

**Projeto de Pesquisa e Planejamento de Atividades**

<b>Aluno:</b> Davi de Souza Fernandes		<b>Data início curso:</b> _01_/_10_/_2022
<b>Orientador:</b> Gabrielle Maria Romeiro Lombardi		<b>Defesa em:</b> Maio/2024
<b>Curso:</b> MBA Data Science e Analytics	<b>Modalidade:</b> Distância	Turma: 222

**1. Título do projeto**

Modelagem genômica para predição de adaptação climática em percevejo-marrom *Euschistus heros* (Fabr. 1798).

**2. Introdução**

As mudanças climáticas têm sido foco de constantes preocupações por ameaçar a continuidade e integridade da biodiversidade e ecossistemas, remodelando a distribuição e ocorrência das espécies (IPCC, 2021). Apesar das ações humanas, alterações climáticas globais já ocorreram diversas vezes, exercendo uma pressão seletiva sobre as espécies, onde aquelas que não foram capazes de se adaptar, acabaram por ser eliminadas, enquanto aquelas que o fizeram, deixaram descendentes e se adaptaram (Avice, 2000; Hegerl et al., 2019). Desse modo, as recentes mudanças climáticas estão e irão exercer uma pressão seletiva sobre a biodiversidade (Halsh et al., 2020).

A adaptação é um processo biológico onde uma característica vantajosa e hereditária, ou seja, que beneficia a reprodução e/ou sobrevivência, é selecionada ao longo do tempo evolutivo em resposta à pressão seletiva (Freeland, 2011). Por outro lado, a evolução se caracteriza pela simples mudança nas frequências gênicas, em escala temporal ou espacial, podendo decorrer da seleção natural e/ou por meios estocásticos, como deriva genética (Lowe et al., 2017). Ambos os processos ocorrem sob uma população, um grupo de organismos da mesma espécie, com potencial de reproduzir e deixar descendentes férteis e que habitam um mesmo tempo e espaço (Freeland, 2011). Já a diversidade genética é o substrato para que os processos evolutivos, como a adaptação, ocorram, selecionando genes capazes de atribuir vantagens aos organismos na população (Dudaniec et al., 2018).

Os insetos são estritamente dependentes dos fatores climáticos, como temperatura, humidade, precipitação, entre outros. (Powell e Logan, 2005). O corpo diminuto, recoberto pelo exoesqueleto, o sistema respiratório traqueal, e a incapacidade de controlar a própria temperatura tornam os insetos dependentes das variáveis climáticas (McCulloch e Waters, 2022). Assim, essas variáveis impactam diretamente a sobrevivência, reprodução e distribuição dos insetos (Loraia et al., 2023; Powell e Logan, 2005). Apesar disso, estima-se que os insetos são o grupo de animais mais diversos e abundantes do planeta, desempenhando inúmeros benefícios ecológicos, como a polinização e ciclagem de nutrientes, por exemplo (Kim, 1993). Por outro lado, muitos insetos atuam como vetores de doenças ou como pragas agrícolas, causando prejuízos na escala de bilhões de dólares anuais. (Kim, 1993). Nos sistemas agrícolas, os insetos pragas encontram abundância de alimentos, e muitas vezes, condições ótimas para seu desenvolvimento (Redlish et al., 2018). Porém a ocorrência de uma

praga é determinada por fatores, bióticos e abióticos, sobretudo pelas características climáticas (Carrière et al., 2010; Haan et al., 2020). Complementarmente, interações biológicas, histórico de ocorrência e aspectos populacionais como a conectividade e diversidade existente, também interferem na distribuição dos insetos (Sexton et al., 2011; Tiffin e Ibarra, 2014).

O percevejo marrom, *Euschistus eros* (Fabr. 1798), é uma das principais pragas da soja no Brasil, se alimentando das vagens e causando danos diretos à cultura (Panizzi, 2015; Sosa-Gomez et al., 2020). Entretanto, até por volta de 1970 o percevejo ainda não figurava entre as pragas primárias da cultura (Soares et al., 2017). Esse fato evidencia a característica intrínseca de como a dinâmica das pragas agrícolas pode mudar ao longo do tempo. As populações de *E. heros* ocorrem ao longo de todo o país, estando sobre pressões de seleção heterogêneas e caracterizadas por grupos genéticos espacialmente estruturados, com variação na diversidade (Singh et al., 2023). Dessa maneira, estudar a composição genética das populações e relacioná-las aos diferentes gradientes climáticos onde ocorrem pode ser uma forma de prever a distribuição da população em um possível cenário de mudanças climáticas (Chen, et al., 2021; Jin et al., 2023).

O estudo da composição genética depende de técnicas e métodos moleculares que permitam o acesso à informação genômica (Lou et al., 2021; Uffelman et al., 2021). Esses métodos possibilitam a leitura e tradução da composição do DNA em informações computacionais passíveis de serem analisadas (Lou et al., 2021; Uffelman et al., 2021). A partir disso, é possível encontrar marcadores moleculares ou variantes que permitem a estimativa da diversidade e composição das populações (Lou et al., 2021; Uffelman et al., 2021). O método GBS “*genotyping by sequence*” é uma técnica de genotipagem por sequenciamento de nova geração. Esse método permite identificar variantes nucleotídicas, SNPs (*Single nucleotide polymorphisms*) existentes entre indivíduos das populações (Lou et al., 2021). Diferentes abordagens e análises podem ser adotadas para se estabelecer modelos preditivos com esses dados (Uffelman et al., 2021). Inferências bayesianas, análises não supervisionadas e supervisionadas são empregadas para estimar a diversidade e estruturação das populações, além de modelos linearmente generalizados, modelos mistos (*Latent Factor Mixed Models*), ou mesmo redes neurais que permitem encontrar as relações entre genótipos e as variáveis ambientais (Calla et al., 2020; Chen et al., 2021; e 2021; Lowe et al., 2017; Slatkin 1987; You et al., 2020).

Ao longo do tempo os genes de uma população tendem a entrar em equilíbrio, em que a frequência de um alelo (variante gênica), se mantém a mesma nas gerações seguintes, a menos que processos evolutivos de deriva genética, ou mecanismos de seleção estejam atuando (Calla et al., 2020; Chen et al., 2021; Péliissié et al., 2022). Por isso, é possível se correlacionar as variações na frequência gênica com os processos adaptativos (Dudaneic et al., 2018). Em algumas espécies de libélulas, alterações na frequência de genes ligados a regulação da temperatura corporal “*heat chock protein*” refletem as mudanças na distribuição geográfica em resposta às mudanças climáticas (Dudaniec et al., 2018). Para a traça das crucíferas, *Plutella xylostella*, a conquista de diversos continentes impôs forte pressão de seleção, especialmente nas populações de regiões Temperadas, em genes ligados a regulações fisiológicas (Chen et al., 2021). Como consequência, considerando o fluxo gênico e os genes adaptativos, é provável que as populações da traça irão se expandir para

além da distribuição atual, em um cenário futuro (Chen et al., 2021). Portanto, o entendimento dos fenômenos biológicos adaptativos depende estritamente do emprego de métodos de análise e visualização dos dados. Sendo assim, o objetivo com o presente trabalho será investigar regiões sob seleção no genoma de *E. heros* e correlacionar essas variações com características climáticas a fim de se criar um modelo preditivo adaptativo para as populações de *E. heros* no Brasil.

### 3. Objetivos

Procurar por variações nucleotídicas (SNPs) presentes no genoma de populações de *E. heros* e encontrar correlações entre as variáveis genômicas com as variáveis climáticas, permitindo o desenvolvimento de um modelo de predição para a adaptação climática dos percevejos em um cenário de aquecimento global futuro.

- 3.1. Encontrar as regiões e SNPs sob seleção nas populações de *E. heros*.
- 3.2. Confirmar a estruturação e diversidade heterogênea das populações de *E. heros* no Brasil.
- 3.3. Relacionar a presença dos SNPs sob seleção com as variáveis climáticas, com a formulação de um modelo de distribuição da espécie.
- 3.4. Construir um modelo de predição para a adaptação das populações a um possível aumento de temperatura futura, considerando a estrutura genética, os SNPs sob seleção e o modelo de distribuição da espécie.

### 4. Material e Métodos

#### Bancos de dados

Os dados utilizados no trabalho são sequências de DNA genômico, gerados por GBS (*genotyping by sequence*) pela plataforma de sequenciamento *Illumina HiSeq 2500*, extraídos de percevejos vivos, oriundos de diferentes regiões do Brasil onde a espécie ocorre. As sequências foram inicialmente tratadas (demultiplexadas) e limpas, removendo adaptadores, contaminação e sequências de baixa qualidade. As sequências foram então alinhadas com um genoma de referência de *E. heros* e filtradas de acordo com os critérios: frequência alélica mínima > 5%, heterozigosidade observada máxima de 75%, máximo de dados faltantes de 20%, máximo número de alelos igual a 2, e pelo menos 3x de cobertura. As sequências foram depositadas nos bancos de dados *National Center for Biotechnology Information* (NCBI), número de acesso PRJNA489772 sob o formato BAM (Singh et al., 2023).

#### Modelo de distribuição da espécie

Para a modelagem ecológica, as coordenadas geográficas, juntamente com dados de distribuição e ocorrência de *E. heros* no Brasil, publicamente disponíveis nos bancos de dados como *CABI* e *GBIF* serão agrupados em um mesmo banco de dados para a análise de distribuição. As variáveis bioclimáticas serão acessadas pelo *World Clim Bioclimate* e analisadas pelo algoritmo MaxEnt (Phillips, 2006). Os dados serão analisados nos softwares R Studio e QGIS. Assim, serão levantados os principais fatores ecoclimáticos que dirigem a distribuição das populações no Brasil (Aguirre-Liguori et al., 2019).

### **Análises populacionais e demográficas**

A estrutura populacional será calculada por meio do software ADMIXTURE -v1.23 (Alexander e Lang, 2011). Adicionalmente serão estimados os valores de *Fst* e Heterozigosidade *Hs* entre as diferentes populações encontradas (Frichot et al., 2013). Para isso, os genótipos serão subdivididos em janelas, e as comparações entre indivíduos e entre populações para o *Fst* levará por conta dessas janelas genômicas. As regiões com baixo *Fst*, em comparação às demais janelas, dentro de um mesmo genótipo, ou grupo, provavelmente indicam regiões sob pressão de seleção (Lou et al., 2021). Uma análise de componentes principais será realizada, com o pacote *ade4* em R.

### **Análise de seleção**

Inicialmente serão utilizados os métodos *Bayescenv* e *Bayescan* para separar os SNPs sob seleção, com os SNPs resultantes das variações demográficas, tomando por base o cálculo do *Fst* entre as populações (Coop et al., 2010; Villemereuil e Gaggiotti, 2015). Esses métodos utilizam estatísticas probabilísticas bayesianas e Markov Chain Monte Carlo, para inferir as probabilidades de um dado SNP estar sob seleção. Possivelmente serão testados a viabilidade de se utilizar os métodos complementares *OutofFlank* e *PCAdapt* (Jombart, 2008; Whitlock e Lotterhos, 2015). A diferenciação genética entre as populações fruto das variações que se comportem semelhantes ao esperado pelo *Fst*, são provavelmente associadas aos processos neutros de variabilidade, como isolamento por distância ou deriva genética (Coop et al., 2010; Villemereuil e Gaggiotti, 2015; Lou et al., 2021). Já a diferenciação genética independente dos fatores demográficos, por distância, possivelmente reflete SNPs sob seleção (Coop et al., 2010; Villemereuil e Gaggiotti, 2015; Lou et al., 2021).

### **Modelagem Ecológica**

A partir dos marcadores sob seleção um banco de dados com a intersecção dos resultados dos diferentes modelos, anteriormente mencionados, será criado. Em seguida, serão testados dois modelos diferentes para correlacionar os marcadores com as variáveis ambientais. O primeiro modelo será o método de aprendizado de máquina não supervisionado Random Forest, que terá como entrada os SNPs identificados sob seleção e as variáveis climáticas e terá como objetivo modelar a probabilidade de distribuição atual das espécies (Chen et al., 2021; Aguirre-Liguori et al., 2019 e 2021; You et al., 2020). A segunda abordagem vai envolver a construção de um modelo de fatores latentes LFMM. Ambas as análises objetivam criar um gradiente ambiental com o denominado *genetic turnover function* ou função de substituição genética, que compara a distribuição atual dos insetos, com a distribuição predita para os genes adaptativos. Quanto maior forem os valores do *genetic turnover* sob uma população, mais sensíveis elas são às mudanças ambientais (Chen et al., 2021; Aguirre-Liguori et al., 2019 e 2021; You et al., 2020). Complementarmente, uma última análise terá como entrada o gradiente gerado e a predição futura dentro do painel intergovernamental IPCC, que estima as mudanças previstas para os fatores climáticos e atmosféricos futuros. Estima-se que em um cenário mais otimista, com a redução dos gases de efeito estufa, haja um aumento na

temperatura média global de 1,5 °C até 2060. Por outro lado, esse aumento pode chegar até 2.5 dentro dos piores cenários, aumentando para mais de 4 °C até 2100. Sendo assim, o modelo final irá integrar os gradientes previstos e comparar os cenários presentes com o cenário de aumento de temperatura em 1.5° e 4°, considerando as mudanças climáticas e o *genetic offset* das populações (Chen et al., 2021).

## 5. Resultados Esperados

Espera-se que as populações de percevejos-marrom no Brasil apresentem a diversidade genética distribuída e com alta estruturação entre elas, como demonstrado por (Singh et al., 2023). Os fatores climáticos que provavelmente serão destacados como fatores limitantes à distribuição de *E. heros* estarão relacionados a umidade e temperatura. Assim, esperamos encontrar genes sob seleção que estejam relacionados à adaptação climática voltadas a esses dois fatores. Dessa forma, é provável que as populações com genes adaptativos associados a essas características apresentem um menor índice *genetic offset* e maior probabilidade de se manter sob as condições futuras. Por outro lado, populações mais isoladas, com menor diversidade e desprovidas de genes adaptativos, provavelmente apresentarão um maior *genetic offset* e possivelmente maior probabilidade de não se expandirem em um cenário futuro, dependendo do fluxo gênico e conectividade com outras populações fonte para levar alelos adaptativos.

## 6. Cronograma de Atividades

(Adicionar as “Atividades planejadas”, assim como o período (tempo para desenvolver cada atividade) planejado para a realização de cada atividade, sendo que deverá ser adequado ao calendário de entregas das etapas do trabalho de conclusão de curso definido pela Coordenação. Marcar com um “x” a coluna que corresponde ao período planejado para desenvolver cada atividade planeja)

[illegible]

<b>Apresentação</b>									<b>07</b>	
<b>Defesa</b>									<b>15</b>	
<b>Entrega versão final do TCC</b>										<b>14</b>
Projeto de Pesquisa; Resultados Preliminares; Entrega do Trabalho de Conclusão de Curso; Entrega da Apresentação da Defesa										

## 7. Referências Bibliográficas

(Listagem das bibliografias citadas no projeto de pesquisa, seguindo rigorosamente as Normas do MBA USP ESALQ – Consulte o manual de “Normas para Elaboração do Trabalho de Conclusão de Curso” disponível no Sistema TCC)

Aguirre-Liguori, J.A.; Ramírez-Barahona, S.; Tiffin, P.; Eguiarte, L.E. 2019. Climate change is predicted to disrupt patterns of local adaptation in wild and cultivated maize. *Proc Biol Sci.* 2019 Jul 10;286(1906):20190486. doi: 10.1098/rspb.2019.0486. PMID: 31290364; PMCID: PMC6650710.

Aguirre-Liguori, J.A.; Ramírez-Barahona, S.; Gaut, B.S. 2021. The evolutionary genomics of species' responses to climate change. *Nat Ecol Evol.* 5(10):1350-1360. doi: 10.1038/s41559-021-01526-9. Epub 2021 Aug 9. PMID: 34373621.

Alexander, D.H.; Lange, K. 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* 12, 246.

Avice, J. C. 2000. *Phylogeography the history and formation of species*. Cambridge, Mass. London: Cambridge, Mass. Harvard University Press London.

Calla, B. et al, 2020. Selective Sweeps in a Nutshell: The Genomic Footprint of Rapid Insecticide Resistance Evolution in the Almond Agroecosystem. *Genome Biology and Evolution*, v.13(1).

Carrière, Y.; Crowder, D.W.; Tabashnik, B.E. 2010. Evolutionary ecology of insect adaptation to Bt crops. *Evol*;3(5-6):561-73. doi: 10.1111/j.1752-4571.2010.00129.x. Epub 2010 Apr 30. PMID: 25567947; PMCID: PMC3352503.

Chen, Y.; Liu, Z.; Régnière, J. et al. 2021. Large-scale genome-wide study reveals climate adaptive variability in a cosmopolitan pest. *Nat Commun*, 12, 7206.

Coop, G.; Witonsky, D.; Di Rienzo, A.; Pritchard, J.K. 2010 Using environmental correlations to identify loci underlying local adaptation. *Genetics* 185, 1411 – 1423. (doi:10.1534/genetics.110.114819)

Dudaniec, R.Y.; Yong, C.J.; Lancaster, L.T. Svensson EI, Hansson B. 2018. Signatures of local adaptation along environmental gradients in a range-expanding damselfly (*Ischnura elegans*). *Mol Ecol.*, v. 27(11):2576-2593.

Freeland, J. *Molecular ecology*. 2nd ed., 1 ed. Oxford Hoboken, NJ: Oxford Wiley-Blackwell Hoboken, NJ, 2011.

Frichot, E., Schoville, S., Bouchard, G., & Francois, O. 2013. Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, 30(7), 1687– 1699. <https://doi.org/10.1093/molbev/mst063>

Haan, N.L.; Zhang, Y.; Landis, D.A. 2020. Predicting Landscape Configuration Effects on Agricultural Pest Suppression. *Trends Ecol Evol.* 35(2):175-186. doi: 10.1016/j.tree.2019.10.003. Epub 2019 Nov 5. PMID: 31699410.

Halsch, C.A.; Shapiro, A.M.; Fordyce, J.A.; Nice, C.C.; Thorne, J.H.; Waetjen, D.P.; Forister, M.L. 2021. Insects and recent climate change. *Proc Natl Acad Sci U S A.* 12;118(2): e2002543117. doi: 10.1073/pnas.2002543117. PMID: 33431560; PMCID: PMC7812774.

Hegerl, G.C.; et al 2019. Causes of climate change over the historical record. *Environ. Res. Lett.* 14 123006.



IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, In press, doi:10.1017/9781009157896.

Jin, M.; North, H.L.; Peng, Y.; Liu, H.; Liu, B.; Pan, R.; Zhou, Y., et al. 2023. Adaptive evolution to the natural and anthropogenic environment in a global invasive crop pest, the cotton bollworm. *Innovation (Camb)*. (4):100454. doi: 10.1016/j.xinn.2023.100454. PMID: 37388193; PMCID: PMC10300404.

Jombart T. 2008. Adegnet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403– 1405. (doi:10.1093/bioinformatics/btn129).

Kim, K.C. 1993. Biodiversity, conservation and inventory: why insects matter. *Biodivers Conserv* 2, 191–214. HYPERLINK "<https://doi.org/10.1007/BF00056668>"<https://doi.org/10.1007/BF00056668>.

Loaiza, J.R.; Bennett, K.L.; Miller, M.J.; León, L.F.D. Unraveling the genomic complexity of sylvatic mosquitoes in changing Neotropical environments. *Current Opinion in Biotechnology*, v. 81, 2023.

Lou, R.N.; Jacobs, A.; Wilder, A.P.; Therikildsen, N.O. 2021. A beginner's guide to low-coverage whole genome sequencing for population genomics. *Mol Ecol*. Dec;30(23):5966–5993. doi:10.1111/mec.16077.

Lowe, W. H.; Kovach, R. P.; Allendorf, F. W. Population genetics and demography unite ecology and evolution. *Trends in Ecology & Evolution*, v. 32, n. 2, p. 141–152, 2017.

McCulloch, G.A.; Waters, J.M. 2023. Rapid adaptation in a fast-changing world: Emerging insights from insect genomics. *Glob Chang Biol*. v. 29(4), pg. 943–954.

Panizzi, A. R. 2015. Growing problems with stink bugs (Hemiptera: Heteroptera: Pentatomidae): species invasive to the US and potential Neotropical invaders. *American Entomologist* 61, 223–233.

Pélissié, B.; Chen, YH.; Cohen, ZP.; Crossley, M.S. 2022. Genome Resequencing Reveals Rapid, Repeated Evolution in the Colorado Potato Beetle. *Mol Biol Evol*, v. 3, 39(2).

Phillips, S.J.; Anderson, R.P.; Schapire, R.E. 2006 Maximum entropy modeling of species geographic distributions. *Ecol. Model*. 190, 231– 259. (doi:10.1016/j.ecolmodel.2005.03.026).

Powell, Ja, logan ja. Insect seasonality: circle map analysis of temperature-driven life cycles. *Theor Popul Biol*., v. 67(3):161–79, 2005.

Redlich, S.; Martin, E.A.; Steffan-Dewenter, I. 2018. Landscape-level crop diversity benefits biological pest control. *J Appl Ecol*. 55:2419–2428.

Singh, K.S.; Cordeiro, E.M.G.; Hunt, B.J.; Pandit, A.A.; Soares, P.L.; Correa, A.S.; Zimmer, C.T.; Zucchi, M.I.; Batista, C. Dow, J.A.T.; Davies, S.A.; Cônsoli, F.L.; Omoto, C.; Bass, C. 2023. The genome sequence of the Neotropical brown stink bug, *Euschistus heros* provides insights into population structure, demographic history and signatures of adaptation. *Insect Biochem Mol Biol*. 152:103890. doi: 10.1016/j.ibmb.2022.103890. Epub 2022 Dec 8. PMID: 36496139.

Sexton, J.P.; Strauss, S.Y.; Rice, K.J. 2011. Gene flow increases fitness at the warm edge of a species' range. *Proc Natl Acad Sci U S A*. Jul 12;108(28):11704–9. doi: 10.1073/pnas.1100404108.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science*, v. 236, n. 4803, p. 787–792.

Soares, P.L.; Cordeiro, E.M.G.; Santos, F.N.S. et al. 2018. The reunion of two lineages of the Neotropical brown stink bug on soybean lands in the heart of Brazil. *Sci Rep* 8, 2496. HYPERLINK "<https://doi.org/10.1038/s41598-018-20187-6>"<https://doi.org/10.1038/s41598-018-20187-6>

Sosa-Gomez, D.R., Corrêa-Ferreira, B.S., Kraemer, B., Pasini, A., Husch, P.E., Delfino Vieira, C.E., Reis Martinez, C.B., Negrao Lopes, I.O., 2020. Prevalence, damage, management and insecticide resistance of stink bug populations (Hemiptera: Pentatomidae) in commodity crops. *Agric. For. Entomol.* 22, 99–118. <https://doi.org/10.1111/afe.12366>.

Tiffin, P.; Ross-Ibarra, J. 2014. Advances and limits of using population genetics to understand local adaptation, *Trends in Ecology & Evolution*, Volume 29, Issue 12, Pages 673-680.

Uffelmann, E.; Huang, Q.Q.; Munung, N.S. et al. 2021. Genome-wide association studies. *Nat Rev Methods Primers* 1, 59 (2021). <https://doi.org/10.1038/s43586-021-00056-9>.

Villemereuil, P.J.; Gaggiotti, O.E. 2015. A new  $F_{ST}$ -based method to uncover local adaptation using environmental variables. *Methods Ecol. Evol.* 6,1248–1258. (doi:10.1111/2041-210X.12418).

Whitlock, M. C., Lotterhos, K. E. 2015. Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of  $F_{ST}$ . *The American Naturalist*, 186, S24–S36. <https://doi.org/10.1086/682949>

You, M.; Ke, F.; You, S. et al. 2020. Variation among 532 genomes unveils the origin and evolutionary history of a global insect herbivore. *Nat Commun* 11, 2321. <https://doi.org/10.1038/s41467-020-16178-9>