Diatraea saccharalis history of colonization in the Americas. The case for human-mediated dispersal

Fabricio J. B. Francischini[1], **Erick M. G. Cordeiro**[2], **Jaqueline B. de Campos**[1], **Alessandro Alves-Pereira**[3], **João Paulo Gomes Viana**[1], **Xing Wu**[4], **Wei Wei**[4], **Patrick Brown**[4], **Andrea Joyce**[5], **Gabriela Murua**[6], **Sofia Fogliata**[6], **Steven J. Clough**[4,7], **Maria I. Zucchi**[1,8]*

1 Graduate Program in Genetics and Molecular Biology, Institute of Biology, University of Campinas, Campinas, Brazil, **2** Department of Entomology and Acarology, University of Sao Paulo, Luiz de Queiroz College of Agriculture (USP/ESALQ), Piracicaba, São Paulo, Brazil, **3** Department of Plant Biology, Institute of Biology, University of Campinas, Campinas, Brazil, **4** Department of Crop Science, University of Illinois, Urbana, Illinois, United States of America, **5** Department of Public Health, University of California, Merced, California, United States of America, **6** Instituto de Tecnología Agroindustrial del Noroeste Argentino, Estación Experimental Agroindustrial Obispo Colombres, Consejo Nacional de Investigaciones Científicas y Técnicas (ITANOA-EEAOC-CONICET), Tucumán, Argentina, **7** US Department of Agriculture-Agricultural Research Service, Urbana, Illinois, United States of America, **8** Laboratory of Conservation Genetics and Genomics, Agribusiness Technological Development of São Paulo (APTA), Piracicaba, São Paulo, Brazil

* mizucchi@apta.sp.gov.br

Abstract

The sugarcane borer moth, *Diatraea saccharalis*, is one of the most important pests of sugarcane and maize crops in the Western Hemisphere. The pest is widespread throughout South and Central America, the Caribbean region and the southern United States. One of the most intriguing features of *D. saccharalis* population dynamics is the high rate of range expansion reported in recent years. To shed light on the history of colonization of *D. saccharalis*, we investigated the genetic structure and diversity in American populations using single nucleotide polymorphism (SNPs) markers throughout the genome and sequences of the mitochondrial gene cytochrome oxidase (COI). Our primary goal was to propose possible dispersal routes from the putative center of origin that can explain the spatial pattern of genetic diversity. Our findings showed a clear correspondence between genetic structure and the geographical distributions of this pest insect on the American continents. The clustering analyses indicated three distinct groups: one composed of Brazilian populations, a second group composed of populations from El Salvador, Mexico, Texas and Louisiana and a third group composed of the Florida population. The predicted time of divergence predates the agriculture expansion period, but the pattern of distribution of haplotype diversity suggests that human-mediated movement was most likely the factor responsible for the widespread distribution in the Americas. The study of the early history of *D. saccharalis* promotes a better understanding of range expansion, the history of invasion, and demographic patterns of pest populations in the Americas.

Introduction

The ability to move from deteriorated environments to more resourceful ones is a fundamental component of the insects' life history and has a significant impact on agricultural ecosystems [1]. As colonizing insects arrive at a new location, both neutral (i.e., bottleneck and founder effect) and adaptive variations (i.e., natural selection) will cause genetic divergence to occur between the newly founded population and the ancestral one. In the case of natural selection, selection pressures imposed by host plants and by the environment promote adaptive evolution upon the insects' morphological, physiological, and behavioral attributes, resulting in divergence of settlers. [2,3]. In other words, the alleles favored by natural selection gradually replace those which are unfavored, changing populations from within [4–6]. The study of genome-wide single nucleotide polymorphisms (SNPs) can be useful to detect traces of selection sweeps and population divergence, and thus be used to address important ecological and evolutionary issues [4–6]. For instance, the careful investigation of the patterns of genetic diversity can contribute to studies on the early demographic history of the pest insects, including past movement and adaptive changes. The mechanism of insect dispersion can be divided into two broad categories: (1) active movement (i.e., migratory and dispersal behavior) [7] and passive movement [8]. Passive movement can be further divided into human-mediated [9–11] and non-human mediated movement—passively transported by boats, trucks or trains in the former, or by wind or water in the latter [12,13]. In both cases of passive movement, the direction of the displacement is biased and correlates with features of the natural environment or human activities.

Substantial advancement has been made in recent years to understand patterns of insect movement [14]. Such studies can be particularly useful in the context of pest insects [15–18], where the information about the origin of infestations and the patterns of dispersal have a direct impact on pest management strategies, including insecticide resistance management. *Diatraea saccharalis* (Fabricius) (Lepidoptera: Crambidae) is one of the most critical sugarcane and maize pests in the Western Hemisphere [19]. This insect is responsible for considerable crop damage and economic losses and is widely distributed throughout South and Central America, the Caribbean region, and the southern United States [20–23]. Although studies have addressed the distribution of this pest, the underlying details of the history of colonization in the Americas remain mostly unknown. The putative center of origin of *D. saccharalis* seems to be floating grasses (*Paspalum* spp. and *Panicum* spp.) along the delta of the Orinoco river in Venezuela [19]. However, while this is plausible, this hypothesis is still mostly uncorroborated.

An intriguing aspect of the life history of *D. saccharalis* is its low dispersal capacity and site fidelity that is suggestive of a more sedentary lifestyle in which matings are mostly restricted to small groups of related insects. Mark-recapture studies recovered over 45% of adults at around 50 meters from the release point, demonstrating the limited dispersal ability of this moth [24,25]. The dispersal distance was slightly increased to around 800 meters when the moths dispersal was aided by the wind. Recent studies have also investigated population structure and gene flow of *D. saccharalis* and detected significant genetic structure in samples from North, Central and South America [21,23,26,27]. However, these studies relied on a limited number of genetic markers (i.e., mitochondrial and SSR markers) and therefore could not directly reveal the dispersion patterns of *D. saccharalis* across the Western Hemisphere [26,28]. To summarize the problem in one question: how can an insect with such a limited dispersal ability have occupied large portions of the American continents?

Our primary goal in this study was to investigate the genetic structure and diversity in populations of *D. saccharalis* in the Americas using SNPs identified through the genotyping-by-sequencing (GBS) method [29]. This approach can be used to analyze differences in frequencies of neutral and adaptive alleles among populations and to provide a powerful tool to

investigate demographic history as well as to identify markers putatively under selection. Such studies give a framework for understanding the evolutionary patterns of the introduced invasive populations of *D. saccharalis* in the Americas.

Material and methods

Insect collections

A total of 250 specimens of *D. saccharalis* were collected in maize and sugarcane production regions in Brazil, Argentina, El Salvador, and the United States during the 2011/2012 and 2012/2013 crop seasons (S1 Table). Upon collection, the individual larvae were kept in Petri dishes containing artificial diet and maintained in controlled condition until pupation (27 ±1°C; 70% of relative humidity, and a photoperiod of 12D:12L). Healthy pupae were then transferred to cylindrical cages (40 x 30 cm) and kept in solitary conditions at 20±1°C until the adults emerged. Species identification was confirmed by the direct examination of the adults' genitalia, and only *D. saccharalis* were used in this study [30,31]. Moths were then stored at -80°C until further sample processing.

DNA extraction

The DNA was extracted from the adult moth tissues following the CTAB protocol described by Doyle and Doyle (1990) [32]. The integrity of DNA was evaluated in 0.8% (w/v) agarose gels with 1X TAE buffer (Tris, acetic acid, EDTA, pH 8.0), by concentrations estimated by comparison with known amounts of a DNA standard (λ phage). The gels were stained in an ethidium bromide bath (0.5 mg/mL), and DNA bands were visualized and photographed under UV light.

Genotyping-by-sequencing

GBS libraries were produced as described by Poland et al. (2012) [33]. DNA concentrations were determined using Picogreen (Molecular Probes, Eugene, Oregon) and a Synergy HT (BioTek, Winooski, Vermont) microplate reader, and adjusted to approximately 50 ng/µl. Five microliters (~250 ng) were pipetted into 96-well plates containing 2.5 µl 0.1 µM specific DNA barcoded *PstI* adaptors. The restriction enzymes *PstI* (New England Biolabs, Ipswich, MA, USA) and *MspI* (New England Biolabs) were used to digest the DNA and to reduce genome complexity. The barcoded *PstI* adapters and a common non-barcoded *MspI* adapter were ligated to the digested DNA and amplified by PCR to create an enriched library. The resulting library was single-end sequenced to 100 bases on a single lane using the Illumina HiSeq 4000 sequencing kit v1 (Illumina, Inc., San Diego, CA, USA). The fastq files obtained were demultiplexed with the bcl2fastq v2.17.1.14 (Illumina) by the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign producing nearly 404 million raw reads. The UNEAK pipeline, a multi-sample, SNP-calling approach developed for analyzing GBS data from species without reference genomes [34], was used to analyze the first 64 bases (beginning with the *PstI* restriction site) to identify SNPs.

TASSEL 5 [35] was used to filter SNP data such that each marker was present in at least 70% of the individuals within each population. Using this filtering, we identified 1,331 SNPs for the 250 individuals included in our GBS library.

Population genomics analyses

File conversion to other population genetics programs was made using PGDSpider v.2.1.0.3 [36]. To identify candidate loci that may have been under selection during the range expansion

of *D. saccharalis* in the Americas, the LOSITAN program [37], which employs the FDIST2 algorithm [38], was used to detect SNPs that are F_{ST} outliers. This method evaluates the relationship between the expected distribution of F_{ST} and heterozygosity assuming an island model of migration. In LOSITAN analysis, 100,000 simulations were run by using the stepwise mutation model with the option of neutral mean F_{ST}. Markers that presented F_{ST} higher than the 95% confidence interval were considered candidates for positive selection, and markers that presented F_{ST} lower than the 95% confidence interval were considered candidates for balancing selection. Genetic diversity and F statistics were estimated under a random model, in which the sampled populations were considered representative of the species and with a common evolutionary history. Allele frequencies, the number of alleles (A), allelic richness (A_R) the observed heterozygosity (H_O) and expected heterozygosity (H_E) and the inbreeding coefficient (F_{IS}) were estimated using the R package *diveRsity* [39]. Cluster analysis was based on the construction of dendrograms using Nei's genetic distance [40], and the neighbor-joining method was performed using the R package *poppr* [41]. The stability of the clusters was tested through 1,000 bootstraps resamples. The genetic structure was investigated employing the non-model-based approach DAPC using the R package *adegenet* [42]. The hierarchical distribution of genomic diversity within and among groups of populations was investigated with Analyses of Molecular Variance (AMOVA) also using the package *poppr*.

Phylogenetic analysis and divergence time estimation

A total of 167 sequences of a fragment of the cytochrome oxidase subunit I (COI) were retrieved from the BOLD repository (http://www.boldsystems.org/) to offer additional support for the study of the early demographic history of *D. saccharalis* in the Americas (S2 Table). DNA sequences were trimmed (533nt) and aligned with MUSCLE in MEGA X [43]. Unique haplotypes were identified in DnaSP software [44], and a phylogenetic analysis based on haplotypes was performed using maximum likelihood (ML) under the HKY+G model was conducted in MEGA X. A total of 1000 bootstraps were used for node support.

Tests for substitution model and molecular clock were also performed in MEGA X (S3 Table). Divergence time estimation was inferred by a Bayesian method implemented in BEAST 1.7.4 [45]. Divergence time calculation has been extensively used for the mitochondrial gene COI; however, the selection of realistic divergence rates is one of the main challenges of such studies. Lack of fossils or a species-specific diverge rate are factors responsible for inconsistencies in the estimation [46]. To overcome those obstacles, we used three different divergence rates frequently used for arthropods. The first rate is the standard rate of 2.3% divergence/My generally used for mitochondria of insects [47]. However, the recent literature suggests examples for slower [48] and faster raters [49,50]. Thus, a second rate of 3.54% [51] and a third rate of 5.2% [52] pairwise divergence per Mya was also used. HKY+G was implemented for the substitution model. The clock model was uncorrelated lognormal and constant size coalescent model for the tree prior. The analysis was run for 50 million generations sampled every 1000 generations in three independent repetitions. A burn-in of 20% was implanted over the total number of trees generated.

Results and discussion

GBS sequencing and detection of outlier SNPs

GBS sequencing data returned over 403 million sequence reads, approximately 1.6 million reads per sample. After cleaning and filtering the sequence data, aligning PCR fragments to reveal SNPs, a total of 1331 SNPs were identified to be used in subsequent population genetic

analyses for 250 individuals. The F_{ST}-outlier analysis showed 125 putative loci under positive selection and 270 under balancing selection (S1 Fig).

The analysis based on outlier SNPs was designed to detect genetic variation putatively driven by natural selection. Populations with shared demographic history should result in similar F_{ST}/H_E values for all loci, but those that deviate from the null model are therefore candidates for loci under selection [53]. Here, we present evidence that natural selection is acting upon population differentiation; it is not clear, however, which selection pressure prevails, whether it may be environmental, crop management or host related adaptions.

Genomic diversity

Overall diversity indexes [i.e., number of alleles (A), allelic richness (A_R), and observed heterozygosity (H_O)] showed a similar trend; that is, the highest values of genetic diversity were found in Brazil and Argentina whereas the lowest values were found in the U.S. and El Salvador (Table 1). The average values of observed (H_O) and expected heterozygosity (H_E) across all populations were 0.12 and 0.16, respectively. The Brazilian population of Jaboticabal-Sugarcane showed the highest observed and expected values ($H_O = H_E = 0.24$) (Table 1). The lowest value for the observed and expected heterozygosity was found in the U.S. population of Belle Glade-Sugarcane ($H_O = 0.03$ and $H_E = 0.08$).

Table 1. *Diatraea saccharalis* genomic diversity estimates based on 1,331 SNP loci by sampling locations.

Location	N	A	A_R	H_O	H_E	F_{IS}
LaCocha_Co	10	1859	1.13	0.1	0.13	0.244
LaCruz_Su	5	1721	1.1	0.1	0.12	0.159
Jujuy_Su	7	1866	1.13	0.11	0.14	0.222
Perga_Co	2	1519	1.05	0.12	0.14	0.183
Qui_Co	4	1552	0.98	0.07	0.13	0.475
Araras_Su	2	1723	1.14	0.16	0.16	0.02
Goias_TBD	4	1856	1.1	0.12	0.17	0.324
Jabo_Su	52	2333	1.41	0.24	0.24	0.006
Morr_Co	4	1910	1.16	0.13	0.17	0.243
MS_TBD	9	2129	1.23	0.14	0.2	0.306
MT_TBD	5	1818	1.12	0.11	0.15	0.225
PAf_Su	3	1751	1.12	0.11	0.15	0.253
Parana_TBD	3	1847	1.16	0.14	0.16	0.175
Pira_Co	25	2118	1.32	0.19	0.19	-0.009
Pira_Su	22	2296	1.38	0.22	0.23	0.046
Rib_Su	4	1916	1.11	0.14	0.18	0.246
Rondo_Co	3	1634	1.03	0.11	0.16	0.347
SHG_Su	4	1916	1.15	0.15	0.19	0.236
SP_TBD	18	2264	1.21	0.13	0.23	0.423
ElNilo_Su	4	1438	1.01	0.05	0.07	0.326
ElPais_Su	2	1277	0.89	0.04	0.13	0.727
BGlade_Su	9	1384	0.95	0.03	0.08	0.629
Louis_Su	15	1526	1.03	0.06	0.09	0.33
Beaum_Su	13	1556	1.07	0.08	0.09	0.149
Wesl_Su	7	1472	1	0.06	0.09	0.361

N = Number of individuals, A = Number of alleles, A_R = Allelic Richness, H_O = Observed heterozygosity, H_E = Expected heterozygosity, F_{IS} = Inbreeding coefficient

https://doi.org/10.1371/journal.pone.0220031.t001

The observed values for heterozygosity were less than the expected values in all but the three Brazilian locations (Jaboticabal-Sugarcane, Araras-Sugarcane, and Piracicaba-Maize). Particularly large differences between observed and expected heterozygosity were observed in populations from Argentina and El Salvador (i.e., Quilmes-Maize and El Pais-Sugarcane). Similar values for H_0 and H_E might indicate that most loci are in Hardy-Weinberg equilibrium, which means that the frequency of the alleles tended to remain unchanged over the generations. On the other hand, differences in observed and expected heterozygosity might indicate critical demographic changes.

Overall fewer heterozygotes than expected were found in populations of *D. saccharalis*, which was also reflected in the inbreeding coefficient (F_{IS}). The average inbreeding coefficient across populations was 0.27, ranging from -0.009 (Piracicaba-Maize) to 0.727 (El Pais-Sugarcane). Departure from the Hardy-Weinberg equilibrium (HWE), as shown previously, and heterozygote deficiency (F_{IS}), can be related to different causes. Limited dispersal, well-structured populations, and the evidence for strong positive selection are all factors that point to a more locally confined population dynamic.

Genetic structure and substructure of *D. saccharalis* populations

The values of pairwise F_{ST} revealed strong genetic structure between the northern (i.e., U.S. and El Salvador) and the southern (i.e., Brazil and Argentina) parts of the American continents (Fig 1, S4 Table). The highest F_{ST} values were found between the U.S. and the Argentinian populations (Louisiana vs. Buenos Aires, F_{ST} = 0.418). In contrast, an overall low degree of differentiation was found among the Brazilian populations (São Paulo *vs.* Paraná, F_{ST} = 0.003; São Paulo *vs.* Minas Gerais, F_{ST} = 0.004), which can be explained by the short geographical distance and therefore higher gene flow among sampled sites. Alternatively, the differences in genetic structure found in Brazil and the U.S. may be the result of a different pattern of colonization of the two geographical regions. This hypothesis will be developed further in the following paragraphs.

Similarly, the neighbor-joining dendrogram analysis revealed distinct geographically separated groups distinguishing samples from Brazil, Argentina, the U.S. and El Salvador (Fig 2). Interestingly, moths from the north and central-west of Brazil were the most related to the Argentina samples indicating similar founding sources or a connecting route for gene flow. Showing a slightly more significant number of groups, the discriminant analysis of principal components (DAPC) also separates *D. saccharalis* samples into four clusters according to the country of origin (Fig 3). The analysis of molecular variance (AMOVAs) supports the structure by geographical regions in which most of the genetic variation was assigned to differences among countries (φst = 0.525, Table 2). The individuals from Argentina and Brazil were more related to each other compared to other locations, confirming that the founding of the two populations is, at least in part, due to related events (i.e., stepping-stone or range expansion from similar source populations) (Fig 2). On the other hand, individuals from Florida were more distant from other samples collected in the United States (i.e., Louisiana and Texas), which could be attributed to different sources of the ancestral founding population (Figs 2 and 3). Our finding supports the hypothesis of the colonization of North America by multiple events from different sources, which agrees with Joyce et al. (2014) [26].

Host plants such as maize and sugarcane seem to be a less critical factor structuring *D. saccharalis* populations (Fig 3, Table 2). This result agrees with the work of Joyce et al. (2016) [23], which did not find significant genetic divergence among populations of *D. saccharalis* collected on maize and sugarcane in El Salvador. Our data indicate that *D. saccharalis*
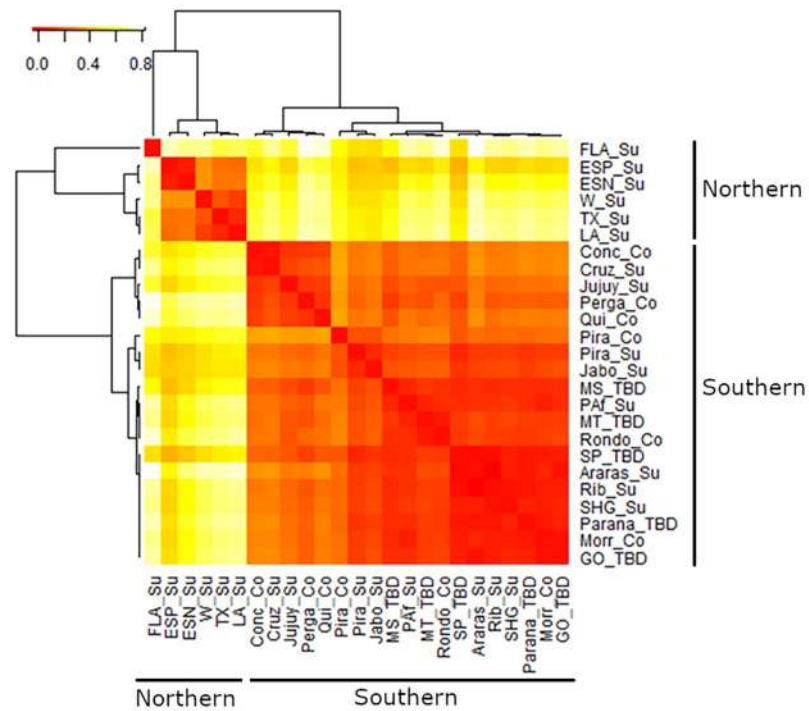
Fig 1. Heatmap of pairwise F$_{ST}$ values between *Diatraea saccharalis* populations, based on 1,331 SNP loci. Dendrograms were plotted using the unweighted-pair-group method with arithmetic mean (UPGMA). Colors indicate the degree of divergence from red (low F$_{ST}$ values, little genetic divergence) to light yellow (high F$_{ST}$ values, strong genetic divergence).

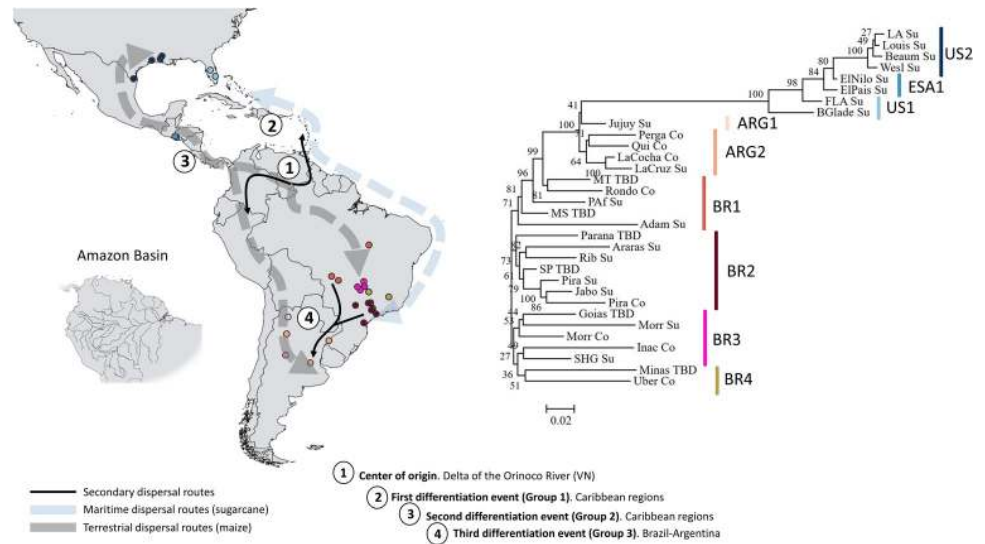https://doi.org/10.1371/journal.pone.0220031.g001

Fig 2. Suggested *D. saccharalis* dispersion routes from its center of origin in the Orinoco Basin, Venezuela, based on historical information about maize dispersal and sugarcane introduction, and the genetic relationships observed in this study. The neighbor-joining dendogram showing the relationships among *D. saccharalis* populations based on Nei's genetic distance. A total of 1,331 SNP loci were used, and confidence of nodes was based on 1,000 bootstraps. On the map, numbers indicate major differentiation events and arrows represent putative routes for dispersal.

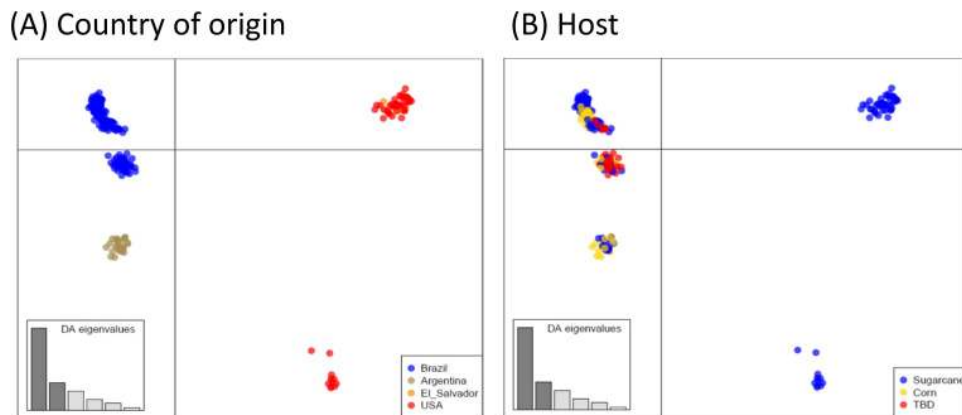https://doi.org/10.1371/journal.pone.0220031.g002

Fig 3. Scatterplots from discriminant analysis of principal components (DAPC) based on 1,331 SNP loci showing the dispersion of 250 *Diatraea saccharalis* individuals across the first two principal components. Individuals (dots) are colored according to (A) their country of origin; and (B) the host crop in which they were collected. Variation represented in x = 44.4% and in y = 15%.

https://doi.org/10.1371/journal.pone.0220031.g003

populations are undifferentiated in exploiting different hosts, suggesting some level of phenotypic plasticity and can also explain the high number of markers under balancing selection [54].

Divergence time estimation

Tree topology generated using the ML approach based on *D. saccharalis* COI sequences was strongly supported by nuclear markers (S2 Fig). The ML tree divided the data into three distinct groups according to geographic regions. The first split within *D. saccharalis* separated Florida from all other populations. A second split separated populations from El Salvador, Mexico, Texas, and Louisiana from those in South America (Brazil and Argentina) indicating a second event of differentiation (S2 Fig). The topology generated from the divergence time analysis based on Bayesian approach revealed a similar pattern of population differentiation. The three different divergence rates used to calibrate the molecular clock rendered different time estimations; however, because there is no specific COI clock rate or fossils records available for *D. saccharalis*, all three values were presented (Table 3, S3 Fig). All estimated divergence times place the separation of the three major groups in the period between 427 thousand

Table 2. Analyses of Molecular Variance (AMOVA) of *Diatraea saccharalis* populations based on the 1,331 SNP markers. The analysis considered different host-plant species and geographic regions. φ statistics are indices of the amount of differentiation among populations, similar to Wright's F statistics. d.f. = degrees of freedom.

Source of Variation	D.f	S.S.	M.S.	Sigma	% of variation	φ
Among countries	3	39716.87	13238.96	329.12	52.59	φst = 0.525**
Within countries	192	56972.76	296.73	296.73	47.41	
Total	195	96689.62	495.84	625.86		
Between hosts	1	5052.01	5052.01	66.05	12.27	φst = 0.122**
Within hosts	194	91637.61	472.36	472.36	87.73	
Total	195	96689.63	495.84	538.41		
Among countries	3	39716.87	13238.96	290.58	46.51	φst = 0.549**
Between host within countries	2	3518.63	1759.32	52.79	8.45	φsc = 0.157**
Within hosts	190	53454.13	281.34	281.34	45.04	φct = 0.465*
Total	195	96689.62	495.84	624.71		

https://doi.org/10.1371/journal.pone.0220031.t002

Table 3. Divergence time based on three arthropod mitochondrial divergence rates. Trees are presented in supplemental material S3 Fig.

	2.5%/my	3.54%/my	5.25%/my
Group 1 (Florida/All other populations)	0.923	0.613	0.427
Group 2 (US, El Salvador, Mex/ Brazil, Argentina)	0.685	0.455	0.316

https://doi.org/10.1371/journal.pone.0220031.t003

and 923 thousand years ago. This period precedes the agriculture expansion period in the Americas, which means that the lineages were formed before maize domestication. However, the period is also posterior to great geological transformation in the regions and might be related to climatic oscillation during the Pleistocene [55–57]. The most recent haplotypes were formed within the period of maize domestication and expansion in the Americas.

Hypotheses for the early history of *D. saccharalis* colonization

Early dispersal strategy—Floating on the Amazon basin rivers. The likely center of origin of *D. saccharalis* is the delta region of the Orinoco River, Venezuela, extending to the lower Amazon River [19,58]. Both basins, Orinoco and Amazon, cover an area of approximately 8,380,000 square kilometers (Fig 2, Amazon Basin), an area already large enough to be considered a remarkable dispersal achievement for insects with limited capacity to disperse. The first reports of the movement of *D. saccharalis* on infested aquatic grasses floating in the Amazon river are from eyewitness accounts [58]. The movement along the Amazon basin rivers would allow an incremental range expansion out of Venezuela throughout northern Brazil. However, this passive dispersal strategy using the rivers was probably more critical before the association of *D. saccharalis* with maize plants which occurred much later for this species.

The early dispersal and the considerable low rate of gene flow possibly created favorable conditions for population isolation and lineage formation as early as ~400–600 thousand years ago. The isolation might have been further enforced by climate oscillation during the late Pleistocene period and prior to the population expansion to more distant parts of the continent. However, the divergence time may vary depending on the clock rate used but consistently points to a period later than any significant geological event (i.e., isthmus of Panama formation ~2.5–3.5 Mya) in the region [59]. An interesting observation is the presence of two distinct lineages of *D. saccharalis* in the U.S., which opens questions about two possible colonization events.

Pre-Columbian era—The hypothesis for human-aided dispersal by terrestrial routes associated with maize. The dispersal history of *D. saccharalis* may have been influenced by human-mediated movements of maize (Fig 2, terrestrial). The '*D. saccharalis*-maize association' hypothesis is based on the early occupation of humans in America followed by the establishment of Pre-Columbian civilizations [60,61]. Approximately 14,000 years ago, or possibly earlier, humans began to domesticate plants and animals in the Americas [62–64]. These new cultivation systems associated with the exchange of goods between early American societies may have been responsible for maize dispersal from Mesoamerica to North and South America and the Caribbean following different routes [65–67]. *D. saccharalis* may have similarly dispersed to areas beyond its center of origin following maize dispersal in Pre-Columbian America. In North America, a lack of haplotype substructure within the second northern genetic group that includes Texas, Louisiana, Mexico, and El Salvador is a reliable indicator of a great movement by land, which mixed haplotypes in all locations. Populations from Louisiana and Texas were in the same cluster, share most haplotypes, and were strictly related to

populations from Central America (El Salvador) indicating a shared route of colonization and gene flow.

Archaeological evidence demonstrates that maize dispersal occurred through the South American lowlands and the Andean region, suggesting two major routes for *D. saccharalis* dispersal [65–70]. Analyzing South American lowlands populations (i.e., Brazilian populations), the closer relationships among the populations from the states of Mato Grosso do Sul, Mato Grosso and Tocantins may be associated with early human migrations(Fig 2). This route of human migration was most likely highly dependent on transportation along rivers, probably following the Xingu, Araguaia or/and the Tocantins rivers. These rivers are connected to the Amazon and Orinoco Basins. Alternatively, those haplotypes could have expanded inland from the coast (i.e., São Paulo) coming from the Caribbean. The hypothesis of maritime expansion will be discussed in the next section.

Regarding the *D. saccharalis* Argentinian populations, a pattern of sub-structure was found as well. While samples from the Pampa region (i.e., Buenos Aires) and the Dry Pampas (i.e., Tucumán and San Luis) grouped together and showed resemblance to the Brazilian populations, populations from remote parts of the northwest Chaco Serrano (i.e., Jujuy) were genetically differentiated. Together, this suggests that populations from the Pampas and the Dry Pampas might be associated with *D. saccharalis* dispersal routes through lowlands, while Jujuy might have a more ancient origin associated with maize dispersal through the Andean region. However, future studies are necessary to confirm this hypothesis.

Post-Columbian era—The hypothesis for the human-aided dispersal by maritime routes in association with sugarcane. In recent years, the long-distance movement of goods including sugarcane by maritime routes were of great importance and may explain some of the genetic structure found within the region (Fig 2, maritime routes). One hypothesis to explain the pattern of genetic divergence found in the Florida population could be the commercial movement of plants during the Columbian Exchange period in the 15th and 16th centuries. The Caribbean region was a crossing point at the beginning of European colonization of America, and served to exchange people and food between the Old and the New World [68]. This new commercial center was a favorable environment for the introduction of new diseases and pests that affected the native human population and their native (maize) and introduced (sugarcane) crops. We speculate that the exchange of infested plant material from the center of origin in Venezuela to the Caribbean islands promoted the spread of *D. saccharalis* by sea routes, northwards to the United States of America, specifically Florida, as well as southwards to Brazil, specifically to the São Paulo state region. The support for this hypothesis comes from the strong genetic structure within the U.S. that could indicate that the Florida population was colonized by an ancestral population probably from the Caribbean region. Alternatively, gradual range expansion of *D. saccharalis* through the Caribbean region could have reached Florida, where the founding population remained relatively isolated throughout the years. More studies including samples from the Caribbean region would be necessary to corroborate this hypothesis.

Recent routes of expansion—Agricultural development and routes of colonization of the Brazilian Cerrado. In Brazil, two genetic clusters were derived from the nuclear markers (Fig 3), which could indicate different introduction routes or a gradual process of colonization. The first and largest genetic cluster included the states of São Paulo and Paraná (southeast and south regions), places on the Brazilian coast that could be linked to the first sugarcane plants that arrived in São Paulo state from Caribbean routes. Moreover, when we consider the mitochondrial marker, haplotypes found in São Paulo (Hap3 and Hap7) and Paraná (Hap3, Hap5, and Hap7) were strongly shared with other locations and have a more recent genealogy. Mitochondrial information (COI analysis) supports the hypothesis of a recent process of expansion

originating from east to west within Brazil. Also part of this cluster, populations from Goiás and Minas Gerais states (southeast/center-west) were likely colonized by insects present in infested material from São Paulo. It is very likely that the demographic event that allowed the colonization of southeast areas belongs to the same putative route for the colonization of areas further south through the lowlands to Argentina (i.e., the F_{ST} between São Paulo and Buenos Aires was 0.02).

The transportation of plants across states seems to be the most critical pathway for *D. saccharalis* dispersion in Brazil countryside, which has been a recent development. This relatively recent introduction of *D. saccharalis* in the Brazilian Cerrado, followed by its population expansion may explain the lower allelic richness observed in populations from central Brazil compared to populations collected in the states of São Paulo, where the gene pool was more diverse. Due to its high commercial value during the last decade, sugarcane had a considerable agricultural expansion in areas of the Cerrado. Interesting, however, is that the oldest mitochondrial haplotype was present only in MT (Hap8) and was not shared with any other locations indicating a second route of colonization, and is probably associated with the earlier lowland colonization from the Amazon region. In summary, the combination of haplotypes present in the Cerrado is likely the result of the sugarcane expansion from the east (i.e., São Paulo) and perhaps an earlier pre-sugarcane expansion associated with the movement of maize along the rivers by native populations. More efforts should be made to characterize *D. saccharalis* populations from northern parts of Brazil to test this hypothesis.

Conclusions

Diatraea saccharalis is a successful invasive species that colonized habitats with very distinct environmental conditions across the Americas. Likely, multiple introductions of populations over a long time period explain the complex process of colonization of new areas. After an initial period of differentiation that separates different lineages, human migration and trade was likely a major factor in explaining the spread of *D. saccharalis* populations throughout the Americas. *Diatraea saccharalis* used maize as the early vehicle for dispersal whereas sugarcane was used in a second historical moment. More information about the genetic background of the Caribbean and South America populations, mainly from Venezuela and Argentina, is needed to support our hypotheses and investigate the lowland and Andean routes of dispersal.

Supporting information

S1 Fig. Loci distribution of *Diatraea saccharalis* populations using the F_{ST} approach implemented in Lositan. Loci plotted in the red area are candidates for being under positive selection while loci plotted in the yellow area are candidates for being under balancing selection. (TIF)

S2 Fig. Phylogeny tree. Maximum likelihood phylogenetic tree was constructed with MEGA X using *D. saccharalis* COI sequences. Bootstrap replication values of 1,000 were used, and the number at each node represents the percentage of bootstrap support for each cluster. Data present three distinct clusters separated by geographical locations. (TIF)

S3 Fig. Time tree of *D. saccharalis* cytochrome oxidase subunit I (COI) sequences using 5.23%/My divergence rate. (TIFF)

S1 Table. Populations of *Diatraea saccharalis* collected and their geographic locations. (DOCX)

S2 Table. Haplotype identification of *D. saccharalis* collected in BOLD database based on 533 bp fragment of cytochrome oxidase subunit I (COI).
(DOCX)

S3 Table. Results from the molecular clock test using maximum likelihood method under Hasegawa-Kishino-Yano (1985) (+G). The null hypothesis of equal evolutionary rate throughout the tree was rejected at a 5% significant level ($p = 0.000$). The analysis involved 16 haplotype sequences.
(DOCX)

S4 Table. Pairwise values of F_{ST} (below diagonal) and P-values (above) of *Diatraea saccharalis* populations. Bold indicates non-significant P-values ($P > 0.05$).
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: Fabricio J. B. Francischini, Steven J. Clough, Maria I. Zucchi.

Data curation: Fabricio J. B. Francischini, Alessandro Alves-Pereira, João Paulo Gomes Viana, Steven J. Clough.

Formal analysis: Fabricio J. B. Francischini, Erick M. G. Cordeiro, Jaqueline B. de Campos, João Paulo Gomes Viana, Xing Wu, Wei Wei, Andrea Joyce, Steven J. Clough.

Funding acquisition: Steven J. Clough.

Investigation: Fabricio J. B. Francischini, Jaqueline B. de Campos, Sofia Fogliata, Steven J. Clough.

Methodology: Fabricio J. B. Francischini, Jaqueline B. de Campos, Alessandro Alves-Pereira, João Paulo Gomes Viana, Steven J. Clough.

Project administration: Steven J. Clough.

Resources: Steven J. Clough.

Supervision: Steven J. Clough, Maria I. Zucchi.

Validation: Gabriela Murua.

Visualization: Fabricio J. B. Francischini, Jaqueline B. de Campos, Alessandro Alves-Pereira, Gabriela Murua.

Writing – original draft: Fabricio J. B. Francischini, Jaqueline B. de Campos, Alessandro Alves-Pereira, João Paulo Gomes Viana, Xing Wu, Wei Wei, Patrick Brown, Andrea Joyce, Gabriela Murua, Sofia Fogliata, Steven J. Clough, Maria I. Zucchi.

Writing – review & editing: Fabricio J. B. Francischini, Erick M. G. Cordeiro, Jaqueline B. de Campos, Alessandro Alves-Pereira, João Paulo Gomes Viana, Xing Wu, Wei Wei, Patrick Brown, Andrea Joyce, Gabriela Murua, Sofia Fogliata, Steven J. Clough, Maria I. Zucchi.

References

1. Dingle H. Migration Strategies of Insects. Science. 1972; 175(4028):1327–1335. https://doi.org/10.1126/science.175.4028.1327 PMID: 17813822

2. Chen MH, Dorn S. Microsatellites reveal genetic differentiation among populations in an insect species with high genetic variability in dispersal, the codling moth, *Cydia pomonella* (L.) (Lepidoptera: Tortricidae). Bull Entomol Res. 2010; 100(1):75–85. https://doi.org/10.1017/S0007485309006786 PMID: 19366473

3. Kirk H, Dorn S, Mazzi D. Molecular genetics and genomics generate new insights into invertebrate pest invasions. Evol Appl. 2013; 6:842–856. https://doi.org/10.1111/eva.12071 PMID: 29387170

4. Black WC, Baer CF, Antolin MF, DuTeau NM. Population genomics: genome-wide sampling of insect populations. Annu Rev Entomol. 2001; 46:441–469. https://doi.org/10.1146/annurev.ento.46.1.441 PMID: 11112176

5. Nosil P, Gompert Z, Farkas TE, Comeault AA, Feder JL, Alex Buerkle C, et al. Genomic consequences of multiple speciation processes in a stick insect. Proc R Soc B Biol Sci. 2012; 279(1749):5058–5065. https://doi.org/10.1098/rspb.2012.0813 PMID: 22696527

6. Kawecki TJ, Ebert D. Conceptual issues in local adaptation. Ecology Letters. 2004; 7(12):1225–1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x

7. Holland RA, Wikelski M, Wilcove DS. How and why do insects migrate? Science. 2006; 313 (5788):794–796. https://doi.org/10.1126/science.1127272 PMID: 16902129

8. Hoffmann BD, Courchamp F. Biological invasions and natural colonisations: are they that different? NeoBiota. 2016; 29:1–14. https://doi.org/10.3897/neobiota.29.6959

9. Robinet C, Roques A, Pan H, Fang G, Ye J, Zhang Y, et al. Role of human-mediated dispersal in the spread of the pinewood nematode in China. PLoS One. 2009; 4(2):e4646. https://doi.org/10.1371/journal.pone.0004646 PMID: 19247498

10. Carlton JT. Community assembly and historical biogeography in the North Atlantic Ocean: The potential role of human-mediated dispersal vectors. Hydrobiologia. 2003; 503:1–8. https://doi.org/10.1023/B:HYDR.0000008479.90581.e1

11. Ward DF, Beggs JR, Clout MN, Harris RJ, O'Connor S. The diversity and origin of exotic ants arriving in New Zealand via human-mediated dispersal. Divers Distrib. 2006; 12(5):601–609. https://doi.org/10.1111/j.1366-9516.2006.00270.x

12. Reynolds AM, Reynolds DR. Aphid aerial density profiles are consistent with turbulent advection amplifying flight behaviours: Abandoning the epithet "passive." Proc R Soc B Biol Sci. 2009; 276(1654):137–143. https://doi.org/10.1098/rspb.2008.0880 PMID: 18782743

13. Kennedy JS. A turning point in the study of insect migration. Nature. 1961; 189:785–791. https://doi.org/10.1038/189785a0

14. Bowler DE, Benton TG. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. Biol Rev. 2005; 80(2):205–225. https://doi.org/10.1017/S1464793104006645 PMID: 15921049

15. Farrow RA, Duly JC. Long-range movements as an adaptive strategy in the *genus heliothis* (Lepidoptera: Noctuidae): A review of its occurrence and detection in four pest species. Aust J Zool. 1987; 35 (1):1–24. https://doi.org/10.1071/ZO9870001

16. Bebber DP, Holmes T, Gurr SJ. The global spread of crop pests and pathogens. Glob Ecol Biogeogr. 2014; 23(12):1398–1407. https://doi.org/10.1111/geb.12214

17. Kim KS, Sappington TW. Boll weevil (*Anthonomus grandis Boheman*) (Coleoptera: Curculionidae) dispersal in the southern United States: evidence from mitochondrial DNA variation. Environ Entomol. 2004; 33(2):457–470. https://doi.org/10.1603/0046-225X-33.2.457

18. Miller NJ, Guillemaud T, Giordano R, Siegfried BD, Gray ME, Meinke LJ, et al. Genes, gene flow and adaptation of *Diabrotica virgifera virgifera*. Agricultural and Forest Entomology. 2009; 11(1):47–60. https://doi.org/10.1111/j.1461-9563.2008.00398.x

19. Myers JG. The Ecological Distribution of some South American Grass and Sugar-cane Borers (*Diatraea* spp., Lep., Pyralidae). Bull Entomol Res. 1935; 26: 335–342. https://doi.org/10.1017/S0007485300036646

20. Box HE. The Crambine Genera *Diatraea* and *Xanthopherne* (Lep., Pyral.). Bull Entomol Res. 1931; 22 (1):1–50. https://doi.org/10.1017/S0007485300029709

21. Lange CL, Scott KD, Graham GC, Sallam MN, Allsopp PG. Sugarcane moth borers (Lepidoptera: Noctuidae and Pyraloidea): phylogenetics constructed using COII and 16S mitochondrial partial gene sequences. Bull Entomol Res. 2004; 94(5):457–464. https://doi.org/10.1079/BER2004320 PMID: 15385065

22. Palacio Cortés AM, Zarbin PHG, Takiya DM, Bento JMS, Guidolin AS, Consoli FL. Geographic variation of sex pheromone and mitochondrial DNA in *Diatraea saccharalis* (Fab., 1794) (Lepidoptera: Crambidae). J Insect Physiol. 2010; 56(11):1624–1630. https://doi.org/10.1016/j.jinsphys.2010.06.005 PMID: 20558173

23. Joyce AL, Sermeno Chicas M, Serrano Cervantes L, Paniagua M, Scheffer SJ, Solis MA. Host-plant associated genetic divergence of two *Diatraea* spp. (Lepidoptera: Crambidae) stemborers on novel crop plants. Ecol Evol. 2016; 6(23):8632–8644. https://doi.org/10.1002/ece3.2541 PMID: 28031813

24. Hayward KJ. A broca da cana-de-açúcar. Bras Açucareiro. 1943; 11: 69–74.

25. Caixeta DF. Dispersão de machos de *Diatraea saccharalis* (Fabricius) (Leipdoptera: Crambidae) em cana-de-açúcar. Universidade Estadual Paulista. 2010.

26. Joyce AL, White WH, Nuessly GS, Solis MA, Scheffer SJ, Lewis ML, et al. Geographic population structure of the sugarcane borer, *Diatraea saccharalis* (F.) (Lepidoptera: Crambidae), in the southern United States. PLoS One. 2014; 9(10):e110036. https://doi.org/10.1371/journal.pone.0110036 PMID: 25337705

27. Pashley DP, Hardy TN, Hammond AM, Mihm JA. Genetic Evidence for Sibling Species within the Sugarcane Borer (Lepidoptera: Pyralidae). Ann Entomol Soc Am. 1990; 83(6):1048–1053. https://doi.org/10.1093/aesa/83.6.1048

28. Silva-Brandão KL, Santos T V., Cônsoli FL, Omoto C. Genetic diversity and structure of brazilian populations of *Diatraea saccharalis* (Lepidoptera: Crambidae): Implications for pest management. J Econ Entomol. 2015; 108(1):307–316. https://doi.org/10.1093/jee/tou040 PMID: 26470135

29. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS One. 2011; 6:e19379. https://doi.org/10.1371/journal.pone.0019379 PMID: 21573248

30. Solis MA, Metz M. An illustrated guide to the identification of the known species of *Diatraea* Guilding (Lepidoptera, Crambidae, Crambinae) based on genitalia. Zookeys. 2016; 565: 73–121. https://doi.org/10.3897/zookeys.565.6797 PMID: 27081337

31. Francischini FJB, De Campos JB, Alves-Pereira A, Gomes Viana JP, Grinter CC, Clough SJ, et al. Morphological and molecular characterization of Brazilian populations of *Diatraea saccharalis* (Fabricius, 1794) (Lepidoptera: Crambidae) and the evolutionary relationship among species of *Diatraea Guilding*. PLoS One. 2017; 12(11):e0186266. https://doi.org/10.1371/journal.pone.0186266 PMID: 29145390

32. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987; 19:11–15.

33. Poland JA, Brown PJ, Sorrells ME, Jannink J-L. Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach. PloS One. 2012; 7(2): e32253. https://doi.org/10.1371/journal.pone.0032253 PMID: 22389690

34. Lu F, Lipka AE, Glaubitz J, Elshire R, Cherney JH, Casler MD, et al. Switchgrass Genomic Diversity, Ploidy, and Evolution: Novel Insights from a Network-Based SNP Discovery Protocol. PLoS Genet. 2013; 9(1):e1003215. https://doi.org/10.1371/journal.pgen.1003215 PMID: 23349638

35. Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, et al. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS One. 2014; 9(2):e90346. https://doi.org/10.1371/journal.pone.0090346 PMID: 24587335

36. Lischer HEL, Excoffier L. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics. 2012; 28:298–299. https://doi.org/10.1093/bioinformatics/btr642 PMID: 22110245

37. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G. LOSITAN: A workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics. 2008; 28(2):298–299. https://doi.org/10.1186/1471-2105-9-323 PMID: 18662398

38. Beaumont MA, Nichols RA. Evaluating Loci for Use in the Genetic Analysis of Population Structure. Proc R Soc B Biol Sci. 1996; 263(1377):1619–1626. https://doi.org/10.1098/rspb.1996.0237

39. Keenan K, Mcginnity P, Cross TF, Crozier WW, Prodöhl PA. DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol. 2013; 4(8):782–788. https://doi.org/10.1111/2041-210X.12067

40. Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics. 1978; 89(3):583–590. PMID: 17248844

41. Kamvar ZN, Tabima JF, Grünwald NJ. *Poppr*: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ. 2014; 2:e281. https://doi.org/10.7717/peerj.281 PMID: 24688859

42. Jombart T. Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics. 2008; 24(11):1403–1405. https://doi.org/10.1093/bioinformatics/btn129 PMID: 18397895

43. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol. 2018; 35(6):1547–1549. https://doi.org/10.1093/molbev/msy096 PMID: 29722887

44. Librado P, Rozas J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics. 2009; 25(11):1451–1452. https://doi.org/10.1093/bioinformatics/btp187 PMID: 19346325

45. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007; 7(214):1–8. https://doi.org/10.1186/1471-2148-7-214 PMID: 17996036

46. Gaunt MW, Miles MA. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. Mol Biol Evol. 2002; 19(5):748–761. https://doi.org/10.1093/oxfordjournals.molbev.a004133 PMID: 11961108

47. Brower A V. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proc Natl Acad Sci. 1994; 91(14):6491–6495. https://doi.org/10.1073/pnas.91.14.6491 PMID: 8022810

48. Ruiz C, Jordal B, Serrano J. Molecular phylogeny of the tribe *Sphodrini* (Coleoptera: Carabidae) based on mitochondrial and nuclear markers. Mol Phylogenet Evol. 2009; 50(2009):44–58. https://doi.org/10.1016/j.ympev.2008.09.023 PMID: 18948213

49. Ribera I, Fresneda J, Bucur R, Izquierdo A, Vogler AP, Salgado JM, et al. Ancient origin of a Western Mediterranean radiation of subterranean beetles. BMC Evol Biol. 2010; 10:29. https://doi.org/10.1186/1471-2148-10-29 PMID: 20109178

50. Andújar C, Serrano J, Gámez-Zurita J. Winding up the molecular clock in the genus *Carabus* (Coleoptera: Carabidae): Assessment of methodological decisions on rate and node age estimation. BMC Evol Biol. 2012; 12:40. https://doi.org/10.1186/1471-2148-12-40 PMID: 22455387

51. Papadopoulou A, Anastasiou I, Vogler AP. Revisiting the insect mitochondrial molecular clock: The mid-Aegean trench calibration. Mol Biol Evol. 2010; 27(7):1659–1672. https://doi.org/10.1093/molbev/msq051 PMID: 20167609

52. Goodman KR, O'Grady P. Molecular Phylogeny and Biogeography of the Hawaiian Craneflies *Dicranomyia* (Diptera: Limoniidae). PLoS One. 2013; 8(9):e73019. https://doi.org/10.1371/journal.pone.0073019 PMID: 24058455

53. Beaumont MA. Adaptation and speciation: What can Fst tell us? Trends Ecol Evol. 2005; 20(8):435:440. https://doi.org/10.1016/j.tree.2005.05.017 PMID: 16701414

54. Via S, Lande R. Genotype-Environment interaction and the evolution of phenotypic plasticity. Evolution. 1985; 39(3):505–522. https://doi.org/10.1111/j.1558-5646.1985.tb00391.x PMID: 28561964

55. Hewitt GM. Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society B: Biological Sciences. 2004; 359(1442):183–195. https://doi.org/10.1098/rstb.2003.1388 PMID: 15101575

56. Hewitt G. The genetic legacy of the Quaternary ice ages. Nature. 2000; 405:907–913. https://doi.org/10.1038/35016000 PMID: 10879524

57. Ohnishi N, Uno R, Ishibashi Y, Tamate HB, Oi T. The influence of climatic oscillations during the Quaternary Era on the genetic structure of Asian black bears in Japan. Heredity. 2009; 579–589. https://doi.org/10.1038/hdy.2009.28 PMID: 19319151

58. Myers JG. The original Habitat and Hosts of three major Sugar-cane Pests of Tropical America (Diatraea, Castnia and Tomaspis). Bull Entomol Res. 1932; 23(2):257–271. https://doi.org/10.1017/S0007485300004181

59. O'Dea A, Lessios HA, Coates AG, Eytan RI, Restrepo-Moreno SA, Cione AL, et al. Formation of the Isthmus of Panama. Science Advances. 2016; 2(8):e1600883. https://doi.org/10.1126/sciadv.1600883 PMID: 27540590

60. Dillehay TD. The late pleistocene cultures of South America. Evol Anthropol. 1999; 7(6):206–216. https://doi.org/10.1002/(SICI)1520-6505(1999)7:6<206::AID-EVAN5>3.0.CO;2-G

61. Fraser B. The first South Americans: Extreme living. Nature. 2014; 514:24–26. https://doi.org/10.1038/514024a PMID: 25279901

62. Lia V V., Confalonieri VA, Ratto N, Cámara Hernández JA, Miante Alzogaray AM, Poggio L, et al. Microsatellite typing of ancient maize: Insights into the history of agriculture in southern South America. Proc R Soc B Biol Sci. 2007; 274:545–554. https://doi.org/10.1098/rspb.2006.3747 PMID: 17476775

63. Diamond J. Evolution, consequences and future of plant and animal domestication. Nature. 2002; 418:700–707. https://doi.org/10.1038/nature01019 PMID: 12167878

64. Matsuoka Y, Vigouroux Y, Goodman MM, Sanchez G. J, Buckler E, Doebley J. A single domestication for maize shown by multilocus microsatellite genotyping. Proc Natl Acad Sci. 2002; 99(9):6080–6084. https://doi.org/10.1073/pnas.052125199 PMID: 11983901

65. Bedoya CA, Dreisigacker S, Hearne S, Franco J, Mir C, Prasanna BM, et al. Genetic diversity and population structure of native maize populations in Latin America and the Caribbean. PLoS One. 2017; 12 (4):e0173488. https://doi.org/10.1371/journal.pone.0173488 PMID: 28403177

66. Vigouroux Y, Glaubitz JC, Matsuoka Y, Goodman MM, Sánchez G. J, Doebley J. Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. Am J Bot. 2008; 95 (10):1240–1253. https://doi.org/10.3732/ajb.0800097 PMID: 21632329

67. Diamond J, Bellwood P. Farmers and their languages: The first expansions. Science. 2003; 300 (5619):597–603. https://doi.org/10.1126/science.1078208 PMID: 12714734

68. Goodman MM, Brown WL. Races of corn. In: Sprague GF, Dudley JW, editors. Corn and corn improvement. 3rd ed. Madison, Wisconsin: ASA, CSSA, and SSSA; 1988. pp. 33–79.

69. Freitas FO, Bustamante PG. Amazonian Maize: Diversity, Spatial Distribution and Historical-Cultural Diffusion. Tipití J Soc Anthropol Lowl South Am. 2013; 11:60–65.

70. Bracco M, Cascales J Hernández JC, Poggio L, Gottlieb AM, Lia VV. Dissecting maize diversity in lowland South America: Genetic structure and geographic distribution models. BMC Plant Biol. 2016; 16:186. https://doi.org/10.1186/s12870-016-0874-5 PMID: 27561710