*Statistical analyses*

All analyses were completed in R version 4.2.2. Three models were used to evaluate the outcomes of the experiment. First, we modeled the effects of aphid biotype, plant species, and virus status on the total number of aphids counted on a host plant after one week (“performance”). All models fitted to these count data followed a negative binomial distribution and were run using the *glm.nb* function in the MASS package in R (Venables and Ripley, 2002). Second, we modeled the effects of aphid biotype, plant species, and viral status on the number of aphids that moved towards a host plant (“preference). Third, we completed a linear model examining the relationship between preference and performance with an interaction term of virus status. In this third model, the predictor variable aphid counts were log-transformed to meet linear regression assumptions. Both the second and third models were normally distributed and used the *glm* function in base R.

For the first two models, we employed a stepwise model selection approach. We started with a fully specified generalized linear model (GLM) including the third-order interaction term between aphid biotype, plant species, and virus status, as well as all underlying second and first-order terms. Given the relatively small size of the dataset to this large model, we sought to reduce model complexity and avoid overfitting *via* stepwise regression. After running these fully specified models, the *stepAIC* function in the MASS package was applied to each model (Venables and Ripley, 2002). When applied to a GLM, this function sequentially reduces the number of parameters to produce a model with the lowest AIC. Such an approach allowed for consistent, unbiased strategy to avoid overfitting across both types of aphid count data. Both full and reduced models significance tests (P-values and critical values) were completed using the car package (Fox and Weisberg, 2018), while estimated marginal means and post-hoc tests *via* Tukey HSD were calculated using the emmeans package (Lenth, 2018).

References (Stats methods only)

R Development Core Team. 2022. R version 4.2.2 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>

Fox, J., & Weisberg S. 2018. Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. *Journal of Statistical Software.* 87:1–27.

Lenth, R. 2018. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.3.0. https://CRAN.R-project.org/package=emmeans

Venables, W.N. & Ripley, B.D. 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York.