GEONOME-WIDE MARKERS CAN IDENTIFY THE ORIGIN AND DEMOGRAPHY OF BOLL WEEVILS RE-INFESTATIONS: A CASE STUDY FROM KINGSVILLE, TX

Lindsey C. Perkin
Charles P.-C. Suh
USDA-ARS ICCDRU
College Station, TX
Raul Ruiz-Arce
USDA-APHIS PPQ
Edinburg, TX
Tyler J. Raszick
Gregory A. Sword
Texas A&M University
College Station, TX

Abstract

A boll weevil re-infestation was recently detected in the Kingsville area of Texas, but the factor(s) contributing to the re-infestation have not been identified. Nearly 250 weevils were obtained from pheromone traps in the re-infested area between July 2018 and February 2019. We used genome-wide SNP markers to determine the most likely geographical origin(s) of the weevils. We also used these same markers along with Coancesty software to estimate the relatedness among captured weevils to infer whether the re-infestation involved continual movement of new weevils into the area or if a few weevils moved into the area followed by local reproduction. Our results indicated the weevils captured in the Kingsville area originated from the Lower Rio Grande Valley (LRGV) production area. Given that none of the weevils appeared to be closely related, our findings suggest the re-infestation in Kingsville likely involved continual movement of weevils from the LRGV throughout the year.

Introduction

The boll weevil, *Anthonomus grandis grandis* Boheman is a destructive pest of commercial cotton throughout the Americas. It has been eradicated from most of the cotton producing areas of the U.S., however, active populations in the Lower Rio Grande Valley (LRGV), TX and Mexico still pose a threat to adjacent areas. Between July 30, 2018 and Feb. 4, 2019, 247 trap-captured weevils were obtained from a previously eradicated area near Kingsville, TX. The objective of our study was to use genome-wide single nucleotide polymorphism (SNPs) to better understand the demography of this infestation by addressing two main questions. First, where is the geographical origin of the captured weevils? Secondly, was the re-infestation due to a few weevils moving into the area followed by local reproduction or were there multiple waves of new weevils coming in throughout the growing and fallow seasons?

Materials and Methods

Weevils were collected weekly from traps near Kingsville, TX by the Texas Eradication Foundation between July 30, 2018 to February 4, 2019 (Figure 1). The location coordinates, field number, county, and date were recorded for each weevil and individuals were put in 100% ethanol and shipped to USDA-ARS, College Station, TX. Upon receipt, weevils were kept at -80°C until further analysis.

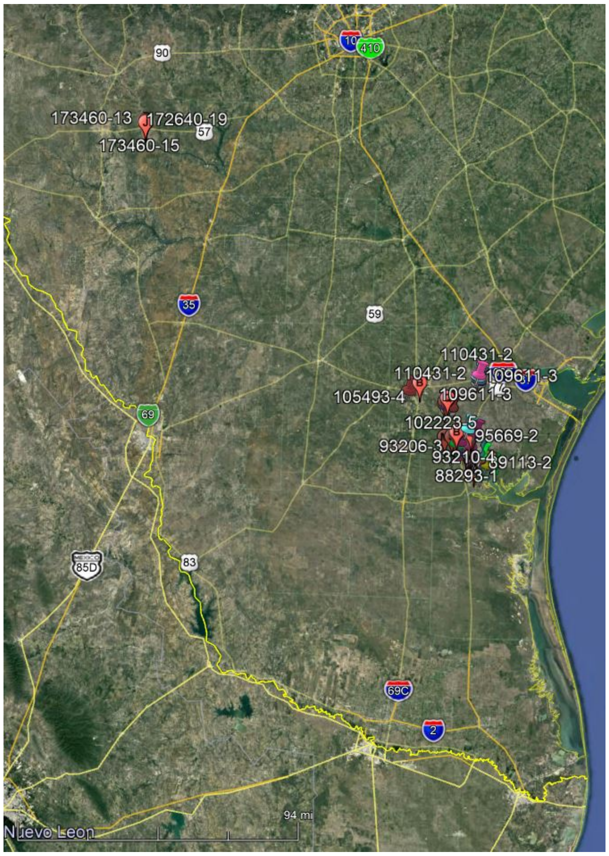

Figure 1. Location of captured weevils near Kingsville, TX.

DNA was extracted from each weevil with the QiaAmp kit (Qiagen, Hilden, Germany) with slight modifications to the extraction protocol. To minimize the shearing of DNA, wide-bore tips were used to mix samples instead of using a vortex. The quantity of DNA was maximized by following an overnight incubation at 56°C and adding carrier RNA as recommended by the manufacturer. All DNA samples were quantified on a Quantus fluorimeter system using the DNA ONE Fluorophore Kit (Promega, Madison, WI). All samples that had enough DNA (186 total) were shipped to the University of Minnesota Genomics Center for ddRAD-seq library preparation and sequencing via the Illumina NovaSeq.

SNPs were identified and filtered with Ddocent (Puritz et al., 2014) to only include biallelic loci present in all 186 weevils. SNPs were then analyzed using R/genepop version 1.0.5 (Rousset, 2008) to calculate F_{ST} values. A PCA plot was generated, which also included population SNP data of weevils obtained from Central Mexico, Western U.S. (AZ), and Eastern (LRGV and Tamaulipas, Mexico). Analyzing the data together allowed us to visualize how the Kingsville weevil population clustered relative to the other population clusters. SNPs were also analyzed with the program Coancestry (Wang, 2011) to determine the relatedness between weevils. Coancestry calculates a related estimate (r=0-1) between all possible paired comparisons of weevils, where r=0 means paired individuals are unrelated, r=0.25 is equivalent to first-cousins, r=0.5 is equivalent to siblings, and r=1 is a parent-offspring relationship.

Results and Discussion

All weevils from the Kingsville re-infestation clustered within the Eastern population cluster (Figure 2, pink), which includes the contiguous population from the Lower Rio Grande Valley (LRGV), TX and Tamaulipas, Mexico. These results indicate the weevils captured in Kingsville were genetically similar to weevils in the LRGV, suggesting the weevils captured in Kingsville likely originated from the LRGV.

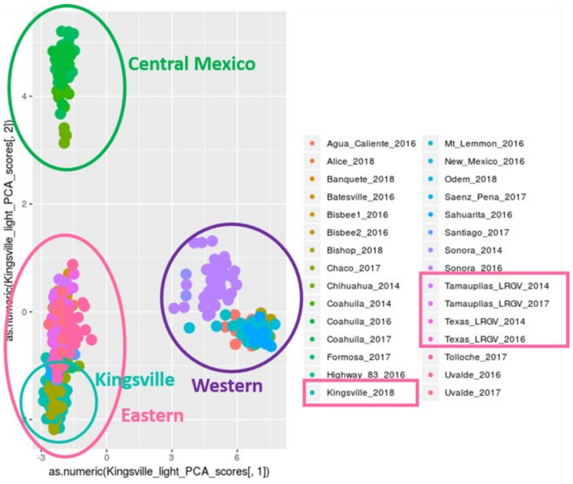

Figure 2. Geographical origin. PCA plot showing the three main population clusters, Central Mexico (green), Western U.S. (purple), and Eastern (pink). The Kingsville weevils (teal) clustered with the Eastern cluster.

A total of 186 individuals and 8,191 SNPs were available for the Coancenstry analysis. Out of the 4,137 possible paired weevil comparisons, 3,432 were unrelated (r=0; Figure 3). There were only 705 instances where weevils showed any level of relatedness (r = 0.0001 - 0.121), but the level of relatedness was equivalent to second cousins or less. Given that 83% of the paired comparisons indicated no relatedness among weevils, our finding supports the scenario where multiple waves of new weevils were coming from the LRGV area throughout the growing and fallow seasons.

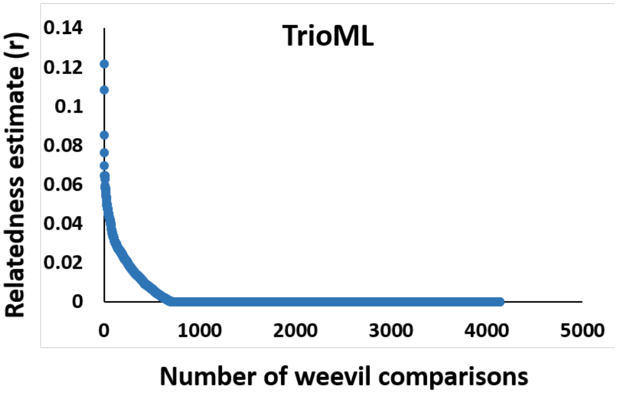

Figure 3. Coancestry results. Relatedness estimates (r) for 186 weevils captured in Kingsville, TX.

Summary

Our results indicated that the weevils captured in Kingsville, TX were likely from the LRGV, TX and Tamaulipas, Mexico region. We also found that these weevils are not related to each other (r = 0) or are very distant relatives, suggesting that multiple waves of weevils from the LRGV were coming into the Kingsville area throughout the year.

Acknowledgements

We thank Jolie Starling for help with DNA extraction, quantification, and data entry. This research was funded through the ARS Innovation Fund. Mention of trade names or commercial products does not imply recommendation or endorsement by the USDA. USDA is an equal opportunity provider and employer.

References

Puritz, J.B., C.M. Hollenbeck, and J.R. Gold. 2014. dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ 2:e431.

Rousset, F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Molecular Ecololgy Resources 8(1): 103-106.

Wang, J. 2011. COANCESTRY: a program for simulating, estimating and analyzing relatedness and inbreeding coefficients. Molecular Ecology 11:141-145.