Ecological Applications

Biotic resistance across a nutrient gradient in experimental wetland mesocosms

Clementina Calvo, Kenneth J. Elgersma, Deborah E. Goldberg, William S. Currie, Evan Batzer & Jason P. Martina

Appendix S1

Additional information on the Methods

The RF models were trained using the morphological traits measured in the field and linear or allometric relationships of the genetically-identified ramets from the three *Typha* species. Based on exploratory graphs, the variables included in the best-fitted models for nonflowering ramets were: leaf width, stem height, basal diameter, dry weight, and stem volume, and the linear relationships of height vs. leaf width, height vs. basal diameter, leaf width vs. basal diameter, and dry weight vs stem volume. For flowering ramets, the variables included were leaf width, height, basal diameter, dry weight, inflorescence gap length, and the allometric relationships of height vs. leaf width, height vs. basal diameter, and leaf width vs. basal diameter. We analyzed flowering and nonflowering ramets separately because previous studies have shown that the allometry between leaf width and height or basal diameter differ between them (Snow *et al.,* 2010) and because additional morphological measurements were available for flowering ramets. For each category, we used 80% of the genetically-identified ramets to train the RF models, while the remaining 20% was used to test the models. The overall accuracy obtained for the nonflowering subset was 78%, while the flowering subset achieved an accuracy of 90%. The RF models were used to classify individually measured and weighed ramets of specimens classified on-site but not genetically identified (nonflowering, n=1145 and flowering, n=377). The classification performance for the three *Typha* species was assessed using confusion matrices, which provided the number of coincidences in the assignments to a class (Table S1). The RF models were developed using the *randomForest* package in R 4.2.2 (Liaw & Wiener, 2002).

The bulk collections of *Typha* comprised individuals categorized into the six field classifications but without individual measurements and then weighed. Before collection, the number of flowering and nonflowering individuals in each field category was counted. To reassign the taxon identity of the biomass of these ramet sets, we multiplied the total biomass of each field category per mesocosm by the biomass proportion of that category obtained from the RF predictions of the ramets with individual measurements but no genetic identification in each site. In other words, we assumed the biomass proportions of the bulk-weighed ramet sets in each mesocosm were the same as the proportions of the individually-weighed ramets from each site. The calculation included the ratio of the number of ramets with and without flowers to account for the higher biomass of flowering individuals. The proportions were calculated separately for each site to consider potential site-level differences in the abundance of the taxa as well as in field classifications resulting from variations in field crew members. The same approach was used to calculate and adjust the number of ramets for each species per site and mesocosm.

References

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R News, 2(3), 18-22.

Snow, A. A., Travis, S. E., Wildová, R., Fér, T., Sweeney, P. M., Marburger, J. E., ... & Mutegi, E. (2010). Species-specific SSR alleles for studies of hybrid cattails (*Typha latifolia* x *T. angustifolia*; *Typhaceae*) in North America. American Journal of Botany, 97(12), 2061-2067.

Figures

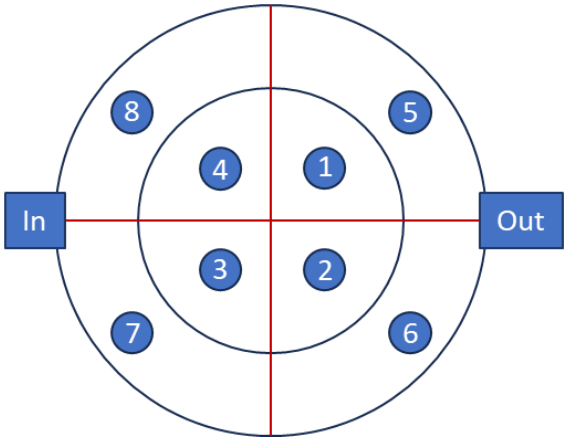


Figure S1. Scheme illustrating the distribution of the eight subplots (20 cm in diameter) within each mesocosm. Red lines indicate the location of the dowel rods. Inflow and outflow of water are specified.

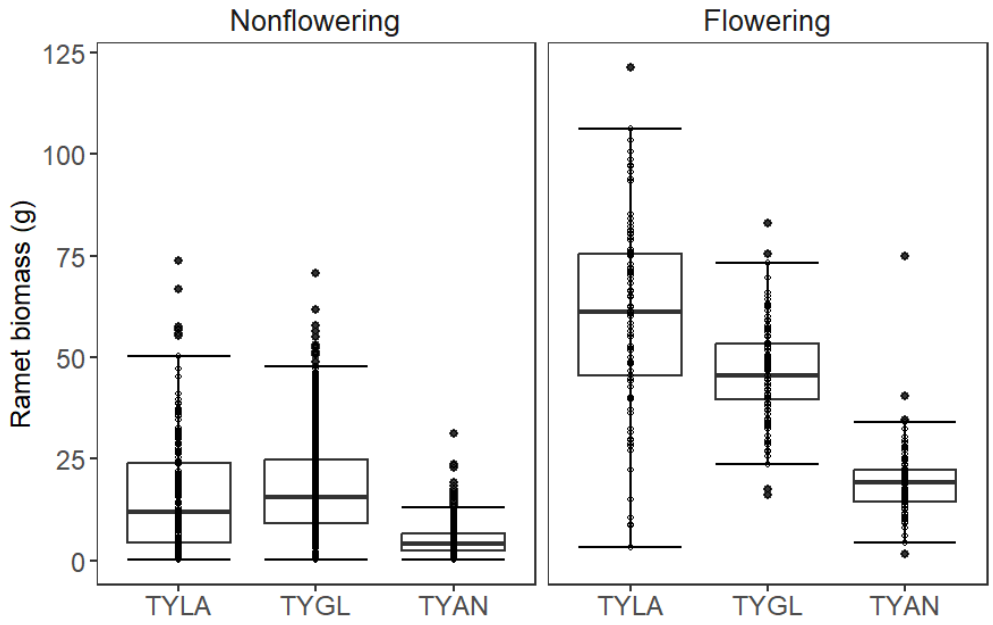


Figure S2. Biomass (g) of *Typha latifolia* (TYLA), *T. x glauca* (TYGL), and *T. angustifolia* (TYAN) nonflowering and flowering ramets. Median values, 0.25 and 0.75 quantiles (boxes), and minimum and maximum values (whiskers) are shown.

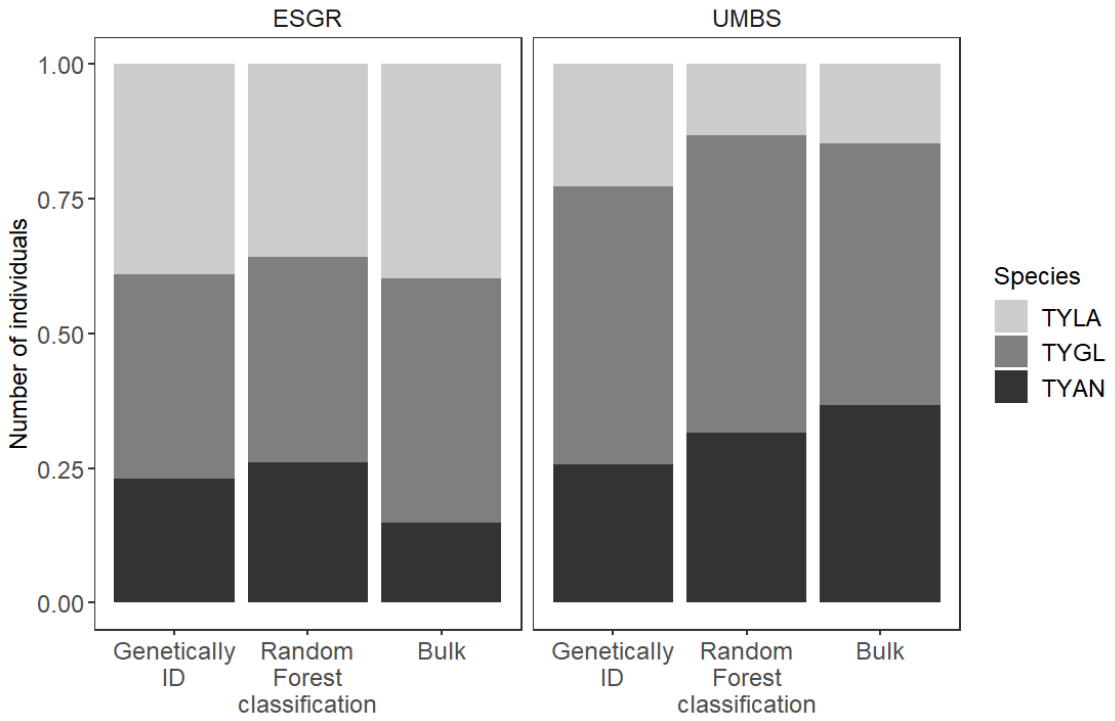


Figure S3. Proportion of individuals of *Typha latifolia* (TYLA), *T. x glauca* (TYGL), and *T. angustifolia* (TYAN) of Genetically ID individuals, ramets classified using random forest (Random Forest classification) and Bulk database, from both experimental sites (ESGR and UMBS).

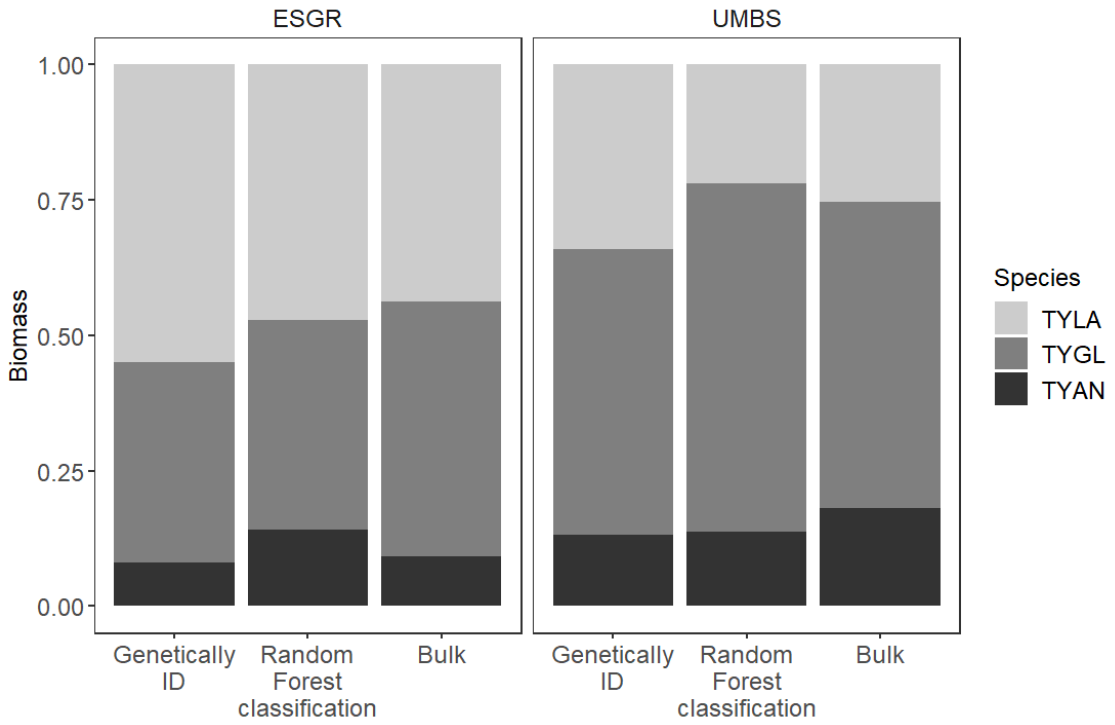


Figure S4. Proportion of biomass of *Typha latifolia* (TYLA), *T. x glauca* (TYGL), and *T. angustifolia* (TYAN) of Genetically ID individuals, ramets classified using random forest (Random Forest classification) and Bulk database, from both experimental sites (ESGR and UMBS).

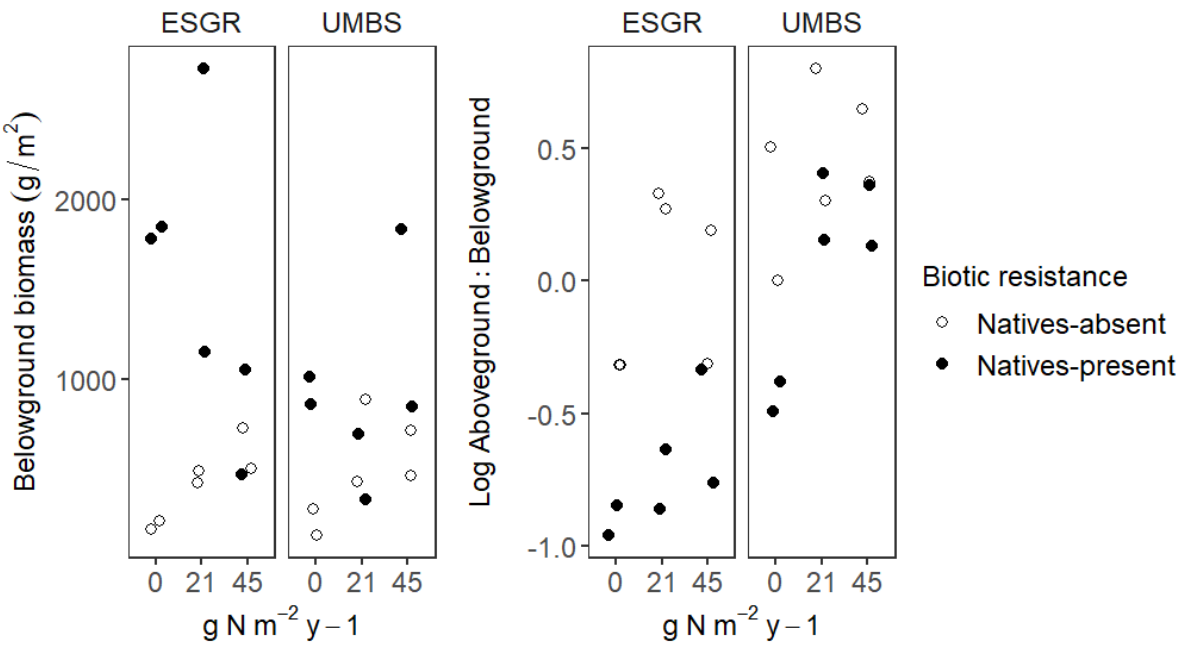


Figure S5. Belowground biomass (g/m^2) and Aboveground: Belowground biomass ratio of the two sites: ESGR and UMBS and for the two biotic resistance treatments: No residents (open circles) and Residents (filled circles), for 0, 21 and 45 g N m^{-2} y^{-1}.

Tables

Table S1. Differences between experimental sites ESGR and UMBS in the chemistry of the groundwater entering the mesocosms (previous to fertilization treatments), and the climate characteristics of the area.

	Water inflow							Climate	
	PO_4-P ug/L	NO_3-N ug/L	NH_4-N ug/L	Ca mg/L	Mg mg/L	Na mg/L	K mg/L	Mean annual temperature (ºC)	Mean annual precipitation (mm/year)
ESGR	21.34 ± 0.8	<1	<8	73.83 ± 0.71	19.24 ± 0.25	2.68 ± 0.13	0.74 ± 0.05	9.27	932
UMBS	<2	83.56 ± 2.6	<8	65.26 ± 2.70	19.75 ± 0.25	5.0 ± 0.10	0.77 ± 0.03	6.07	749

Table S2. Confusion matrix for the ramets classified on-site (6 categories: TYLA, TYLA?, TYGL, TYGL?, TYAN, and TYAN?) and the prediction for those ramets (3 categories: TYLA, TYGL, and TYAN) based on random forest models (RF), for both nonflowering and flowering ramets. Accuracy (%) is indicated for each case.

	RF predictions							
	Nonflowering				**Flowering**			
On-site classification	TYAN	TYGL	TYLA	Accuracy (%)	TYAN	TYGL	TYLA	Accuracy (%)
TYAN	291	125	21	66.6	112	6	1	94.1
TYAN?	28	109	23	17.5	1	10	4	6.7
TYGL	17	303	130	67.3	3	129	63	66.2
TYGL?	0	30	39	43.5	0	6	6	50
TYLA	0	0	1	100	0	0	7	100
TYLA?	0	2	6	75	0	0	15	100

Table S3. Response of total biomass per mesocosm (including both *Typha* and Residents when present) to explanatory variables from analysis of variance (ANOVA) output of generalized linear models for each site (GLM; family: gaussian, link: log). Only variables retained in each model after AIC selection are shown. Explanatory variables entered into each model include Biotic resistance, Nitrogen level, Nitrogen level-quadratic, Litter mass (g m^{-2}), and a subset of interactions. The percentage of deviance explained by the models is shown (D^2). ESGR and UMBS are the two study sites as in Fig. 1.

	ESGR (D^2=0.62)			**UMBS (D^2=0.75)**		
	Df	F	Pr(>F)	Df	F	Pr(>F)
Biotic resistance (BR)	1,41	19.298	**<0.001**	1,43	4.516	**0.039**
Nitrogen level (N)	1,41	14.307	**<0.001**	1,43	21.993	**<0.001**
(Nitrogen level)^2	1,41	5.646	**0.022**			
Litter (L)	1,41	1.329	0.257	1,43	8.491	**0.006**
BR*N	1,41	6.330	**0.016**			
N*L	1,41	5.609	**0.023**	1,43	19.767	**<0.001**

Note: p values <0.05 appear in bold.

Table S4. Response of *Typha* biomass to explanatory variables from analysis of variance (ANOVA). Explanatory variables: Biotic resistance, Nitrogen level, Nitrogen level-quadratic, and their interactions. ESGR and UMBS are the two study sites as in Fig. 1.

	ESGR			**UMBS**		
	Df	F	Pr(>F)	Df	F	Pr(>F)
Biotic resistance (BR)	1,42	78.525	**<0.001**	1,42	43.782	**<0.001**
Nitrogen level (N)	1,42	18.571	**<0.001**	1,42	55.233	**<0.001**
Nitrogen level 2 (N2)	1,42	7.196	**0.010**	1,42		
BR*N	1,42	14.023	**<0.001**	1.42		
BR*N2				1,42	9.457	**0.003**

Note: The p values <0.05 appear in bold.

Table S5. Response of *Typha* biomass to explanatory variables from analysis of variance (ANOVA) output of generalized linear models is shown (GLM, family: gaussian, link: log). Only variables retained in each model after AIC selection are shown. Explanatory variables entered into each model include Biotic resistance, Nitrogen level, Nitrogen level-quadratic, Resident biomass (g m^{-2}), Litter mass (g m^{-2}), and a subset of interactions. The percentage of deviance explained by the models is shown (D^2). ESGR and UMBS are the two study sites as in Fig. 1.

	ESGR (D^2=0.80)			**UMBS (D^2=0.78)**		
	Df	F	Pr(>F)	Df	F	Pr(>F)
Biotic resistance (BR)	1,40	9.807	**0.003**	1,39	0.534	0.469
Nitrogen level (N)	1,40	14.509	**<0.001**	1,39	8.603	**0.006**
Resident Biomass (RB)	1,40	0.085	0.772	1,39	6.7316	**0.013**
Litter mass (L)	1,40	5.029	**0.031**	1,39	4.983	**0.031**
BR*N				1.39	0.255	0.616
BR*L				1,39	1.597	0.214
N*RB	1,40	10.130	**0.003**	1,39	0.079	0.779
N*L	1,40	12.878	**<0.001**			
RB*L	1,40	12.217	**0.001**	1,39	1.741	0.195

Note: p values <0.05 appear in bold.

Table S6. Response of *Typha latifolia* (TYLA), *T. x glauca* (TYGL), and *T. angustifolia* (TYAN) biomass (g dry weight m^{-2}) to explanatory variables from analysis of variance (ANOVA) output of generalized linear models is shown (GLM, family: gaussian, link: log). Only variables retained in each model after AIC selection are shown. Explanatory variables entered into each model include Biotic resistance, Nitrogen level, Nitrogen level-quadratic, Resident biomass (g m^{-2}), Litter mass (g m^{-2}), and a subset of interactions. The percentage of deviance explained by the models is shown (D^2). ESGR and UMBS are the two study sites as in Fig. 1.

	ESGR			**UMBS**		
	Df	F	Pr(>F)	Df	F	Pr(>F)
	TYLA (D^2=0.72)			**TYLA (D^2=0.77)**		
Biotic resistance (BR)	1,40	13.796	**<0.001**	1,40	0.372	0.545
Nitrogen level (N)	1,40	16.209	**<0.001**	1,40	18.456	**<0.001**
Resident biomass (RB)				1,40	9.521	**0.004**
Litter (L)	1,40	0.073	0.788	1,40	0.181	0.673
N*RB	1,40	0.678	0.415	1,40	1.673	0.203
N*L	1,40	13.997	**<0.001**	1,40	37.733	**<0.001**
RB*L	1,40	13.399	**<0.001**	1,40	12.645	**<0.001**
	TYGL (D^2=0.65)			**TYGL (D^2=0.81)**		
Biotic resistance (BR)	1,43	55.416	**<0.001**	1,38	6.625	**0.014**
Nitrogen level (N)	1,43	6.258	**0.016**	1,38	2.169	0.149
Resident biomass (RB)				1,38	11.082	**0.002**
Litter	1,43	5.594	**0.023**	1,38	17.002	**<0.001**
BR*N				1,38	0.583	1.000
BR*L				1,38	0.862	0.359
N*RB				1,38	0.016	0.902
N*L	1,43	2.773	0.103	1,38	8.735	**0.005**
RB*L				1,38	0.842	0.365
	TYAN (D^2=0.26)			**TYAN (D^2=0.58)**		
Biotic resistance	1,42	13.174	**<0.001**			
Nitrogen level	1,42	2.123	0.153	1,44	25.536	**<0.001**
Nitrogen level 2	1,42	3.010	0.090	1,44	21.811	**<0.001**
Resident biomass				1,44	45.608	**<0.001**
Litter	1,42	0.637	0.429			
BR*L	1,42	0.000	0.992			

Note: p values <0.05 appear in bold.

Table S7. Response of Native biomass (g/m^2) to explanatory variables from analysis of variance (ANOVA) output of generalized linear models is shown (GLM, family: gaussian, link: log). Main effects of explanatory variables for both experimental sites are shown. Only variables retained in each model after AIC selection are shown. Explanatory variables: Nitrogen level, Nitrogen level-quadratic, *Typha* biomass (g/m^2), Litter mass (g/m^2), and a subset of interactions. The percentage of deviance explained by the models is shown (D^2).

	ESGR (D^2=0.51)			**UMBS (D^2=0.86)**		
	Df	F	Pr(>F)	Df	F	Pr(>F)
Nitrogen level (N)	1,18	10.093	**0.005**	1,19	44.762	**<0.001**
Nitrogen level 2	1,18	11.842	**0.003**	1,19	24.120	**<0.001**
Typha mass (T)	1,18	2.479	0.133	1,19	36.932	**<0.001**
Litter mass (L)	1,18	1.214	0.285	1,19	0.001	0.975
N*T				1,19	0.091	0.767
N*L				1,19	6.348	**0.022**
T*L	1,18	3.127	0.094			

Note: The p values <0.05 appear in bold.