). This analysis was performed without timing and synchrony in leaf production because we lacked data on these traits for some species. The analyses were run using sampling effort as a covariate. First, we ran a global test using all of the explanatory variables, and, because this analysis was statistically significant ($P \leq 0.05$, 9,999 permutations), we performed variable selection using the function *ordistep*. We did this to avoid overfitting and to select the most parsimonious model. Adjusted R^2 values for the selected model were computed using the function *varpart*. We also ran a more restricted analysis that included only species for which we had data on phenology of leaf production. This analysis produced very similar results. We performed these analyses in the *vegan* package.

To investigate whether past host shifts have occurred more often on *Inga* that are more similar in defenses or on *Inga* that are more closely related, we examined the congruence of the herbivore phylogenies with *Inga* phylogeny and *Inga* defenses using ParaFit (69) from the APE package. This statistical tool tests the significance of a hypothesis of coevolution between parasites and hosts using phylogenetic distance matrices of associated taxa and a set of host–parasite associations. Distance matrices for herbivores and plants were derived from their phylogenetic trees and from a cluster dendrogram we obtained from the total defense distance matrix using the "cophenetic" function in the APE package. We ran this analysis with 9,999 permutations. We perform the above analyses for the entire Lepidoptera community and for the three major clades associated with *Inga* separately: Gelechioidea, Erebidae, and Riodinidae.

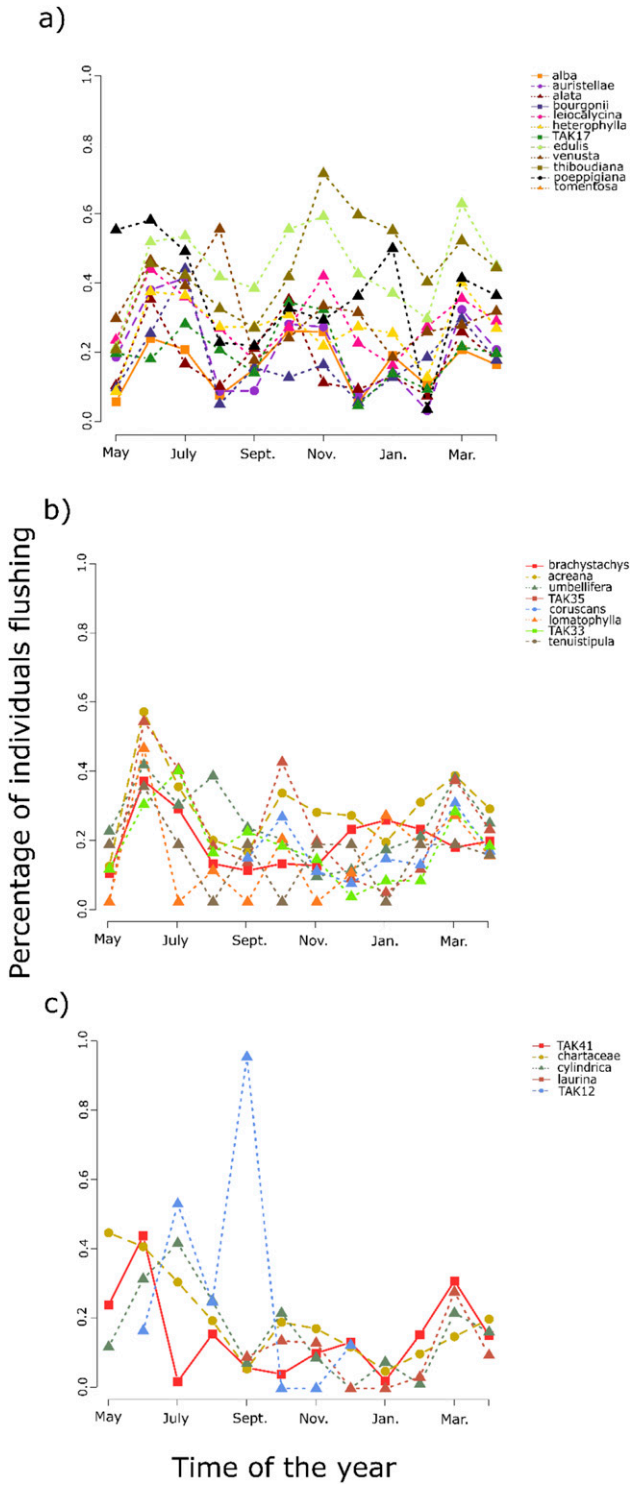

Fig. S1. Leaf production of 25 *Inga* species in Los Amigos between 2007 and 2008 as the proportion of individuals flushing (*A*) continuously, (*B*) moderately continuously, and (*C*) episodic (highly synchronized).

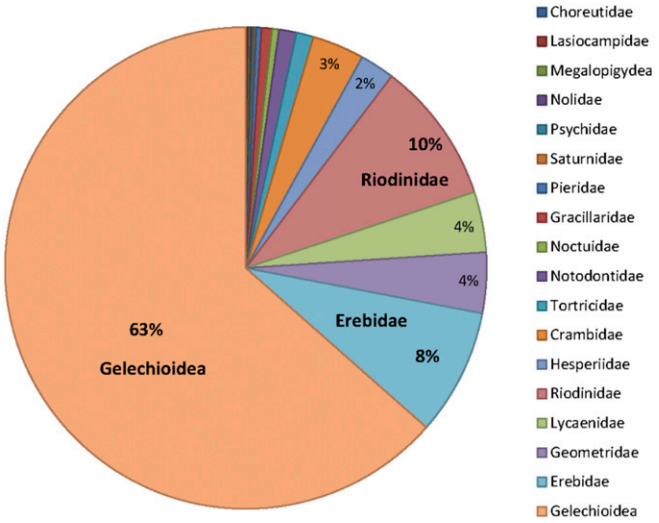

Fig. S2. Lepidopteran herbivore families associated with *Inga* in Los Amigos. Percentages represent the relative abundances. All individual herbivores were barcoded with the mitochondrial locus *COI*. In addition, nuclear genes (elongation-factor 1α, *EF-1α*, and wingless *Wg*) were sequenced for the Gelechioidea, Riodinidae, and Erebidae.

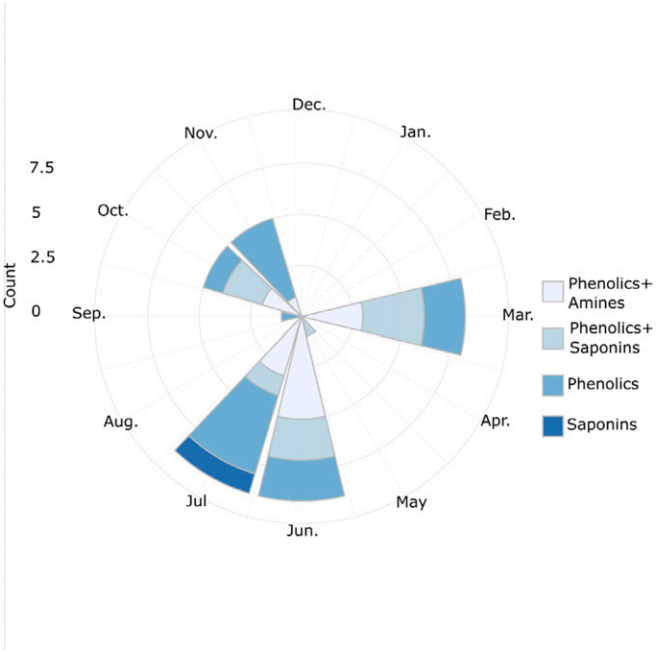

Fig. S3. Stacked circular histogram of peak flushing events for *Inga* species at Los Amigos between 2007 and 2008. Flushing events are grouped by defense chemistry.

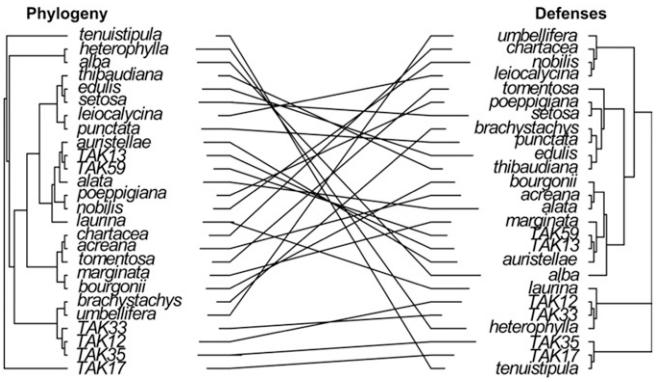

Fig. S4. Comparison between the phylogenetic tree (*Left*) and the defensogram (defense dendrogram; *Right*) for *Inga* species.

Table S1. Pairwise correlations between defense traits among *Inga* species

Leaf defense traits	Trichome length	Leaf expansion rate	Leaf chlorophyll content	Mean number of ants per nectary	Ant visitor community to extrafloral nectaries	Chemistry	Timing of leaf production	Synchrony in leaf production
Trichome density, no. of hairs per 2 cm^2	0.79†,*	0.1†	−0.07†	−0.08†	0.08	0.01	−0.013	0.31†
Trichome length, mm		−0.06†	−0.15†	−0.05†	0.22	0.07	−0.04	−0.2†
Leaf expansion rate, % per d^{-1}			−0.55†,*	0.09†	0.08	0.20	−0.012	0.15†
Leaf chlorophyll content, mg per m^{-2}				−0.24†	0.04	0.07	−0.04	−0.22†
Mean number of ants per nectary					0.28*	0.08	−0.05	0.14†
Ant visitor community to extrafloral nectaries						0.02	0.03	0.002
Chemistry (presence/absence of secondary compounds)							0.06	0.03
Timing of leaf production (mean angle)								0.07

Correlation coefficients with † are based on the phylogenetic generalized linear model (PGLS); the rest are partial Mantel *r*. Significant values ($P < 0.05$) are marked with an asterisk.

Table S2. Correlations between chemical defenses and phylogenetically PCA-derived axes (PPCA) of defense variation

PPCA axes	*R*	*P* (reps = 9,999)
PPCA 1 (physical defenses)	0.06	0.20
PPCA 2 (developmental defenses)	0.09	0.10
PPCA 3 (timing in leaf production)	0.15	0.06
PPCA 4 (biotic defenses)	0.11	0.82
PPCA 5 (synchrony in leaf production)	0.03	0.51

Correlation coefficients are partial Mantel *r*.

Table S3. Measure of phylogenetic signal for each phylogenetically PCA-derived axis (PPCA) and the principal coordinates of the chemistry and ant species distance matrices (PCO) using Blomberg's *K*

Defensive traits	*K* statistic	*P* (reps = 9,999)
PPCA 1 (physical defenses)	0.48	0.05
PPCA 2 (developmental defenses)	0.32	0.32
PPCA 3 (timing in leaf production)	0.24	0.81
PPCA 4 (biotic defenses–ant number)	0.23	0.81
PPCA 5 (synchrony in leaf production)	0.30	0.40
Chemistry PCO1 (46% of variation)	0.34	0.52
Chemistry PCO2 (19% of variation)	0.27	0.50
Biotic PCO1 (ant species, 17% of variation)	0.30	0.50
Biotic PCO2 (ant species, 15% of variation)	0.30	0.42
Biotic PCO3 (ant species, 12% of variation)	0.30	0.53

For PCO components, values in parentheses represent the percentage of variation explained by each component.

Table S4. *Inga* chemotypes

	Chemotypes													
Inga species	QAG	DHM	C/E	C/EG	C/ERG	GC/GE	GC/GEG	GC/GEC	FG	TYRG	TYMG	C\E PYR GAL	TYR	SAP
acreana	0	0	0	0	0	0	0	1	0	0	0	0	0	1
alata	0	0	0	0	0	1	0	0	0	0	0	0	0	1
alba	0	0	0	0	0	0	0	0	0	0	0	0	1	1
auristellae	0	0	0	0	0	0	1	0	0	0	0	0	0	1
bourgonii	0	0	1	0	0	0	0	0	0	0	0	0	0	1
brachystachys	0	0	0	1	0	0	0	0	1	0	0	0	0	0
capitata	1	0	0	0	0	0	0	0	1	0	1	0	0	0
chartacea	0	0	1	0	0	0	0	0	0	0	0	0	0	1
edulis	0	0	1	0	0	0	0	0	0	0	0	0	0	0
heterophylla	0	0	0	0	0	0	0	0	1	1	0	0	0	0
laurina	0	0	0	0	0	0	0	0	1	1	0	0	0	0
leiocalycina	0	0	0	0	1	0	0	0	0	0	0	0	0	0
marginata	0	0	0	0	0	0	0	1	1	0	0	0	0	1
nobilis	0	0	1	0	0	0	0	0	0	0	0	0	0	0
poeppigiana	0	0	0	0	0	0	1	0	1	0	0	0	0	0
punctata	0	0	0	0	0	0	1	0	1	0	0	0	0	0
sapindoides	0	0	0	0	0	0	0	0	0	0	0	0	0	1
setosa	0	0	0	0	0	1	0	0	0	0	0	0	0	0
TAK12	1	0	0	0	0	0	0	0	1	0	0	0	0	0
TAK13	0	0	0	0	0	0	0	0	0	0	0	0	0	1
TAK17	0	1	0	0	0	0	0	0	0	0	0	0	0	0
TAK35	1	0	0	0	0	0	0	0	0	1	1	0	0	0
TAK59	0	0	0	0	0	0	1	0	0	0	0	0	0	1
tenuistipula	0	0	1	0	0	0	0	0	0	0	0	0	0	0
thibaudiana	0	0	0	0	0	0	1	0	0	0	0	0	0	0
tomentosa	0	0	0	0	0	0	0	0	1	0	0	0	0	1
umbellifera	0	0	0	0	0	0	0	1	0	0	0	0	1	0
venusta	0	0	0	0	0	1	0	0	0	0	0	0	0	1

C/E, catechin/epicatechin; C/EG, catechin/epicatechin gallate; C/ERG, catechin/epicatechin–rhamnose–gallate; C/E PYR GAL, catechin/epicatechin pyranose gallate; DHM, dihydromyricetin; FG, flavone glycoside; GC/GE, gallocatechin/galloepicatechin; GC/GEC, gallocatechin/epigallocatechin–coumarate; GC/GEG, gallocatechin/epigallocatechin gallate; QAG, quinic acid–gallate; SAP, saponins; TYMG, tyramine gallate; TYR, tyrosine; TYRG, tyrosine gallate.

Table S5. PCR protocol for lepidopteran herbivores

Reagents	Volume in standard PCRs, μL
Multiplex Master Mix (Qiagen)	10
Water	6.6
Q-solution (Qiagen)	2.0
Forward 1 (20 μM)	0.2
Reverse 2 (20 μM)	0.2
DNA template	1.0
Total	20

Table S6. Taxon-specific primer pairs used in this study with target marker, optimized annealing temperature, and amplified product size

Primer pairs	Sequence, 5′–3′	Target group	Locus	Annealing temperature, °C	Expected length, bp	Source
LepF1	ATT CAA CCA ATC ATA AAG ATA TTG G	Lepidoptera	COI	49	658	Ref. 90
LepR1	TAA ACT TCT GGA TGT CCA AAA AAT CA					
EF51.9	CAR GAC GTA TAC AAA ATC GG	Gelechioidea	EF1α	53	511	Ref. 91
EFrcM4	ACA GCV ACK GTY TGY CTC ATR TC	Erebidae		55		
		Riodinidae		55.5		
LepWg1	GAR TGY AAR TGY CAY GGY ATG TCT GG	Gelechioidea	Wg	60	403	Ref. 92
LepWg2	ACT ICG CAR CAC CAR TGG AAT GTR CA	Erebidae		57		
		Riodinidae		54		
LepWg1a_mod	GAA TGT AAR TGT CAY GGY ATG TCY GG	Erebidae	Wg	53	403	This study
Lepwg2a_mod	ACT GCG CAG CAC CAR TKG AAT GTG CA	Riodinidae		54.3		
LepWg1	GAR TGY AAR TGY CAY GGY ATG TCT GG	Erebidae	Wg	54	348	This study
Ere_Wg_mod	GAT ACC CTC KIC CRC ARC	Riodinidae		51.5		

Table S7. Species tree parameters for reconstruction of phylogenetic relationships for Gelechioidea, Riodinidae, and Erebidae

Clades	Substitution models	Species tree priors	Tree rooting
Gelechioidea	COI (1): TN93	Strict clock	MOTUs from family Tortricidae
	COI (2): TN93	Constant population size	
	COI (3): TN93	Yule speciation model	
	EF1α (1, 2): HKY		
	EF1α (3): HKY		
	Wg (1, 2): HKY		
	Wg (3): HKY		
Riodinidae	COI (1): TN93+I	Strict clock	MOTUs from families Lycaenidae and Pieridae
	COI (2): HKY	Constant population size	
	COI (3): TN93+G	Yule speciation model	
	EF1α (1): HKY+I		
	EF1α (2): HKY		
	EF1α (3): HKY+G		
	Wg (1, 2): HKY+G		
	Wg (3): HKY		
Erebidae	COI (1): HKY+G	Strict clock	MOTUs from families Noctuidae, Nolidae, and Notodontidae
	COI (2): HKY+I	Linear population size	
	COI (3): HKY+G	Yule speciation model	
	EF1α (1, 2): HKY+I		
	EF1α (3): GTR+G		
	Wg (1, 2): GTR+G		
	Wg (3): GTR+G		

Numbers in parentheses represent codon positions.

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (XLSX)