

# Supporting Information

Asner et al. 10.1073/pnas.1401181111

## SI Methods

**Canopy Foliar Sampling.** Our sampling strategy focused on exhaustive surveys of as many sunlit canopy species, both common and rare, as possible over forest community areas of up to 600 ha and was broadly directed by historical surveys from the same or similar locations (1–3). Individual canopies meeting the full sunlight criterion were marked, and a voucher specimen was collected. Vouchers were matched by Carnegie Institution taxonomists to type specimens kept at the National Agrarian University La Molina Herbarium in Peru and the Missouri Botanical Garden. We also matched genus names to information provided by Kew Botanic Gardens. Family-level taxonomy followed the Angiosperm Phylogeny Group 3 (4). Because Angiosperm Phylogeny Group 3 uses detailed genetic information, our taxonomic analyses approximate phylogenetic analyses.

The foliar database is distributed among 106 families, 425 genera, and 2,420 species. Because of high species turnover between forest communities, the taxonomic partitioning within the sites ranged from 6 to 49 families, from 7 to 146 genera, and from 9 to 282 species (Table S2). Analyses of intraspecific variation were performed on a subset of 393 species, and each species contained between 3 and 13 individuals. Detailed information and maps for all species and sites are provided on the Carnegie Spectranomics Project website (<http://spectranomics.ciw.edu>). The website also lists species identities with taxon information. Duplicate vouchers for all samples are held in the Carnegie Institution herbarium section of the National Agrarian University La Molina Herbarium in Peru and the Carnegie Spectranomics Library at Stanford University.

Leaf collection campaigns were conducted using tree-climbing techniques. Only fully sunlit branches of mature leaves were taken and sealed in large polyethylene bags to maintain moisture, stored on ice in coolers, and transported to a local site for processing within 3 h, (usually less than 30 min). A subset of leaves was selected from the branches for scanning and weighing. Leaf area was determined on a 600 dots/in flatbed top-illumination optical scanner using enough leaves to fill two scan areas of 21 × 25 cm (up to about 75 leaves per sample depending on leaf size). Petioles were removed from each leaf before scanning, and midveins were cut out of the leaves when they exceeded 1 mm in diameter. Leaves exceeding the surface area of the scanner were cut into sections (without petiole or midvein if >1 mm in diameter) until two full scan areas were completed. The scanned leaves were then dried at 70 °C for a minimum of 72 h before dry mass (DM) was measured. Leaf mass per area was then calculated as grams DM meter<sup>-2</sup>. Also, from this subset of leaves, leaf discs (at least 30 per leaf) were immediately taken from 12 randomly selected leaves and transferred to –80 °C cryogenic containers and then climate-controlled –80 °C freezers until chemical assays were performed in the laboratory. The remaining leaves were detached from the branches, and subsamples were selected to represent the range of colors and conditions found among all leaves collected. When epiphylls were encountered, they were removed, along with dust and debris, before drying. These subsamples were dried in mobile ovens at 70 °C for a minimum of 72 h before vacuum sealing for transport to the laboratory for redrying before chemical analysis.

**Chemical Assays.** Chemical analysis protocols, along with instrument and standards information, are downloadable from the Carnegie Spectranomics Project website (<http://spectranomics.ciw.edu>) and summarized here. Dried foliage was ground in a

20-mesh Wiley mill, and subsets were analyzed for a variety of elements and carbon fractions. Total element concentrations of P and Ca were determined in 0.4 g dry leaf tissue by inductively coupled plasma spectroscopy (Therma Jarrel-Ash) after microwave digestion in 10 mL concentrated (~70% vol/vol) nitric acid solution (CEM MARSXpress). One blank and two reference standards (Peach NIST SRM 1547 and internal lemon leaf) were digested and measured with each set of 40 foliar samples to track the reproducibility and accuracy of the method.

Carbon fractions, including soluble C (composed of amino acids, pectins, simple sugars, and starch), hemicellulose, cellulose, and lignin, were determined in 0.5 g dry ground leaf tissue through use of sequential digestion of increasing acidity (5) in a fiber analyzer (Ankom Technology). C fractions are presented on an ash-free DM basis after ignition of the remaining sample at 500 °C for 5.5 h. Internal lemon leaf standard was used as a reference with each run to ensure consistency across runs. A subset of the ground material was further processed to a fine powder for determination of total C and N concentrations by combustion–reduction elemental analysis (Costec Analytical Technologies Inc.). A portion of the combustion gas from each sample was routed through an isotope ratio mass spectrometer (Finnigan S19; Thermo Scientific) for determination of  $\delta^{13}\text{C}$  in the sample. Reference standards (Peach NIST SRM 1547 and internal lemon leaf) were included with every set of 20 samples.  $\delta^{13}\text{C}$  was calculated on a per mil basis (‰) with respect to the Pee Dee Belemnite standard.

Frozen leaf disks were used for the total phenolic, chlorophyll, and carotenoid determinations. For phenols, disks were ground in 95% methanol on the high-throughput tissue homogenizer. A portion of the solution was further diluted and incubated on an orbital shaker at room temperature (15–18 °C) in the dark for 48 h to ensure proper phenol extraction (6). The total phenolic concentration in solution was determined colorimetrically using the Folin–Ciocalteu method. Phenol concentrations were measured in gallic acid equivalents relative to an eight-point Gallic acid standard curve. Total chlorophyll and total carotenoid concentrations were quantified using two frozen leaf disks (0.77 cm<sup>2</sup> area each). These disks were rapidly ground (90 s) in 1.5-mL centrifuge tubes containing 0.75 mL 100% acetone on a high-throughput tissue homogenizer (Troemner) with a small amount of MgCO<sub>3</sub> to prevent acidification. After dilution and centrifugation for 3 min at 2,000 × g, the absorbance of the supernatant was measured using a dual-beam scanning UV-VIS spectrometer (Lambda 25; Perkin-Elmer).

**Analyses.** We used ordinary least squares regression to assess relationships between log-transformed leaf traits, elevation, mean annual temperature, mean annual precipitation, and their interactions. We also assessed intra- and interspecific variation using coefficients of variation calculated with untransformed data. We used ANOVA tests to compare chemical traits on lower- vs. higher-fertility sites based on US Department of Agriculture soil taxonomy.

With the goal of examining how variance in chemical data can be explained by taxonomic grouping, we developed nested ANOVA models with random effects using the *lme4* (residual maximum likelihood) package in R (7, 8). We included the phylogenetic levels of family (f), genus nested within family (g), and species nested within genus within family (s) as well as an environmental component incorporated as site (T). All effects were treated as random. In each model, y is any chemical trait

for each canopy sample. This value was modeled as the sum of the mean value for the entire dataset  $\mu$  (or subset, when specified), the nested genetic effects (family  $i$ , genus  $j$  within family  $i$ , and species  $ijk$  within genus  $j$ ), the site effect ( $T$ ), and the residual error of the measurement  $e$ :

$$y = \mu + f_i + g_{ij} + s_{ijk} + T_i + e_{ijkl}.$$

The total variance about the mean for a given trait was, therefore, quantitatively parsed into the variance explained by families ( $\sigma_f^2$ ), genera within families ( $\sigma_g^2$ ), species within genera ( $\sigma_s^2$ ), site ( $\sigma_T^2$ ), and specimens within species ( $\sigma_e^2$ ):

$$\sigma_{\text{total}}^2 = \sigma_f^2 + \sigma_g^2 + \sigma_s^2 + \sigma_T^2 + \sigma_e^2.$$

If, in a given model, the last term ( $\sigma_e^2$ ) accounted for a high percentage of the total variance, then we concluded that site characteristics and taxonomy did not explain the data well. We refer to this component as the model residual.

1. Gentry AH (1988) Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Ann Mo Bot Gard* 75(1):1–34.
2. Phillips OL, et al. (2003) Habitat association among Amazonian tree species: A landscape-scale approach. *J Ecol* 91(5):757–775.
3. Silman MR (2006) Plant species diversity in Amazonian forests. *Tropical Rain Forest Responses to Climate Change*, eds Bush M, Flenly J (Springer-Praxis, London).
4. Stevens PF (2001) *Angiosperm Phylogeny Website*. Available at [www.mobot.org/MOBOT/research/APweb/](http://www.mobot.org/MOBOT/research/APweb/). Accessed November 10, 2013.
5. Van Soest PJ (1994) *Nutritional Ecology of the Ruminant* (Cornell Univ Press, Ithaca, NY), 2nd Ed.
6. Ainsworth EA, Gillespie KM (2007) Estimation of total phenolic content and other oxidation substrates in plant tissues using Folin-Ciocalteu reagent. *Nat Protoc* 2(4):875–877.

One limitation of this analysis is that it describes the overall variation explained by each input variable. We acknowledge that not all taxa have equal variance; some may have tightly clumped chemical signatures, whereas others may vary widely. This analysis will not pick up such trends; instead, the method quantifies the entire pattern of phylogenetic grouping or lack thereof relative to site and residual effects. Previous work successfully tested the validity of nested random effects modeling for analysis of phylogenetic partitioning of foliar chemical traits (9–11).

To compare rates of change of multiple canopy chemicals, we computed the gradient-normalized trait values at each site. This procedure was done by subtracting the mean chemical trait value of the entire gradient ( $M_{\text{gradient}}$ ) from the mean value of each site ( $m_{\text{site}}$ ) and dividing the difference by the gradient SD ( $SD_{\text{gradient}}$ ):

$$(m_{\text{site}} - M_{\text{gradient}}) / SD_{\text{gradient}}.$$

We repeated this procedure for chemical traits expressed on a mass basis (Fig. 1) and an area basis (Fig. S2).

7. Faraway JJ (2005) *Extending the Linear Model with R: Generalized Linear, Mixed Effects, and Nonparametric Regression Models* (Chapman & Hall/CRC, New York).
8. Bates D, Maechler M, Bolker B (2012) Package 'lme4' [documentation file]. Available at <http://cran.stat.sfu.ca/web/packages/lme4/lme4.pdf>. Accessed August 8, 2012.
9. Fyllas N, et al. (2009) Basin-wide variations in foliar properties of Amazonian forest: Phylogeny, soils and climate. *Biogeosciences* 6:2677–2708.
10. Asner GP, Martin RE (2011) Canopy phylogenetic, chemical and spectral assembly in a lowland Amazonian forest. *New Phytol* 189(4):999–1012.
11. Asner G, Martin R, Suhaili A (2012) Sources of canopy chemical and spectral diversity in lowland Bornean forest. *Ecosystems* 15(3):504–517.





**Table S2. Taxonomic partitioning of foliar samples from canopy trees in 19 sites along the Andes–Amazon elevation gradient in Peru**

Site name	Individuals	Families	Genus	Species
Sucusari	334	49	124	230
Allpahuayo 1	338	44	140	222
Jenaro Herrera 1	437	55	146	282
Jenaro Herrera 2	84	25	48	55
Jenaro Herrera 3*	62	17	37	44
Allpahuayo 2	344	48	120	213
Inkaterra*	336	48	108	178
Tambopata 1	344	39	107	198
Tambopata 2*	204	40	94	129
Los Amigos 1*	178	34	80	120
Los Amigos 2	206	36	76	120
Paujil 1	208	40	87	146
Paujil 2	46	19	29	33
Huampal*	310	49	120	186
San Pedro 1*	130	34	54	76
San Pedro 2*	143	38	63	104
Tres Cruces 1*	72	22	23	37
Tres Cruces 2*	60	15	20	33
Tres Cruces 3*	20	6	7	9

\*Site considered to be higher fertility in this study (Table S1).



**Table S4. Summary of standard least squares regression models relating mass-based leaf traits to environment along the Andes–Amazon elevation gradient**

	Adjusted $r^2$ (RMSE)	Elevation (m)	Adjusted $r^2$ (RMSE)	MAT	MAP	MAT × MAP
<b>All forests</b>						
$\delta^{13}\text{C}$	0.93 (0.36)	15.37*	0.98 (0.2)*	-18.67*	-3.50 <sup>†</sup>	nr
LMA	0.52 (15.51)	4.49*	0.52 (15.49) <sup>†</sup>	nr	nr	nr
Total C	nr		nr	nr	2.42 <sup>†</sup>	nr
Soluble C	0.82 (2.79)	9.15*	0.91 (1.97)*	-10.53*	-4.01*	nr
Chl	0.22 (1.07)	-2.47 <sup>†</sup>	0.38 (0.95) <sup>†</sup>	nr	nr	-2.74 <sup>†</sup>
Car	nr		0.31 (0.17) <sup>†</sup>	nr	nr	-2.85 <sup>†</sup>
N	nr		nr	nr	nr	nr
P	0.25 (0.04)	2.62 <sup>†</sup>	0.43 (0.04) <sup>†</sup>	-3.46 <sup>†</sup>	-2.27 <sup>†</sup>	nr
Ca	nr		nr	nr	nr	nr
Phenols	nr		nr	nr	nr	nr
Lignin	0.49 (3.60)	-4.32*	0.68 (2.87)*	5.05*	2.91 <sup>†</sup>	nr
Cellulose	0.88 (0.76)	-11.75*	0.95 (0.51)*	14.7*	3.48 <sup>†</sup>	2.76*
<b>Higher-fertility soils</b>						
$\delta^{13}\text{C}$	0.92 (0.44)	10.11*	0.99 (0.09)*	-7.22*	nr	5.82*
LMA	0.85 (11.20)	7.28*	0.84 (11.6) <sup>†</sup>	nr	nr	nr
Total C	0.68 (0.76)	4.44 <sup>†</sup>	0.65 (0.78) <sup>†</sup>	nr	nr	nr
Soluble C	0.86 (2.31)	7.50*	0.93 (1.67) <sup>†</sup>	nr	nr	nr
Chl	0.73 (0.76)	-5.05*	0.86 (0.54) <sup>†</sup>	nr	nr	nr
Car	0.68 (0.13)	-4.51 <sup>†</sup>	0.85 (0.09) <sup>†</sup>	nr	nr	nr
N	0.50 (0.15)	-3.18 <sup>†</sup>	nr	nr	nr	nr
P	nr		nr	nr	nr	nr
Ca	0.63 (0.32)	-4.01 <sup>†</sup>	0.66 (0.31) <sup>†</sup>	nr	nr	nr
Phenols	nr		nr	nr	nr	nr
Lignin	0.41 (2.67)	-2.71 <sup>†</sup>	0.65 (2.07) <sup>†</sup>	nr	nr	nr
Cellulose	0.90 (0.69)	-9.30*	0.97 (0.41)*	5.97*	nr	nr

Adjusted  $r^2$  values show RMSE in parentheses, and the  $t$  values for model variables are provided. nr, no relationship at the  $P = 0.05$  level; RMSE, root mean square error.

\* $P < 0.001$  significance value.

<sup>†</sup> $P < 0.05$  significance value.

**Table S5. Summary of standard least squares regression models relating leaf traits calculated on an area basis to elevation for all forests and higher-fertility forests (Table S1)**

	All forests		Higher-fertility forests	
	Adjusted $r^2$ (RMSE)	$t$	Adjusted $r^2$ (RMSE)	$t$
d13C	0.47 (0.04)	4.15*	0.92 (0.22)	9.53*
LMA	0.52 (15.5)	4.51*	0.90 (0.29)	8.49*
Total C	0.43 (0.06)	-3.85*	0.90 (0.21)	-8.40*
Soluble C	nr	nr	0.65 (0.20)	-3.95 <sup>†</sup>
Chl	0.21 (0.02)	-2.38 <sup>†</sup>	0.85 (0.23)	-6.93*
Car	0.17 (0.01)	-2.15 <sup>†</sup>	0.85 (0.22)	-6.80*
N	0.19 (0.01)	-2.27 <sup>†</sup>	0.87 (0.18)	-7.51*
P	nr	nr	0.71 (0.17)	-4.59 <sup>†</sup>
Ca	nr	nr	0.70 (0.37)	-4.42 <sup>†</sup>
Phenols	0.29 (0.17)	-2.9 <sup>†</sup>	0.46 (0.24)	-2.81 <sup>†</sup>
Lignin	0.76 (0.03)	-7.65*	0.90 (0.19)	-8.51*
Cellulose	0.67 (0.03)	-6.18*	0.92 (0.21)	-9.63*

Adjusted  $r^2$  values show RMSE in parentheses, and the  $t$  values for model variables are provided.

\* $P < 0.001$  significance value.

<sup>†</sup> $P < 0.05$  significance value.

**Table S6. Summary of least squares regression models relating phylogenetic components of leaf traits to elevation along the Andes–Amazon elevation gradient**

	Family		Genus		Species		Residual	
	Adjusted $r^2$ (RMSE)	Elev. (m)	Adjusted $r^2$ (RMSE)	Elev. (m)	Adjusted $r^2$ (RMSE)	Elev. (m)	Adjusted $r^2$ (RMSE)	Elev. (m)
<b>All forests</b>								
$\delta^{13}\text{C}$	0.01 (10.38)	1.07	-0.03 (20.6)	-0.72	-0.07 (13.02)	0.15	-0.06 (9.75)	0.18
LMA	0.19 (12.59)	2.16*	-0.05 (12.3)	-0.50	-0.06 (16.47)	0.22	0.06 (17.31)	-1.42
Total C	-0.07 (15.04)	-0.06	-0.06 (9.4)	0.34	-0.05 (14.11)	-0.50	-0.02 (5.84)	0.82
Soluble C	0.1 (13.95)	1.65	-0.03 (13.72)	0.76	0.18 (13.41)	-2.12*	-0.04 (7.93)	-0.61
Chl	0.23 (9.39)	2.42*	0.01 (16.78)	-1.06	0.00 (12.38)	1.02	0.02 (15.36)	-1.14
Car	0.10 (14.41)	1.66	-0.03 (16.96)	-0.69	-0.06 (14.28)	0.21	0.00 (15.45)	-0.98
N	0.01 (14.28)	1.11	-0.06 (10.83)	-0.31	0.03 (11.67)	-1.24	-0.06 (6.57)	0.31
P	-0.06 (14.96)	-0.19	-0.05 (11.16)	-0.46	-0.04 (12.01)	0.59	-0.07 (11.6)	0.08
Ca	-0.04 (17.66)	0.61	0.06 (12.51)	-1.44	-0.06 (18.07)	-0.16	0.02 (9.08)	1.12
Phenols	0.15 (16.28)	1.94	0.00 (17.51)	-0.99	0.04 (15.09)	-1.27	-0.06 (13.56)	0.36
Lignin	0.04 (15.36)	1.32	-0.03 (14.15)	-0.71	-0.01 (16.86)	-0.91	-0.04 (7.93)	0.64
Cellulose	0.07 (12.14)	-1.49	0.01 (19.17)	1.06	-0.07 (16.4)	0.06	-0.04 (5.74)	-0.57
<b>Higher-fertility soils</b>								
$\delta^{13}\text{C}$	0.04 (14.41)	-1.14	-0.16 (10.63)	0.17	-0.04 (12.37)	0.84	0.06 (3.52)	1.21
LMA	-0.06 (21.19)	0.78	0.07 (14.69)	-1.23	-0.08 (8.34)	-0.71	-0.04 (8.99)	0.84
Total C	0.24 (13.96)	1.80	0.00 (13.71)	-0.99	-0.10 (13.5)	0.61	0.11 (14.34)	-1.38
Soluble C	0.08 (5.61)	1.26	-0.16 (13.64)	-0.24	-0.10 (10.15)	-0.60	-0.14 (6.55)	0.36
Chl	-0.12 (18.73)	-0.49	-0.05 (9.28)	-0.80	-0.11 (16.81)	0.55	-0.10 (11.71)	0.63
Car	-0.02 (17.06)	0.92	-0.02 (14.08)	-0.93	-0.11 (18.17)	0.54	0.07 (9.89)	-1.25
N	0.23 (11.91)	1.75	0.08 (15.99)	-1.27	0.17 (11.98)	1.55	0.06 (15.82)	-1.21
P	0.12 (19.61)	1.41	-0.01 (16.53)	-0.96	0.02 (15.59)	-1.06	-0.12 (9.9)	0.48
Ca	0.11 (10.82)	1.36	-0.06 (22.17)	-0.78	-0.13 (13.95)	0.45	-0.13 (8.05)	-0.46
Phenols	-0.13 (16.68)	0.46	-0.13 (13.2)	0.43	0.34 (8.48)	-2.13	-0.04 (5.32)	0.87
Lignin	0.14 (12.03)	-1.48	-0.07 (23.97)	0.72	-0.17 (14.16)	-0.06	-0.15 (4.48)	0.29
Cellulose	-0.13 (10.75)	0.41	-0.02 (13.84)	-0.93	-0.10 (15.86)	0.59	-0.17 (12.22)	-0.07

Adjusted  $r^2$  values show RMSE in parentheses, and the  $t$  values for model variables are provided.

\* $P < 0.05$  significance relationship.







