



1  
2

**Table S1:** Statistical models testing the genetic specificity of the plant-insect food web.

Response	df	<i>F</i> or $\chi^2$	<i>P</i>
Gall size <sup>1</sup>			
Leaf gall	23,57	2.17	<b>0.009</b>
Bud gall	21,44	0.98	0.504
Apical-stem gall	16,12	0.29	0.988
Gall abundance <sup>2</sup>	25,119	202.40	<b>0.001</b>
Leaf gall		74.60	<b>0.001</b>
Bud gall		55.02	<b>0.006</b>
Apical-stem gall		44.47	<b>0.042</b>
Mid-stem gall		28.27	0.295
Composition of gall community <sup>3</sup>	22,89	1.96	<b>0.001</b>
Abundance of gall-parasitoid interactions <sup>2</sup>	25,119	357.10	<b>0.001</b>
Leaf gall			
<i>Platygaster</i> sp.		79.51	<b>0.001</b>
<i>Mesopolobus</i> sp.		50.00	<b>0.009</b>
<i>Torymus</i> sp.		60.11	<b>0.001</b>
<i>Tetrastichus</i> sp.		32.96	0.105
Mymarid sp. A		6.37	0.448
Bud gall			
<i>Platygaster</i> sp.		18.04	0.276
<i>Mesopolobus</i> sp.		6.37	0.497
<i>Torymus</i> sp.		39.81	0.079
<i>Tetrastichus</i> sp.		18.09	0.492
<i>Lestodiplosis</i> sp.		16.05	0.552
Apical-stem gall			
<i>Torymus</i> sp.		23.13	<b>0.048</b>
Mid-stem gall			
<i>Platygaster</i> sp.		6.64	0.452
Composition of gall-parasitoid interactions <sup>3</sup>	12,45	1.57	<b>0.007</b>
Proportion of galls parasitized <sup>4</sup>			
Leaf gall	23,58	75.79	<b>&lt;0.001</b>
<i>Platygaster</i> sp.		93.47	<b>&lt;0.001</b>
<i>Mesopolobus</i> sp.		42.56	<b>0.008</b>
<i>Torymus</i> sp.		42.92	<b>0.007</b>
<i>Tetrastichus</i> sp.		29.55	0.163
Mymarid sp. A		3.97	0.999
Bud gall	21,46	49.84	0.072
Apical-stem gall	18,12	15.69	0.614
Composition of trophic interactions in the plant-insect food web <sup>3</sup>	22,89	1.90	<b>0.001</b>

Notes: <sup>1</sup>GLM (error distribution = Gaussian, link function = identity), log-transformed;  
<sup>2</sup>multivariate GLM (error distribution = negative binomial, link function = log);  
<sup>3</sup>PERMANOVA on Bray-Curtis dissimilarities (999 permutations);  
<sup>4</sup>GLM (error distribution = binomial, link function = logit). *P*-values in bold ( $P < 0.05$ ), italics ( $P < 0.10$ ), and normal font ( $P > 0.10$ ) denote degree of statistical significance.

**Table S2:** Statistical models explaining insect food web responses to genetic variation in coastal willow (*Salix hookeriana*). We report the coefficients of all predictor variables that were included in the final statistical models, which were determined using AIC and likelihood-ratio tests.

Response	Predictors			
Gall size <sup>1</sup>	Salicylates/ Tannins PC1	Flavones/ Flavonols PC1		
Leaf gall	<b>-0.20</b>	<b>-0.26</b>		
Gall abundance <sup>2</sup>	C:N	Flavanones/ Flavanonols PC1	Plant size	
Leaf gall	<i>0.04</i>	-0.03	-0.36	
Bud gall	<i>0.08</i>	-0.07	<b>-1.01</b>	
Apical-stem gall	0.01	<b>0.46</b>	0.26	
Mid-stem gall	0.02	-1.81	<i>-4.77</i>	
Abundance of gall-parasitoid interactions <sup>2</sup>	Leaf gall size	Leaf gall abundance	Bud gall abundance	Apical-stem gall abundance
Leaf gall				
<i>Platygaster</i> sp.	<b>-0.22</b>	<b>1.22</b>	0.20	-0.15
<i>Mesopolobus</i> sp.	<b>-0.27</b>	<b>0.90</b>	-0.26	0.44
<i>Torymus</i> sp.	<i>0.19</i>	<b>0.76</b>	-0.30	0.72
<i>Tetrastichus</i> sp.	<i>-0.24</i>	0.71	0.45	-1.09
Mymarid sp. A	-1.67	<b>20.83</b>	-2.07	3.35
Bud gall				
<i>Platygaster</i> sp.	0.43	0.23	<b>5.81</b>	-14.25
<i>Mesopolobus</i> sp.	0.16	0.30	0.77	1.95
<i>Torymus</i> sp.	<b>-0.17</b>	0.31	<b>1.39</b>	-0.43
<i>Tetrastichus</i> sp.	0.15	0.51	<b>1.83</b>	0.08
<i>Lestodiplosis</i> sp.	0.04	-0.61	<i>1.46</i>	1.75
Apical-stem gall				
<i>Torymus</i> sp.	-0.12	0.05	-0.64	<b>4.09</b>
Mid-stem gall				
<i>Platygaster</i> sp.	1.54	<i>-15.03</i>	0.53	-9.23

Notes: <sup>1</sup>GLM (error distribution = Gaussian, link function = identity), log-transformed;  
<sup>2</sup>multivariate GLM (error distribution = negative binomial, link function = log). *P*-values in bold ( $P < 0.05$ ), italics ( $P < 0.10$ ), and normal font ( $P > 0.10$ ) denote degree of statistical significance.

**Table S3:** Generalized linear models (error distribution = binomial, link function = logit) explaining the proportion of leaf galls parasitized. Final models were determined using AIC and likelihood-ratio tests.

Response	Predictor	df	$\chi^2$	P
Total parasitism	Gall size	1,79	22.28	<0.001
<i>Platygaster</i> sp.	Gall size	1,77	17.58	<0.001
	Gall abundance	1,77	0.73	0.394
	Gall size x abundance	1,77	8.71	0.003
<i>Mesopolobus</i> sp.	Gall size	1,77	7.28	0.007
	Gall abundance	1,77	0.29	0.588
	Gall size x abundance	1,77	4.21	0.040
<i>Torymus</i> sp.	Gall size	1,78	3.83	0.050
	Gall abundance	1,78	5.24	0.022

**Relatedness and functional-trait diversity of willow genotypes** – The matrix of microsatellite markers for the 26 willow genotypes used in this study was published in Table S1 of (1); however, since the willow genotyping was only based on 2 markers, they were unable to infer the relatedness among genotypes. If certain genotypes are more closely related to each other, and consequently have very similar phenotypes, this could introduce spurious confidence in our associations between willow traits and gall abundances/phenotypes. We can examine this phenotypic similarity by measuring the functional evenness and divergence of the 26 willow genotypes in multivariate trait space (2). To do this, we calculated the average trait value for each of the 40 traits we measured for each willow genotype. We then calculated functional evenness and functional divergence using the ‘FD’ package in R. For both indices, values close to zero correspond to functional redundancy, while values close to one indicate functional distinctiveness. We found that functional evenness and divergence were equal to 0.94 and 0.87, respectively, suggesting that the phenotypes (in multivariate trait space) of each genotype are quite distinct from each other. Therefore, we argue that not knowing the relatedness among the 26 genotypes probably introduces little bias in our trait associations with the abundances and sizes of galls.

**Sampling interactions in gall-parasitoid network** – The total number of potential gall-parasitoid interactions in this bipartite network is 24 (i.e. each of the 4 galls could interact with each of the 6 parasitoids,  $6 \times 4 = 24$ ). Interspecific differences among gall species (e.g. differences in gall morphology, phenology, plant part galled) and sampling effort likely constrain the number of potential interactions observed to considerably less than 24. While it was not the focus of our study to examine interspecific differences, it is important to demonstrate that we have sampled the majority of interactions in the gall-parasitoid network. To demonstrate this, we considered unique gall-parasitoid interactions as ‘species’ and used Chao 1 (3) to estimate the total number of interactions. While we documented 12 unique gall-parasitoid interactions, Chao 1 estimated the number of interactions to be 14.98 (std. error = 4.49), suggesting that we have sampled the majority of interactions in the gall-parasitoid network.

**Calculating quantitative-weighted linkage density (food-web complexity)** –

Quantitative-weighted linkage density,  $LD_q$ , was calculated using the following equations (4). Given an  $s$ -by- $s$  food web matrix  $\mathbf{b} = [b_{ij}]$ , with  $b_{ij}$  corresponding to the number of individuals of species  $j$  (galls or parasitoids) emerging from species  $i$  (willow or galls) per willow branch over a single growing season,  $b_{i.}$  is the sum of row  $i$ ,  $b_{.j}$  is the sum of column  $j$ , and  $b_{..}$  is the total sum. The Shannon indices for the prey and predatory interactions were calculated as,

$$H_j = - \sum_{i=1}^s \frac{b_{ij}}{b_{.j}} \ln \frac{b_{ij}}{b_{.j}}$$

$$H_i = - \sum_{j=1}^s \frac{b_{ij}}{b_{i.}} \ln \frac{b_{ij}}{b_{i.}}$$

The effective number of prey and predatory interactions were calculated as  $N_j^* = \exp(H_j)$  and  $N_i^* = \exp(H_i)$ , respectively. Finally, quantitative-weighted link density was calculated as,

$$LD_q = \frac{1}{2b_{..}} \left( \sum_{i=1}^s b_{i.} N_i^* + \sum_{j=1}^s b_{.j} N_j^* \right)$$

**Asymptotic vs. non-asymptotic models** – We fit both asymptotic and non-asymptotic phenomenological models (5) to extrapolate our estimates of food-web complexity. While more sophisticated and accurate models have been developed to extrapolate species richness (3), nothing has been developed for extrapolating food-web complexity. These phenomenological models have the advantage that they make no assumptions about the processes generating the data (3); therefore, they are likely a good starting point for extrapolating food-web complexity.

For our asymptotic model we used a scaled and shifted Michaelis-Menten function (6) of the form,

$$LD_{q,N} = \frac{a(N-1)}{b+(N-1)} + \overline{LD}_{q,1},$$

where  $N$  represents either the number of plants (sampling effort simulation) or the number of genotypes (genetic variation simulation).  $LD_{q,N}$  is the predicted complexity at  $N$ , while  $a$  and  $b$  are phenomenological parameters that scale  $LD_{q,N}$  and  $N$ , respectively.  $\overline{LD}_{q,1}$  is a constant parameter, representing the average complexity for mixtures of either 1-genotype 1-plant (sampling effort simulation) or 1-genotype 4-plants (genetic variation simulation). Adding the constant,  $\overline{LD}_{q,1}$ , and subtracting the constant, 1, shift the function so that when  $N = 1$ ,  $LD_{q,N} = \overline{LD}_{q,1}$ . We used non-linear least squares to estimate parameters  $a$  and  $b$ . For the non-asymptotic models, we fit log-log ( $\log(LD_{q,N}) = m * \log(N) + intercept$ ) and log-linear ( $LD_{q,N} = m * \log(N) + intercept$ ) models. The asymptotic and non-asymptotic models we chose have been widely used for extrapolating species richness (5), which is why we used them for food-web complexity.

**Results for simulations of sampling effort and genetic variation** – We fit the asymptotic and non-asymptotic models to our sampling effort simulations of 1-genotype mixtures of 1 to 4 plants (1,000 estimates per level of sampling effort, details in *Materials and Methods*). We found that all of the models gave a similar fit to the data; however, they gave very different predictions for the complexity of 1-genotype 100-plant mixtures (Table S4). Therefore, to evaluate which of these models was more realistic, we re-fit these models to our genetic variation simulations of 1 to 25 genotypes (grey circles in Fig. 6 of main text). We found that the asymptotic model provided a much better fit ( $R^2 = 0.96$ ) and more accurate predictions than either of the non-asymptotic models (Table S5). In particular, the asymptotic model's predicted complexity of 25-genotype 100-plant mixtures deviated less than a tenth of 1% from the observed average ( $LD_q = 2.209$ ), whereas the non-asymptotic models overestimated complexity by 2.4% (log-linear) and 3.1% (log-log).

**Table S4:** Comparing asymptotic and non-asymptotic models for predicting the complexity of 1-genotype 100-plant mixtures. Note that for these data (sampling effort simulation),  $N$  represents the number of plants.

Model type	Equation	$R^2$	Predicted $LD_q$ 1-genotype 100-plant mixture
Asymptotic (Michaelis-Menten)	$LD_{q,N} = \frac{0.62(N-1)}{(3.62 + (N-1))} + 1.25$	0.885	1.84
Non-asymptotic (log-log)	$\log(LD_{q,N}) = 0.15 * \log(N) + 0.22$	0.881	2.45
Non-asymptotic (log-linear)	$LD_{q,N} = 0.20 * \log(N) + 1.24$	0.884	2.17

**Table S5:** Comparing asymptotic and non-asymptotic models for predicting the complexity of 25-genotype 100-plant mixtures. The observed complexity of the 25-genotype 100-plant mixture was 2.209. Note that for these data (genetic variation simulation),  $N$  represents the number of genotypes.

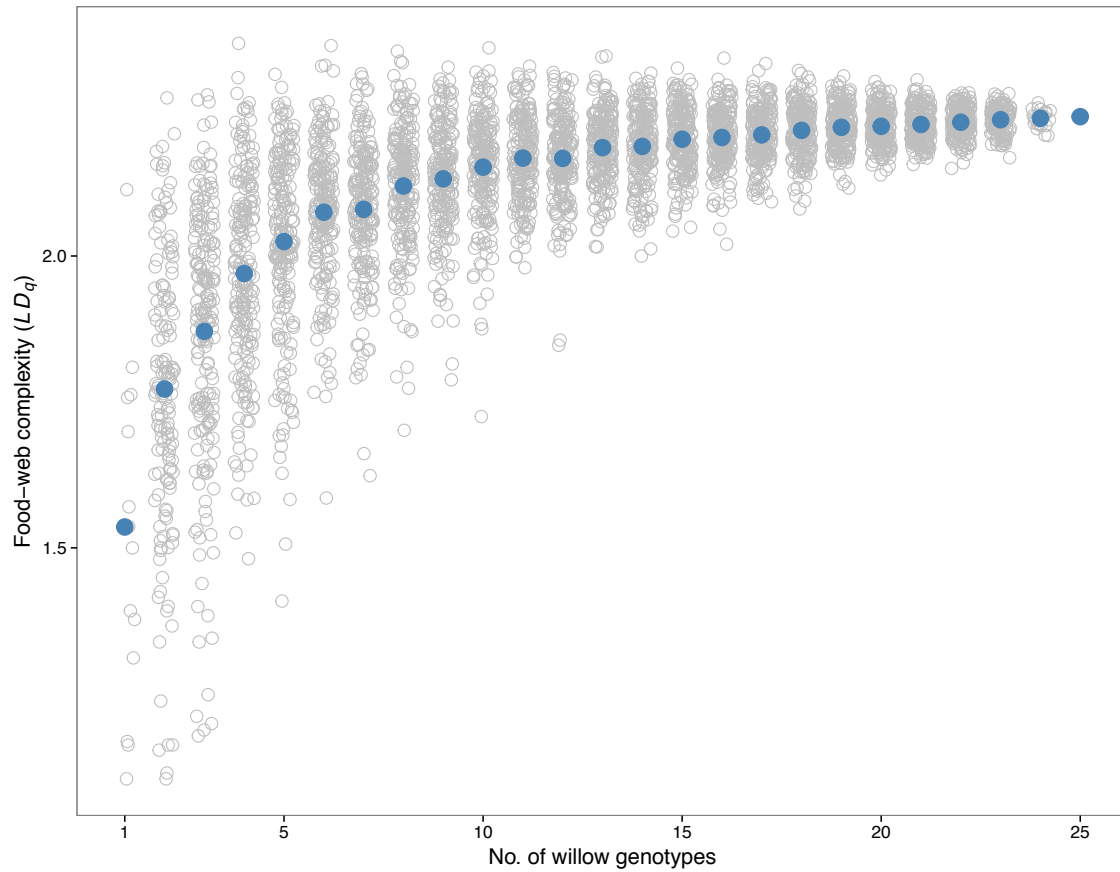
Model type	Equation	$R^2$	Predicted $LD_q$ 25-genotype 100-plant mixture
Asymptotic (Michaelis-Menten)	$LD_{q,N} = \frac{0.76(N-1)}{(2.25 + (N-1))} + 1.52$	0.96	2.210
Non-asymptotic (log-log)	$\log(LD_{q,N}) = 0.10 * \log(N) + 0.50$	0.87	2.277
Non-asymptotic (log-linear)	$LD_{q,N} = 0.19 * \log(N) + 1.65$	0.89	2.262

**Assessing the accuracy of the asymptotic model** – After we identified the asymptotic model as the most appropriate for our data, we wanted to evaluate whether the model was likely to over- or under-estimate the complexity of 1-genotype 100-plant mixtures. To do

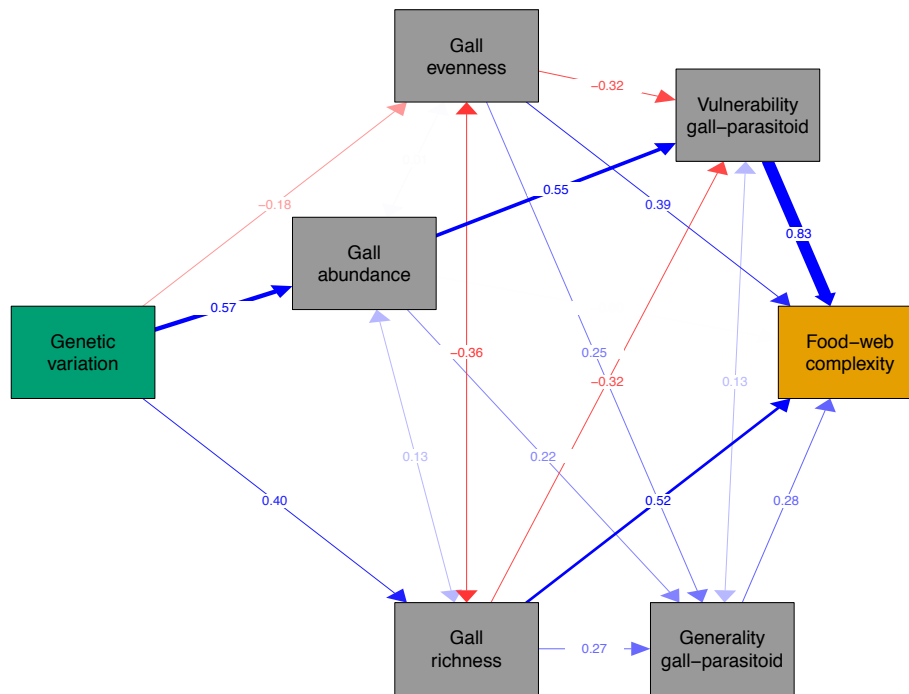
this, we took advantage of the complete data set we had for the genetic variation simulation. Specifically, we refit the asymptotic model with smaller fractions of data to examine how accurately it extrapolated to predict the complexity of 25-genotype 100-plant mixtures. When we did this, we found that the model began to increasingly overestimate food-web complexity, but only slightly. For example, using only the first 40% of the data (i.e. 1 to 10 genotypes), the model overestimated food-web complexity by less than 0.5%, while, using only the first 16% of the data (e.g. 1 to 4 genotypes), the model overestimated food-web complexity by 0.9%. Since our asymptotic model for the sampling effort simulation is extrapolating based on 4% of the potential data (4 of 100 plants), the predicted complexity of 1-genotype 100-plant mixtures is likely an overestimate. This suggests that the reported effect of 20% is a conservative estimate of the additive effects of genetic variation.

**Structural equation model of food-web complexity** – For our plant-insect food web, complexity is principally determined by three components: (i) the effective number of gall species per willow (i.e. Shannon diversity of galls); (ii) the effective number of parasitoid species per gall (vulnerability,  $V_q$ ); and (iii) the effective number of gall species per parasitoid (generality,  $G_q$ ) (4). Increases in any of these 3 components, all else equal, will directly increase food-web complexity. Moreover, the total abundance and diversity of galls may indirectly affect complexity by influencing the vulnerability and generality of the gall-parasitoid network. Therefore, we built our structural equation model to incorporate these different pathways. In addition, since species diversity is determined by both the evenness and richness of a community, we partitioned gall diversity into its evenness ( $E^l = \exp(\text{Shannon diversity})/\text{richness}$ ) and richness components (7) before building the model. Given the non-linear relationship between genetic variation and food-web complexity (Fig. S1), we restricted our analysis to the first 4-levels of genetic variation. We feel this was justified for two reasons: (i) this was the portion of the relationship that increased the most; and (ii) this was the only portion of the relationship that was mostly linear with constant variance, thereby satisfying the assumptions of the linear regression models that made up our structural equation model. Finally, we used a test of directed separation (8), which essentially tests whether there are any significant paths missing from the model. For tests of directed separation,  $P > 0.05$  indicates that the model provides a good fit to the data (i.e. no missing paths), whereas  $P < 0.05$  indicates a model with missing paths.

Fig. S1 shows the data from the one replicate simulation that we used to evaluate the structural equation model in Fig. S2. We found that this model provided a good fit to the data (Fisher  $C = 11.61$ ,  $k = 6$ ,  $P = 0.071$ ). In particular, we found that genetic variation increased food-web complexity primarily by: (i) an increase in gall richness that directly increased complexity ( $0.40 \times 0.52 = 0.21$ ); and (ii) an increase in gall abundance that indirectly increased complexity by increasing gall vulnerability ( $0.57 \times 0.55 \times 0.83 = 0.26$ ). Interestingly, gall evenness had a small overall negative effect on complexity ( $(-0.18 \times 0.39) + (-0.18 \times -0.32 \times 0.83) + (-0.18 \times 0.25 \times 0.28) = -0.03$ ).



**Figure S1.** One of 50 replicate simulations, showing the positive relationship between willow genetic variation and food-web complexity. Grey circles represent estimates of food-web complexity for specific samples, whereas blue circles represent the average complexity at each level of genetic variation. These data were used in the structural equation model (Fig. S2).



**Figure S2.** Structural equation model of the paths by which genetic variation influences food-web complexity. Blue and red arrows indicate positive and negative relationships, respectively. One-way arrows indicate modelled paths, whereas double-headed arrows indicate correlated relationships. Numerical values in the middle of each path represent the standardized path coefficients and can be used to determine the magnitude of direct and indirect effects.

## References

1. Barbour MA et al. (2015) Multiple plant traits shape the genetic basis of herbivore community assembly. *Functional Ecology* 29:995–1006.
2. Villéger S, Mason N, Mouillot D (2008) New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology* 89:2290–2301.
3. Gotelli NJ, Colwell RK (2010) *Estimating species richness* (Oxford University Press, Oxford).
4. Bersier L-F, Banašek-Richter C, Cattin M-F (2002) Quantitative descriptors of food-web matrices. *Ecology* 83:2394–2407.
5. Colwell RK, Coddington J (1994) Estimating terrestrial biodiversity through extrapolation. *Phil Trans Roy Soc B* 345:101–118.



6. Bolker BM (2008) *Ecological Models and Data in R* (Princeton University Press, Princeton, New Jersey).
7. Tuomisto H (2012) An updated consumer's guide to evenness and related indices. *Oikos* 121:1203–1218.
8. Shipley B (2000) A New Inferential Test for Path Models Based on Directed Acyclic Graphs. *Structural Equation Modeling: A Multidisciplinary Journal* 7:206–218.