Multiple incursion pathways for *Helicoverpa armigera* in Brazil show its genetic diversity spreading in a connected world

Jonas Andre Arnemann1,2,3*, Stephen Roxburgh1, Tom Walsh1, Jerson Guedes2, Karl Gordon1, Guy Smagghe3 & Wee Tek Tay1

The Old World cotton bollworm *Helicoverpa armigera* was first detected in Brazil with subsequent reports from Paraguay, Argentina, Bolivia, and Uruguay. This pattern suggests that the *H. armigera* spread across the South American continent following incursions into northern/central Brazil, however, this hypothesis has not been tested. Here we compare northern and central Brazilian *H. armigera* mtDNA COI haplotypes with those from southern Brazil, Uruguay, Argentina, and Paraguay. We infer spatial genetic and gene flow patterns of this dispersive pest in the agricultural landscape of South America. We show that the spatial distribution of *H. armigera* mtDNA haplotypes and its inferred gene flow patterns in the southwestern region of South America exhibited signatures inconsistent with a single incursion hypothesis. Simulations on spatial distribution patterns show that the detection of rare and/or the absence of dominant mtDNA haplotypes in southern *H. armigera* populations are inconsistent with genetic signatures observed in northern and central Brazil. Incursions of *H. armigera* into the New World are therefore likely to have involved independent events in northern/central Brazil, and southern Brazil/Uruguay-Argentina-Paraguay. This study demonstrates the significant biosecurity challenges facing the South American continent, and highlights alternate pathways for introductions of alien species into the New World.

Biological invasions are major ecological phenomena that influence the worldwide distribution of species. They can be a major driver of ecological change, affecting conservation (loss of biodiversity and species extinction1,2), human health (e.g., the Zika virus transmitted by the invasive *Aedes aegypti*3) and agriculture (e.g., the spread of the cotton boll weevil, *Anthonomus grandis* through the Americas4; the introduction of fall army worm *Spodoptera frugiperda* into Africa from the New World5). In Brazil alone, 35 new pests have been detected6 or confirmed7,8 in the last 10 years.

Invasive insects cost a minimum of US$70 billion per year globally for goods and services9. The Old World cotton bollworm, *Helicoverpa armigera* is considered a major agricultural pest with an estimated annual global cost to agriculture of over USD$5 billion10, and Kriticos et al.11 estimated that the arrival of *H. armigera* into North America would put at risk an extra USD$78 billion of agricultural output. A number of life history traits predispose *H. armigera* to be a highly successful insect pest12. These include: (i) high polyphagy, where larvae of this insect pest are known to feed on over 180 plant hosts from at least 68 plant families13; (ii) its long distance migratory ability with migrations of up to 2,000 km14–17; (iii) the ability to enter a facultative diapause as pupae under unfavourable environmental conditions such as extreme high or low temperatures12; and (iv) high fecundity and a short generation time, capable of completing up to 10 to 11 generations per year12,18,19. Such fast generation times could aid in building population size, and thus contribute to a successful invasion20.

In Brazil, *H. armigera* was confirmed in January-February 201321,22, and the incursion has resulted in over USD$800 million in losses and control costs since 20126,23,24. The presence of *H. armigera* was also reported in rapid succession in neighbouring western and south-western countries, including Paraguay in October 201325,26, Argentina in August-October 201327, and in 2014/2015 in Uruguay28. *H. armigera* was also confirmed in the

1CSIRO, Black Mountain Laboratories, Clunies Ross Street, ACT 2601, Canberra, Australia. 2Department of Crop Protection, Universidade Federal de Santa Maria, Santa Maria, Brazil. 3Plants and Crops, Faculty of Bioscience Engineering, Ghent University, Ghent, Belgium. *email: jonasarnemann@gmail.com
northern/central Brazilian incursions. haplotype diversity ($h$)

| Location          | Nucleotide diversity ($\pi$) | Haplotype diversity ($h$) |
|-------------------|------------------------------|---------------------------|
| Asia              | 0.00360 ± 0.00034            | 0.912 ± 0.024             |
| Europe            | 0.00238 ± 0.00040            | 0.738 ± 0.082             |
| Australia         | 0.00403 ± 0.00057            | 0.882 ± 0.047             |
| South America     | 0.00244 ± 0.00015            | 0.769 ± 0.018             |
| Asia + Europe     | 0.00314 ± 0.00027            | 0.862 ± 0.032             |
| Asia + Europe + Australia | 0.00380 ± 0.00026       | 0.894 ± 0.024             |

Table 1. Comparison of *Helicoverpa armigera* partial mtDNA COI gene nucleotide diversity ($\pi$ ± s.e.) and haplotype diversity ($h$ ± s.e.) between different countries/continents.

Caribbean countries of Puerto Rico\textsuperscript{29} and Dominican Republic\textsuperscript{30,31} and on the mainland of United States of America\textsuperscript{32}. The detection of *H. armigera* from southern Florida down to Argentina over just 3 years demonstrates the speed at which this species has established in the New World. It clearly has the ability to spread through a connected landscape and via an island hopping, stepping-stone dispersal model. The pathway of *H. armigera* into the South American continent may be linked with commodity movements involving importations of agricultural and horticultural products from multiple Old World destinations\textsuperscript{33}. Whilst multiple origins of *H. armigera* in Brazil are likely, it remains unclear whether these founder individuals arrived as a single or as multiple event(s). By examining the genetic signature of the populations found in the South American continent using mtDNA markers, it may be possible to utilise spatial genetic signatures of the insect to infer frequencies of introductions. It is almost certain that *H. armigera* had been present in South America for a period of time [e.g.\textsuperscript{22–31}] prior to its first identification in Brazil\textsuperscript{34}, remaining undetected due to its close morphology with the New World species *H. zeas*, and the difficulty of detecting invasive pests at the early stages of incursions\textsuperscript{35}. It is unknown whether the *H. armigera* populations detected across South America arose from a single, or multiple, original introductions, and where these were. A detailed population genetic study is needed to test the hypothesis that there were multiple unrelated introduction events into South America, which would have significant implications for biosecurity preparedness for the South American continent and the potential reintroduction of novel adaptive ecotypes\textsuperscript{36} into the Old World\textsuperscript{33,34}. To test this hypothesis, we undertook the present study, in which we show that gene flow and spatial distribution patterns of *H. armigera* mtDNA haplotypes support multiple introductions of the *H. armigera* into South America, with the incursion(s) in the southern regions likely independent from the northern/central Brazilian incursions.

**Results**

**PCR amplification and sequence analysis.** All specimens from the southern/south-western regions of South America were successfully sequenced for the mtDNA COI fragment (GenBank accession numbers MG230495–MG230526; KU255535–KU255543 from\textsuperscript{37}) using the Noc–COI–F/R primer pairs. Sequence identity searches against the NCBI GenBank database confirmed that all suspected moths matched (i.e., 99–100% nucleotide identity) published *H. armigera* sequences, and did not contain premature stop codons.

Unique amino acid substitutions were detected in three haplotypes (i.e., Harm_BC47, Harm_BC42, Harm_BC43). For Harm_BC47, this involved an L/V change, and in Harm_BC42 a I/M change. Both substitutions involved amino acids with hydrophobic side chain. In Harm_BC43 the unique amino acid involved an A to G change where both amino acids belonged to the ‘small’ category. All remaining unique haplotypes (e.g., Harm_BC06, BC23, BC24, BC34, BC37, BC39, BC44, BC45, BC46) had nucleotide transition substitutions at 3\textsuperscript{rd} codon positions. The most significant unique non-Brazilian haplotypes detected were Harm_BC13, Harm_BC16, and Harm_BC17. All three unique haplotypes shared SNPs with other haplotypes, indicating that they did not have unexpected base changes. Furthermore, these haplotypes were also detected multiple times in separate sequencing efforts.

The range of genetic distances (i.e., measures of genetic divergence/degree of differentiation) of *H. armigera* within Asia (China, India and Pakistan) and within Australia were both 0.00–0.04%, while within Europe (Germany and unknown sites) and within South America (Brazil, Argentina, Uruguay and Paraguay) were both 0.00–0.02%. Estimates of evolutionary divergence between all *H. armigera* sequences from Australia, Asia, Europe and South America were therefore likewise low and ranged between 0.00–0.04%. Observed nucleotide diversity between countries/continents in *H. armigera* ranged from 0.0024 ± 0.0004 (s.e.) to 0.0040 ± 0.0006 (s.e.) (Table 1).

**Haplotypes.** Haplotypes were collated from previous works\textsuperscript{37,38} and from this study. A total of 47 haplotypes were identified from 314 individuals that consisted of 44 individuals from Asia, 18 from Australia, 28 from Europe and 226 from South America (Suppl. Table 3). Four most prevalent mtDNA COI haplotypes identified in this study were designated Harm_BC01, Harm_BC02, Harm_BC03, and Harm_BC04.

Brazil shares 5 mtDNA COI haplotypes with Asia (Harm_BC01, Harm_BC02, Harm_BC04, Harm_BC06 and Harm_BC07) and also 5 haplotypes with Europe (Harm_BC01, Harm_BC02, Harm_BC03, Harm_BC04 and Harm_BC06). Haplotypes Harm_BC01 and Harm_BC02 were present in all Brazilian states, and the Harm_BC01 haplotype was shared with all other locations, except Argentina, Paraguay and Uruguay, which had unique haplotypes (Harm_BC43, Harm_BC44, Harm_BC46 and Harm_BC47). The haplotype Harm_BC02 in Brazil was shared with Asia, Europe and Uruguay. Brazil also had 12 unique haplotypes (Harm_BC05, Harm_BC14, Harm_BC23, Harm_BC24, Harm_BC34, Harm_BC35, Harm_BC36, Harm_BC37, Harm_BC38, Harm_BC39,
Harm_BC42 and Harm_BC45), and 80% of Brazilian individuals belonged to haplotypes Harm_BC01, Harm_BC02, Harm_BC03, Harm_BC04 and Harm_BC05.

**Helicoverpa armigera haplotypes distribution and validation of the null model.** Results from validation of the null model indicated no tendency for the test to generate either Type I or Type II errors (i.e., having an almost perfect rectangular distribution of P-values across 10,000 statistical tests, with each test based on 10,000 random realisations of pseudo-observed data). At the 5% confidence level, almost exactly 5% of tests yielded statistically significant results with randomly generated data for the full dataset (501/10,000 = 0.0501 (Suppl. Fig. 1a); and when testing for Brazil vs. non-Brazil 496/10,000 = 0.0496 (Suppl. Fig. 1b)).

**Whole table analysis.** Figure 1 shows the distribution of $\chi^2_{OBS}$ under the null model. No random realisations were found to have a value of the test statistic greater than or equal to that observed ($\chi^2_{OBS} = 719.2$), thereby indicating highly significant non-randomness within the observations ($P = 0.000$), and strong support that at least one haplotype/location observation has observations that are either more or less than expected by chance alone.

**Sub-table analysis.** Evidence from above (Whole Table Analysis) strongly supported some mtDNA haplotypes were differentially distributed across sampling sites, and the $T_{SDIFF}$ analysis was therefore used to further identify those haplotypes that were unduly rare or common across the locations. The results indicated haplotype Harm_BC01 and Harm_BC02 were both simultaneously under-represented in Argentina (ARG), Paraguay (PRY) and one Brazilian state (BA), and overly represented in another Brazilian state (PI) (Fig. 2). The analysis also found evidence to support haplotypes Harm_BC03 and Harm_BC04 as being sporadically overly represented in three Brazilian states (BA, MT and RS) (dark blue cells, Fig. 2).

**Brazil vs. Non-Brazil.** In the ‘Brazil vs. non-Brazil’ treatment of haplotype distribution data (Suppl. Table 4; Fig. 3), non-randomness of haplotype distribution within the matrix was again confirmed by the $\chi^2_{OBS}$ analysis ($\chi^2_{OBS} = 164.0$, $P$-value < 0.000; Fig. 3). The irregular distribution of the test statistic in Fig. 2 reflected a smaller dataset and, therefore, fewer possible combinations of allowable observations to fulfill the row and column constraints.

Consistent with the Whole Table Analysis (i.e., Fig. 2), at the scale of the individual haplotypes, Harm_BC01 and Harm_BC02 were simultaneously under-represented in non-Brazilian and over-represented in Brazilian. For haplotypes Harm_BC13, Harm_BC16 and Harm_BC17 the opposite was true. A trend for over-representation of unique haplotypes Harm_BC43, Harm_BC44, Harm_BC46 and Harm_BC47 at Non-Brazil locations was also detected in this analysis, but deemed statistically insignificant under the Benjamin and Hochberg (1995) false discovery rate correction. The difference in the strength of statistical significance between the two analyses (Figs. 2 and 3) is due to a smaller allowable number of permutations when allocating observations at random to the two aggregated spatial categories.

**AMOVA and FST analysis.** Overall $F_{ST}$ estimates based on the partial mtDNA COI gene, when treating our data set as Brazilian, non-Brazilian (Uruguay, Argentina and Paraguay), and Old World samples, showed that a significant $F_{ST}$ value (0.2879) was detected between Brazilian and non-Brazilian samples, indicating significant population sub-structure (i.e., low gene flow) between these populations (Suppl. Table 5). In contrast, the low $F_{ST}$ estimate (0.0742, see also Tay et al.31) suggested consistent gene flow between Brazilian and Old World populations. Between the Non-Brazilian and the Old World populations, the significant pairwise $F_{ST}$ estimate (0.2022) obtained suggested limited gene flow between these populations as inferred from the maternally inherited mtDNA COI gene.

At the fine-scale level, population pairwise $F_{ST}$ (Suppl. Table 6) indicated limited gene flow between the Paraguay population and all of Brazilian populations, but not between the Paraguay, Argentina and Uruguay populations. Interestingly, limited gene flow also occurred between the Brazilian state of Rio Grande do Sul and the northern states of Piauí and Bahia, as well as with central Brazilian state of Mato Grosso. Similarly, the Argentinian population showed limited gene flow with Maranhão, Piauí and Bahia. Significant population sub-structure exhibited by Paraguay analysed with the whole Brazilian population would suggest the H. armigera population from Paraguay is unlikely to be recently established from Brazilian populations.

Analysis of molecular variance (AMOVA) among populations of H. armigera indicated that low levels of genetic variation existed between populations. Approximately 7% of variation detected could be attributed to that
between individuals within groups, while the majority (>90%) of variation could be explained by the heterogeneous populations present across different geographic regions in the South American continent (Suppl. Table 7).

**Discussion**

The rapid spread and establishment of *H. armigera* across much of the South American continent has generated a very large population with significant impacts on agricultural production. In contrast to usual invasions such as the African incursion of the New World fall armyworm *S. frugiperda* [34,39–42], the invasive *H. armigera* population in the New World appears to be very diverse. Furthermore, the spatial distribution of this diversity strongly suggests that the population has spread from two different regions of introduction. Statistical analyses of the haplotype distribution patterns show that some of the *H. armigera* haplotypes most commonly found in Brazil appeared to be uncommon in Argentina (i.e., Harm_BC01) and Paraguay (Harm_BC02), and conversely, that some of the less common and/or unique haplotypes found in the non-Brazilian countries (e.g., Harm_BC13, Harm_BC16, Harm_BC17) appeared disproportionately uncommon in Brazil. Furthermore, *F*~ST~ analyses suggest reduced gene flow between populations from the Cone Sul region (Southern Brazil, Argentina, Paraguay and Uruguay) and populations from either northern/central Brazil or from the Old World. Within individual countries, disproportionally over-represented haplotypes were identified (e.g., see Suppl. Table 2 and Fig. 2), and

|       | Non-Brazil |       | Brazil states |
|-------|------------|-------|----------------|
|       | ARG | URY | PRY | BA | MA | MT | PI | RR | PR | SC | RS |
| Harm_BC01 | -1.54 | -0.58 | 0.07 | -4.21 | 2.03 | 0.15 | -0.19 | 1.90 | 0.24 | 1.99 | 0.14 |
| Harm_BC02 | -1.37 | 0.28 | -4.02 | 1.49 | 1.59 | -0.85 | -0.15 | 0.00 | -0.11 | -0.41 | -0.74 |
| Harm_BC03 | 0.65 | -0.22 | -1.05 | 2.06 | -1.26 | 0.90 | -1.39 | -0.49 | -0.29 | -0.68 | -1.22 |
| Harm_BC04 | 0.78 | -0.14 | -0.58 | 2.07 | -1.44 | -1.34 | -1.83 | -0.92 | -0.20 | -0.44 | -0.75 |
| Harm_BC05 | -0.13 | -0.08 | 0.34 | 1.06 | -0.69 | -0.67 | 0.35 | 0.46 | 0.87 | -0.29 | -0.55 |
| Harm_BC06 | -0.01 | -0.01 | -0.04 | -0.40 | 0.92 | -0.08 | -0.18 | -0.07 | -0.02 | -0.04 | -0.07 |
| Harm_BC07 | -0.07 | -0.04 | -0.19 | -0.23 | -0.36 | -0.36 | 2.13 | -0.31 | -0.08 | -0.17 | -0.33 |
| Harm_BC13 | -0.04 | -0.03 | 1.89 | -1.38 | -0.21 | -0.21 | -0.49 | -0.18 | -0.05 | 0.90 | -0.20 |
| Harm_BC14 | -0.03 | -0.02 | -0.08 | 1.23 | -1.16 | -0.16 | -0.37 | -0.14 | -0.04 | -0.08 | -0.15 |
| Harm_BC16 | 0.97 | -0.02 | 0.92 | -0.77 | -0.16 | -0.16 | -0.37 | -0.14 | -0.04 | -0.08 | -0.15 |
| Harm_BC17 | -0.03 | -0.02 | 1.92 | -0.77 | -0.16 | -0.16 | -0.37 | -0.14 | -0.04 | -0.08 | -0.15 |
| Harm_BC23 | -0.01 | -0.01 | -0.04 | -0.38 | -0.08 | -0.08 | 0.81 | -0.07 | -0.02 | -0.04 | -0.08 |
| Harm_BC24 | -0.01 | -0.01 | -0.04 | -0.38 | -0.08 | -0.08 | -0.19 | 0.93 | -0.02 | -0.04 | -0.08 |
| Harm_BC34 | -0.01 | -0.01 | -0.04 | 0.62 | -0.08 | -0.08 | -0.19 | -0.07 | -0.02 | -0.04 | -0.08 |
| Harm_BC35 | -0.01 | -0.01 | -0.04 | 0.61 | -0.08 | -0.08 | -0.18 | -0.07 | -0.02 | -0.04 | -0.08 |
| Harm_BC36 | -0.01 | -0.01 | -0.04 | 0.62 | -0.08 | -0.08 | -0.19 | -0.07 | -0.02 | -0.04 | -0.08 |
| Harm_BC37 | -0.01 | -0.01 | -0.04 | -0.38 | 0.92 | -0.08 | -0.18 | -0.07 | -0.02 | -0.04 | -0.08 |
| Harm_BC38 | -0.01 | -0.01 | -0.04 | -0.38 | -0.08 | 0.92 | -0.19 | -0.07 | -0.02 | -0.04 | -0.08 |
| Harm_BC39 | -0.01 | -0.01 | -0.04 | 0.62 | -0.08 | -0.08 | -0.19 | -0.07 | -0.02 | -0.04 | -0.08 |
| Harm_BC42 | -0.01 | -0.01 | -0.04 | -0.38 | -0.08 | -0.08 | -0.18 | -0.07 | -0.02 | -0.04 | 0.91 |
| Harm_BC43 | -0.01 | -0.01 | 0.96 | -0.39 | -0.08 | -0.08 | -0.18 | -0.06 | -0.02 | -0.04 | -0.10 |
| Harm_BC44 | -0.01 | 0.99 | -0.04 | -0.38 | -0.08 | -0.08 | -0.17 | -0.06 | -0.01 | -0.04 | -0.11 |
| Harm_BC45 | -0.01 | -0.01 | -0.04 | -0.38 | -0.08 | -0.08 | -0.16 | -0.06 | -0.01 | -0.03 | 0.85 |
| Harm_BC46 | -0.01 | -0.01 | 0.96 | -0.33 | -0.07 | -0.06 | -0.11 | -0.03 | -0.01 | -0.02 | -0.32 |
| Harm_BC47 | 1.00 | -0.00 | -0.01 | -1.13 | -0.05 | -0.05 | -0.16 | -0.11 | -0.04 | -0.11 | -0.32 |

**Figure 2.** TSDIFF values and indication of statistically significant deviations from the null model based of a false discovery rate (FDR) of 0.05 [54]. Pale red cells are cells with a non-significantly lower number of observations than expected at random, and pale blue cells non-significantly higher numbers of observations. Dark red cells indicate significantly lower numbers of observations. Dark blue cells indicate significantly higher numbers of observations.
within Brazil, the rare Harm_BC13 haplotype was detected in the State of Santa Catarina at sites located only approximately 350 km from Paraguay where this rare haplotype was also identified. Two other unique haplotypes (Harm_BC42, Harm_BC45) were also only identified in the southern state of Rio Grande do Sul in Brazil.

It is necessary to keep in mind that the northern and central Brazilian *H. armigera* may contain the same rare haplotypes as our southern region populations, although studies involving greater sample sizes from similar sampling periods\(^\text{37,38}\) did not detect these rare haplotypes. Nevertheless, rare haplotypes detected in Argentina, Paraguay and Uruguay have also not been detected in other Brazilian samples to-date. We were unable to include the mtDNA COI sequences of Tay *et al.*\(^\text{31}\) due to different mtDNA COI regions being characterised. The study of Tay *et al.*\(^\text{31}\) analysed substantial Brazilian populations from similar period and had identified similar haplotype frequency patterns as this study. Furthermore, Tay *et al.*\(^\text{31}\) showed that within Brazil, the *H. armigera* population contained both globally common and globally rare haplotypes. At the national level, there were 13 rare haplotypes identified in Brazil, and is similar to the total number of 12 rare haplotypes identified in this study. Despite using different parts of the mtDNA COI gene, both studies have detected similar patterns of rare and common partial mtDNA COI haplotypes, suggesting that the results of our analyses are representative of the diversity present in Brazil at sampling time.

Population structure studies in *H. armigera* based on the mtDNA genes have found a general lack of substructure even for populations separated by considerable geographic distances (e.g.\(^\text{43,44}\)), and this finding is supported by studies based on limited nuclear markers (e.g.\(^\text{35,37,45–49}\)). Anderson *et al.*\(^\text{36}\) demonstrated differences between sub-species of *H. armigera* present in Australia/New Zealand (i.e., *H. armigera conferta*) and the Old

### Table 3

| Haplotypes  | Non-Brazil | Brazil  |
|-------------|------------|---------|
|             | TS\(\text{DIFF}\) | P-values | TS\(\text{DIFF}\) | P-values |
| Harm_BC01   | -1.90      | 0.000   | 1.90      | 0.000   |
| Harm_BC02   | -5.20      | 0.000   | 5.20      | 0.000   |
| Harm_BC03   | -0.70      | 0.895   | 0.70      | 0.895   |
| Harm_BC04   | -0.04      | 1.000   | 0.04      | 1.000   |
| Harm_BC05   | -0.50      | 1.000   | 0.50      | 1.000   |
| Harm_BC06   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC07   | -0.28      | 1.000   | 0.28      | 1.000   |
| Harm_BC13   | 1.83       | 0.014   | -1.83     | 0.014   |
| Harm_BC14   | -0.12      | 1.000   | 0.12      | 1.000   |
| Harm_BC16   | 1.89       | 0.005   | -1.89     | 0.005   |
| Harm_BC17   | 1.89       | 0.004   | -1.89     | 0.004   |
| Harm_BC23   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC24   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC34   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC35   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC36   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC37   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC38   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC39   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC42   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC43   | 0.94       | 0.116   | -0.94     | 0.116   |
| Harm_BC44   | 0.94       | 0.115   | -0.94     | 0.115   |
| Harm_BC45   | -0.06      | 1.000   | 0.06      | 1.000   |
| Harm_BC46   | 0.94       | 0.113   | -0.94     | 0.113   |
| Harm_BC47   | 0.95       | 0.108   | -0.95     | 0.108   |

Figure 3. TS\(\text{DIFF}\) for the Brazil vs Non-Brazil comparison of associated P-values.
World sub-species (i.e., H. armigera armigera), but not between global H. armigera armigera populations using genome-wide SNP markers. In northern/central Brazil, gene flow patterns of H. armigera showed non-significant levels of population substructure but exhibited reduced gene flow with southern South American populations of H. armigera. These southern South American populations, and particularly those of Paraguay and Argentina, also exhibited significant population substructure with Old World populations, suggesting that their origins differed from the origins of the founding populations present in northern/central Brazil. 

Given that sufficient gene flow to prevent population structure had previously been detected in Brazilian H. armigera populations, and that H. armigera in Argentina, Paraguay and Uruguay, was confirmed shortly after the Brazilian detections, it would be a logical assumption that this likely represented natural migration, and/or movements of contaminated agricultural commodities from Brazil to the southern regions of South America. It was therefore unexpected to find unique mtDNA haplotypes in the southern South American populations that were as yet unreported in Brazil. Equally as unexpected was the lack of the most common haplotypes (e.g., Harm_BC01, Harm_BC02) in these countries given that they represented the major Brazil, and in fact, global, haplotypes. 

To further explain the observed heterogeneous haplotype patterns across Paraguay, Uruguay, Argentina, and the southern Brazilian states of Rio Grande do Sul, Santa Catarina and Paraná (i.e., the Cone Sul region), two hypotheses may be put forward: (I) intrinsic factors associated with new biological incursions (e.g., stochastic lineage sorting, survival/reproductive variability, etc.) in a new environment, and (II) independent incursion pathways of H. armigera into South America. 

In Brazil where the incursion of H. armigera was first reported, stochastic lineage sorting of founding populations (i.e., hypothesis I) could lead to the observed heterogeneous haplotype distribution patterns. This could involve factors such as lower population density and variability at population introduction phase and/or the lag-phase, high variability of reproductive success rates (e.g., see Gaither et al., variable adaptation success rates (e.g., differential response to attacks by parasitoids and/or predation rates), susceptibility to viral/bacterial/fungal pathogen attacks, climatic stress, etc. to the novel New World environments at the early incursion stages. 

In Brazil, across a number of studies, unexpectedly high genetic diversity of H. armigera has been detected as represented by multiple maternal lineages (Suppl. Table 2 e.g., Suppl. Fig. 2). This would likely indicate a complex incursion history of H. armigera into Brazil, and may be associated with agricultural commodity movements from Old World regions into Brazil. For example, repeated incursions/releases at different Brazilian sites could function as population source increasing propagule pressure, raising and maintaining diversity which is important in sustaining an incipient population (i.e. hypothesis II). These suggested scenarios involving the highly volatile and variable periods of an exotic organism’s biology contrast the scenarios offered by Leite et al. where repeated bottleneck effects such as potentially associated with differential pest control management strategies were deemed likely factors that underpinned the rapid population expansion signatures in both H. armigera and the New World endemic and closely related H. zea (but see). In fact, H. zea in the New World was hypothesised as the outcome of an earlier incursion and the subsequent divergence (ca. 1.5 million years ago) from its common ancestor with H. armigera, and involved a founder population with limited genetic diversity. The H. zea genome, sequenced prior to the recent arrival of H. armigera in the New World, showed no evidence for subsequent introgression with H. armigera, and no evidence for the gain of additional genes affecting host use, but rather for the loss of genes already present in H. armigera.

Repeated introductions and high propagule pressure are increasingly being recognised as important factors that underpin the establishment of an alien species. With repeated introduction events, the likelihood of diverse maternal lineages that ultimately contribute to propagule pressure is high. Together with lineage sorting and stochastic processes (e.g., demographic, environmental) experienced by the invasive species in the new environment, sampling of the mtDNA COI gene and the construction of a haplotype network will likely appear similar to one of a rapid population expansion (i.e., a ‘star-shaped’ haplotype network). This scenario of a ‘star-shaped’ haplotype network, as detected in H. armigera populations in the South Americas (see Fig. 2 of), differed fundamentally to that reported for H. zea (i.e., Fig. 1 of; Fig. 2 of). Multiple introductions of an invasive pest insect that resulted in a mtDNA genetic signature similar to a rapid population expansion signature, has also been previously reported in Brazil (e.g. the Asian citrus psyllid Diaphorina citri, see).

With the migration and dispersal ability of H. armigera in mind, the high frequency (i.e., 68%) of the two most common Harm_BC01 and Harm_BC02 haplotypes in Brazil populations, and a lack of population structures in northern/central Brazil (e.g.,) and the rest of the world, it was perhaps unexpected to observe statistically significant spatial mtDNA COI haplotype patterns and FST estimates in the Cone Sul region. The significant over- and underrepresented haplotypes in the Cone Sul region suggest that this population likely originated from somewhere outside the extensively sampled areas of central and northern Brazilian populations (i.e., hypotheses II). For example, haplotypes Harm_BC13 and Harm_BC17 for Paraguay and Harm_BC44 and Harm_BC47 for Uruguay and Argentina, respectively, were over-represented in these countries and to a lesser extent, also the over-representation of haplotypes Harm_BC43, Harm_BC44, Harm_BC46 and Harm_BC47, although the P-values (0.100–0.114) lie outside of the range usually considered significant, thereby adding support that these maternal lineages likely originated from non-Brazil source populations (i.e., hypothesis II). 

The detection of the H. armigera genetic spatial signatures identified in this work provide the first insights into potential transnational patterns of the incursion of this species into the New World. This reflects opportunities to improve biosecurity protocols relating to phytosanitary practices of agricultural and horticultural commodity movements in the Cone Sul region. Although the haplotype distribution pattern from Rio Grande do Sul was excluded by our spatial analysis, pairwise FST estimates indicated a limited gene flow of the Rio Grande do Sul population with populations from three Brazilian states (see Suppl. Table 5), and coupled with the statistically significant over-representation of unique haplotypes in the Cone Sul region, our results therefore added support to the hypothesis of multiple introduction pathways of H. armigera into South America.
Similarly, the presence of limited gene flow between Rio Grande do Sul and Argentina populations and the various Brazilian populations might suggest different degrees of admixture between *H. armigera* from non-Brazilian and Brazilian populations. This observation should be further explored and should consider utilizing genome-wide SNP markers to increase detection efficiency of population admixture. Although the low sample size (n = 2) in Uruguay has prevented meaningful interpretations of gene flow patterns based on the mtDNA COI marker, future genome-wide SNP markers studies on these individuals may enable migration patterns and the introduction history to be interpreted with more confidence. Taken as a whole, these *Fst* results suggest that non-Brazilian *H. armigera* populations, and particularly those from Paraguay, followed by Argentina, and to a small extend that of the Rio Grande do Sul population, likely represent populations that were wholly or partially derived from alternative incursions event(s) from that detected in northern/central Brazil. Single locus markers have clear limitations, however recent studies from multiple mtDNA markers and from genome-wide SNP markers have also demonstrated multiple introductions from globally diverse *H. armigera* populations in the invasive Brazilian populations. Taken as a whole, these findings provided evidence to support multiple independent introductions of *H. armigera* into the South American continent over the effect of stochastic lineage sorting. Our finding is not without precedent, with the globally invasive hemipteran whitefly *Bemisia tabaci* MED species (i.e., the real *B. tabaci*62) also being shown to have been independently introduced into this southern region of South America in addition to an earlier introduction elsewhere in Brazil.

Findings that populations of *H. armigera* in the South American continent likely also involved multiple introduction pathways was fortuitous, because the pattern identified in this study is unlikely to be maintained over time as mixing with other populations in South America and further incursions are likely. That incursions of *H. armigera* from the Old World potentially involved multiple pathways will have significant implications to pest and resistance management strategies in the New World. For example, populations of *H. armigera* around the world have developed resistance to conventional pesticides (e.g., and, the South American populations have also arrived at least with resistance to pyrethroids but also potentially with various enhanced allelochemical detoxification traits. Increasing genetic diversity is a key factor that underpins increasing invasion success.69,70 While significant levels of genetic diversity now exist in Brazil, and propagule pressure has also therefore decreased, the genetic make-up of these populations could be further bolstered by likely unrelated source populations from other parts of the Old World and this will further complicate and challenge management strategies. As pointed out by De Barro et al., measures to restrict the recruitment of additional genetic diversity should be maintained even after establishment and spread have occurred, so as to avoid increasing the genetic diversity of damaging invasive pests.

**Material and Methods**

**Sample collection and DNA extraction.** Suspected *H. armigera* adults were collected using delta traps baited with the female sexual pheromone Iscalure armigera (ISCA Tecnologias LTDA, Ijui, RS, Brazil) randomly installed in soybean fields in the Brazilian states Rio Grande do Sul, Santa Catarina, and Paraná, and also in Uruguay, Argentina and Paraguay, in the 2014/15 cropping season (specifically, early April 2014; Table 2). Preservation of specimens, gDNA extraction procedures and PCR amplification and sequencing of partial mitochondrial DNA COI gene were done following PCR conditions as detailed in Arnemann et al. using the primers Noc-COI-F (S′-GGCGAAAATGTACCTTTATTCAC-3′) and Noc-COI-R (S′-CCAAAATCTAAAATAATGGTGG-3′).

**Selection of published mtDNA COI haplotype dataset.** For our analysis, the following criteria underpinned our selection of mtDNA COI from publicly available sources: (A) only include published sequences from studies that had undergone critical review processes to avoid inclusion of untested haplotypes; (B) the sequences must originate from samples collected at a similar sampling time period as our material, (C) the published sequences must match our characterised partial mtDNA COI gene region, and (D) the populations must include northern/central Brazil to enable spatial comparisons to our Southern populations. Based on these criteria, sequences from three studies were chosen, with the published data from Arnemann et al. also here included as part of the southern populations of *H. armigera*, while the studies of Leite et al. and Mastrangelo et al. represented the most comprehensive population diversity surveys at the mtDNA COI region in Brazil and fulfilled all four criteria.

| Code/state | Geographical coordinates Lat/Lon | Sampling date | GenBank Accession. |
|------------|---------------------------------|---------------|-------------------|
| ARG        | −32.1963222, −61.716597222      | 09-Apr-2014   | MG230495-MG230498 |
| PRT        | −25.25568055; −57.56711111      | 12-Apr-2014   | MG220499-MG220502 |
| URF        | −34.90434444; −54.93699166      | 03-May-2014   | MG220525-MG220526 |
| BRA-BS     | −29.72635611; −53.56116111      | 10-Apr-2014   | MG220356-MG220357 |
| BRA-SC     | −26.462400; −53.508684444       | 06-Apr-2014   | MG220358-MG220352 |
| BRA-PR     | −24.158361111; −49.821575       | 25-Mar-2014   | MG220303-MG220305 |

**Table 2.** Collection sites, dates, and mtDNA COI GenBank accession numbers of *Helicoverpa armigera* specimens from Brazil (BRA), Uruguay (URY), Argentina (ARG) and Paraguay (PRY). Brazilian States are Rio Grande do Sul (RS), Santa Catarina (SC) and Paraná (PR).
Sequence analysis of partial mtDNA COI gene. The programs Pregap and Gap4 within the Staden sequence analysis package were used for editing DNA trace files and to assemble sequence contigs (i.e., haplotypes). Assembled mtDNA COI haplotypes were checked for premature stop codons that may be indicative of pseudogenes. Categorisation of global mtDNA COI haplotypes at the 5′ gene region, (Suppl. Table 1) and estimates of evolutionary divergence between all H. armigera individuals (i.e., between South America vs. Europe vs. Asia vs. Australia; n = 314) involved 548 bp of the mtDNA COI partial gene in the final dataset. The evolutionary divergence between haplotypes was estimated using the maximum composite likelihood model, and included all (i.e., 1st + 2nd + 3rd) codon positions using MEGA. Estimates of haplotype diversity (h ± SE) and nucleotide diversity (π ± SE) were carried out using the molecular evolution software package DNA sequence polymorphism (DnaSP) v. 5.10.01. Assessments for presence of PCR errors, or nuclear mitochondrial sequences (NUMTs) involved (1) examination for presence of premature stop codons in the sequenced partial mtDNA COI gene region, (2) ascertaining for conservation of amino acid substitution patterns where presence of unique amino acid changes were further assessed for conservation of biochemical properties and/or their molecule sizes, and (3) sequence characterisation of rare haplotypes as confirmed by multiple independent PCR and sequencing efforts.

Analysis of Helicoverpa armigera haplotypes spatial distribution patterns. To better investigate the spatial distribution patterns of H. armigera haplotypes in the South American continent, a matrix table was prepared for the frequencies of the 25 H. armigera mtDNA COI haplotypes identified to-date from 11 South American locations (i.e., from Brazilian samples sites: Bahia (BA), Maranhão (MA), Mato Grosso (MT), Piauí (PI), Roraima (RR), Parana (PR), Santa Catarina (SC), and Rio Grande do Sul (RS); and from Argentina, Uruguay, and Paraguay; Suppl. Table 2), prior to performing a contingency table analysis using the χ² statistic to detect departures, as detailed below. The statistical test was based on the randomisation of haplotypes x locations generated according to an appropriate null model (see ‘null model’ below), similar to Gotelli’s analysis of species co-occurrence data.

Null and alternative hypotheses. Due to the perceived unevenness of mtDNA COI haplotype spatial patterns, a statistical test was applied to ascertain whether the diversity and frequencies of individual haplotypes had occurred independently and at random across South American sampling sites, or whether there was evidence for spatial segregation. For such a test the null hypothesis is therefore that mtDNA COI haplotypes are distributed randomly across the locations. The alternative hypothesis therefore considers at least one haplotype as being either more or less common, in at least one location, than expected due to chance alone (i.e. haplotypes were randomly distributed across locations).

Test statistics used. Equation (1) below was used for the χ² statistic to detect departures from randomness:

\[ \chi^2_{\text{Obs}} = \sum_{ij} \left( \frac{Z_{\text{Obs},ij} - Z_{\text{Ran},ij}}{Z_{\text{Ran},ij}} \right)^2 \]

where \( Z_{\text{Obs},ij} \) is the observation for row i and column j within the matrix (i.e. Suppl. Table 2), and \( Z_{\text{Ran},ij} \) is the expected value for the same table entry calculated under the null hypothesis. If certain haplotypes and/or locations were disproportionately under or over-represented, then \( \chi^2_{\text{Obs}} \) would be expected to fall within the extreme tail of the distribution of \( \chi^2_{\text{Obs}} \), (i.e., the calculated value of the statistic when the null hypothesis is known to be true).

Although an analysis based on \( \chi^2_{\text{Obs}} \) provides an overall test of haplotype randomness across the landscape, information on which location and/or haplotype combinations underpin any deviation from randomness will require an additional test statistic (i.e., Eq. (2)) to be applied at the scale of each haplotype/location combination:

\[ TS_{\text{DIFF}} = Z_{\text{Obs},ij} - Z_{\text{Ran},ij} \]

Equation (2) implies that when \( TS_{\text{DIFF}} < 0.0 \) a haplotype would be less common than expected in location j due to chance alone, while \( TS_{\text{DIFF}} > 1.0 \) the combination would be more common than expected in location j by chance, and when \( TS_{\text{VAR}} = 0.0 \) the observed data conformed with the null model. Analysis of \( TS_{\text{DIFF}} \) can therefore be used to further explore primary circumstances (i.e., combinations of haplotype and location) and directionality (i.e., whether haplotypes were unexpectedly rare or common across the sites) if/when evidence of differentially distributed haplotypes across the landscapes was detected (i.e. to identify the genotypes that are unduly rare or common across the locations).

Determining the statistical significance of individual table entries is problematic because multiple comparisons are being simultaneously assessed. For example in Fig. 2, there are 11 locations x 25 haplotypes = 275 cell-level tests, of which approximately 14 would be expected to show significance simply by chance (i.e. at the 0.05 level of confidence). To control for this ‘familywise’ error rate the method of Benjamini and Hochberg was applied, using a false discovery rate of 0.10. This identified seven cell entries of interest in the full analysis (Fig. 2), and five pairwise (Brazil /non-Brazil) comparisons of interest when data were pooled across locations (Fig. 4).

Null model. The null model generated for this study involved distributing all 226 observations (i.e., the complete mtDNA COI dataset of H. armigera in South America; Suppl. Table 2) across haplotypes and locations at random, with the constraint of fixed row and column totals (i.e. the same number of observations of each haplotype, and the same number of observations for each location, are both retained in the randomised matrices). This treatment is necessary to ensure that the null model explicitly accounts for both unequal survey efforts at different sites and the overall lower frequencies of some haplotypes. The algorithm AS159 of Patefield was used to generate random tables that constrained both row and column marginal totals.
This null model represents a non-parametric alternative to Fisher’s exact test, which is also based on fixed marginal totals, but which relies upon a chi-squared distribution to approximate the underlying exact distribution. This approximation can be poor if the data are sparse (see Suppl. Table 3), hence the decision to calculate P-values via random permutation.

Calculating P-values. P-values for \( \chi^2_{\text{Obs}} \) were calculated by enumerating the number of times \( \chi^2_{\text{Ran}} \) was less than or equal to \( \chi^2_{\text{Obs}} \) \((N_{\text{LessEqual}})\), and also the number of times \( \chi^2_{\text{Ran}} \) was greater than or equal to \( \chi^2_{\text{Obs}} \) \((N_{\text{GreaterEqual}})\), based on 10,000 randomly generated tables. The two-tailed P-value \(^7^9\) is given by:

\[
P = \text{value} = \frac{\min(N_{\text{GreaterEqual}}, N_{\text{LessEqual}})}{10,000} \times 2
\]

P-values for \( T_{\text{DIFF}} \) were calculated similarly, but at the scale of individual table cells. Note that as the quantities \( N_{\text{GreaterEqual}} \) and \( N_{\text{LessEqual}} \) can potentially overlap in their included values (i.e., both include an equality term), P-values can potentially exceed 1.0; in such cases values are rounded down to 1.0.

Validating the null model. The statistical test was validated by creating \( n \) pseudo-observed data sets under the null model, followed by application of the test to each of the created data sets to confirm appropriate Type I and Type II error rates. At the 5% confidence level only 5% of pseudo random data sets should be statistically significant, and the expected P-value distribution across the \( n \) tests should be rectangular. If it is found that more than 5% of test results are significant at the 5% confidence level, then this indicates an elevated Type I error rate (i.e., a chance that the null hypothesis is incorrectly rejected when no real differences are actually present). Conversely, when less than the nominal number of pseudo random data sets yield a significant result it can lead to incorrect acceptance of the null hypothesis (i.e., concluding no significant difference even when one is actually present; i.e., an elevated Type II error rate).

The null model was validated through repeated analyses (10,000 times) using random pseudo-data reconstructed to match the \( n = 266 \) observations (Suppl. Table 2), but allowing observations to be randomly allocated to the matrices, and thereby ensuring conformity with the null hypothesis. Tables that had no observations for a given row or column during the constructing of pseudo-data were excluded. The code, the program and test data are provided as supplementary material (Suppl. Material 1).

Brazil vs. Non-Brazil. The above analyses were repeated, but combining the location data to consider just two categories of samples – the Brazilian and the non-Brazilian (Argentina, Paraguay, Uruguay) samples. Note that this analysis was carried out as a guide to assist interpretation for the observed within-table patterns of deviations, and have not impacted our main findings which was the result from the full dataset (i.e., Fig. 2) analysis.

AMOVA and FST analysis. The population pairwise \( F_{ST} \) and AMOVA estimates were carried out using Arlequin 3.5.2.2\(^{39}\), and significance values were estimated with 10,000 permutations. For the AMOVA, populations were assigned into four groups to test molecular variation across geographical regions of Brazil (North: Roraima, Maranhão, Piauí; Central/Eastern: Bahia, Mato Grosso; South: Rio Grande do Sul, Santa Catarina, Paraná) and neighbouring countries (Argentina, Paraguay, Uruguay).

References
1. Lambertini, M. et al. Invasives: a major conservation threat. Science 333, 404–405 (2011).
2. Simberloff, D. et al. Impacts of biological invasions: what’s what and the way forward. Trends. Ecol. Evol. 28, 58–66 (2013).
3. Zanluca, C. First report of autochthonous transmission of zikavirus in Brazil. Mem. Inst. Oswaldo Cruz 110, 569–572 (2015).
4. Barbosa, S., Braga, S. R., Lukefahr, M. J. & Beingolea, G. O. Invasives: a major conservation threat. Science 333, 404–405 (2011).
5. Goergen, G., Kumar, P. L., Sankung, S. B., Togola, A. & Tamò, M. First report of outbreaks of the fall armyworm Spodoptera frugiperda (J. E. Smith) (Lepidoptera, Noctuidae), a new alien invasive pest in west and central Africa. PLoS ONE 11, e0165632, 10.1371/journal.pone.0165632 (2016).
44. Behere, G. T.

39. Goergen, G., Kumar, P. L., Sankung, S. B., Togola, A. & Tamò, M. First report of outbreaks of the fall armyworm

38. Castiglioni, E.

33. EUROPHYT. European Union Notification System for Plant Health Interceptions Annual Report 2014. Luxembourg: Food and Veterinary Office (2015).

32. Sosa-Gómez, D. R.

31. Senave. Senave en alerta tras ingreso de peligrosa plaga agrícola. Available at, http://www.abc.com.py/edicion-im7resa/economia/

30. EUROPHYT. European Union Notification System for Plant Health Interceptions Annual Report 2014. Luxembourg: Food and Veterinary Office (2015).

29. Mastrangelo, T.

28. Castiglioni, E.

27. Murúa, M. G.

26. SENAVE. Senave en alerta tras ingreso de peligrosa plaga agrícola. Available at, http://www.abc.com.py/edicion-im7resa/economia/

25. Senave. Senave en alerta tras ingreso de peligrosa plaga agrícola. Available at, http://www.abc.com.py/edicion-im7resa/economia/

24. Mastrangelo, T.

23. Goergen, G., Kumar, P. L., Sankung, S. B., Togola, A. & Tamò, M. First report of outbreaks of the fall armyworm

22. Tayeh, A.

21. Senave en alerta tras ingreso de peligrosa plaga agrícola. Available at, http://www.abc.com.py/edicion-im7resa/economia/

20. Tayeh, A.

19. Raulston, J. R.

18. Cunningham, J. P. & Zalucki, M. P. Understanding Heliothine (Lepidoptera: Heliothinae) pests: what is a host plant? J. Econ. Entomol.

17. Behere, G. T., Tay, W. T., Russell, D. A., Kranthi, K. R. & Batterham, P. Population genetic structure of the cotton bollworm Helicoverpa armigera (Hübner) (Lepidoptera: Noctuidae) in India as inferred from EPIC-PCR DNA markers. PLoS ONE 8, e53448, https://doi.org/10.1371/journal.pone.0053448 (2013).

16. Jones, C. M. et al. Genomewide transcriptional signatures of migratory flight activity in a globally invasive pest. Mol. Ecol. 24, 4901–4911 (2015).

15. Nibouche, S., Buès, R., Toubon, J. F. & Poitout, S. Allozyme polymorphism in the American cotton bollworm Helicoverpa armigera (Hübner) (Lepidoptera: Noctuidae): comparison of African and European populations. Hereditas 80, 438–445 (1998).

14. Widmer, M. W. & Schofield, P.

13. Cunningham, J. P. & Zalucki, M. P. Understanding Heliothine (Lepidoptera: Heliothinae) pests: what is a host plant? J. Econ. Entomol.

12. Fitt, G. P. Ecology of Heliothis species in relation to agroecosystems. Annu. Rev. Entomol. 34, 17–52 (1989).

11. Nagoshi, R. N.

10. Sharma, H. C. Heliothis/Helicoverpa management: emerging trends and strategies for future research. New Delhi: Oxford& IBH Publishing Co. Pvt. Ltd (2005).

9. Kriticos, D. J. et al. The potential distribution of invading Helicoverpa armigera in North America: is it just a matter of time? PLoS ONE 10, e0133224, https://doi.org/10.1371/journal.pone.0133224 (2015).

8. Bradhaw, C. J. A. et al. Massive yet grossly underestimated global costs of invasive insects. Nat. Commun. 7, 12986, https://doi.org/10.1038/ncomms12986 (2016).

7. Arnemann, J. A. et al. Mitochondrial DNA COI characterization of Helicoverpa armigera (Lepidoptera: Noctuidae) from Paraguay and Uruguay. Genet. Mol. Res. 15, 15028292, https://doi.org/10.4238/gmr.15028292 (2016).

6. Lopes-da-Silva, M., Sanches, M. M., Stancioli, A. R., Alves, G. & Sugayama, R. The role of natural and human mediated pathways for invasion of agricultural pests: a historical analysis of cases from Brazil. Agricultural Sciences 5, 634–646 (2014).

5. Fitt, G. P. Ecology of Heliothis species in relation to agroecosystems. Annu. Rev. Entomol. 34, 17–52 (1989).

4. Widmer, M. W. & Schofield, P.

3. Cunningham, J. P. & Zalucki, M. P. Understanding Heliothine (Lepidoptera: Heliothinae) pests: what is a host plant? J. Econ. Entomol.

2. Tay, W. T., Russell, D. A., Kranthi, K. R. & Batterham, P. Population genetic structure of the cotton bollworm Helicoverpa armigera (Hübner) (Lepidoptera: Noctuidae) in India as inferred from EPIC-PCR DNA markers. PLoS ONE 8, e53448, https://doi.org/10.1371/journal.pone.0053448 (2013).

1. Coaker, T. H. Investigations on Heliothis armigera in Uganda. B. Entomol. Res. 50, 487–506 (1959).

0. Cunningham, J. P. & Zalucki, M. P. Understanding Heliothine (Lepidoptera: Heliothinae) pests: what is a host plant? J. Econ. Entomol.
45. Daly, J. C. & Grepp, P. Genetic variation in Heliothis in Australia: species identification and gene flow in the two pest species H. armigera (Hübner) and H. punctigera Wallengren (Lepidoptera: Noctuidae). B. Entomol. Res. 75, 169–184 (1985).
46. Zhou, X., Faktor, O., Applebaum, S. W. & Coll, M. Population structure of the pestiferous moth Helicoverpa armigera in the eastern Mediterranean using RAPD analysis. Heredity (Edinb) 85, 251–256 (2000).
47. Xiao-feng, C., Sheng-jiang, T., Ren-yi, L., Ying, W. & Diamo, L. Study on the genetic variation of the cotton bollworm Helicoverpa armigera (Hübner) populations in China. Insect Sci. 7, 243–249 (2000).
48. Endersby, N. M., Hoffmann, A. A., McKechnie, S. W. & Weeks, A. R. Is there genetic structure in populations of Helicoverpa armigera from Australia? Entomol. Exp. Appl. 122, 253–263 (2007).
49. Vass, J. M., Brevault, T., Achaleke, J. & Menozzi, P. Genetic structure of the polyphagous pest Helicoverpa armigera (Lepidoptera: Noctuidae) across the sub-Saharan cotton belt. Commun. Agric. Appl. Biol. Sci. 73, 433–437 (2008).
50. Gaither, M. R., Bowen, B. W. & Tsonen, R. J. Population structure in the native range predicts the spread of introduced marine species. P. Roy. Soc. B-Biol. Sci. 280, 20130409, https://doi.org/10.1098/rspb.2013.0409 (2013).
51. Oliveira, H. N., Simonato, J., Glaeser, D. F. & Pereira, F. F. Parasitism of Helicoverpa armigera pupae (Lepidoptera: Noctuidae) by Tetrastichus howardi and Trichoplistusviarautae (Hymenoptera: Eulophidae). Semina. Ciencias Agràrias 37, 111–115 (2016).
52. Lockwood, J. L., Casse, P. & Blackburn, T. The role of propugel pressure in explaining species invasions. Trends. Ecol. Evol. 20, 223–228 (2005).
53. Simberloff, D. The role of propugel pressure in biological invasions. Annu. Rev. Ecol. Evol. S. 40, 81–102 (2009).
54. Chown, S. L. et al. Biological invasions, climate change and genomics. Evol. Appl. 8, 23–46 (2014).
55. Pearce, S. L. et al. Genomic innovations, transcription plasticity and gene loss underlying the evolution and divergence of two highly polyphagous and invasive Helicoverpa pest species. BMC Biol. 15, 69 (2017).
56. Mallet, J., Korman, A., Heckel, D. G. & King, P. Biochemical genetics of Heliothis and Helicoverpa (Lepidoptera, Noctuidae) and evidence for a founder event in Helicoverpa zea. Ann. Ent. Soc. Amer. 86, 189–197 (1993).
57. Grille, G., Gauthier, N., Buenahora, J., Basso, C. & Bonato, O. First report of the Q biotype of Bemisia tabaci in Argentina and Uruguay. Phytoparasitica 39, 235–238 (2011).
58. Guidolin, A. S., Fresia, P. & Cônsoli, F. L. The genetic structure of an invasive pest, the Asian citrus psyllid Diaphorina citri (Hemiptera: Liviidae). PLoS ONE 9, e115749, https://doi.org/10.1371/journal.pone.0115749 (2014).
59. Anderson, C. J., Tay, W. T., McGaughran, A., Gordon, K. & Walsh, T. K. Population structure and gene flow in the global pest Helicoverpa armigera. Mol. Ecol. 25, 5296–5311 (2016).
60. Tay, W. T., Evans, G. A., Boykin, L. M. & De Barro, P. J. Will the real Bemisia tabaci please stand up? PLoS ONE 7(11), e50550, https://doi.org/10.1371/journal.pone.0050550 (2012).
61. de Moraes, L. A., Carli, Glauber R. Sturmer, Luis E. Curioletti and Rubens A. Fiorin assisted with Brazilian and DNA preparation. This study was funded by CAPES, UFSM, CSIRO Health and Biosecurity ‘Genes of the world’ project. (2015).
62. Durigan, M. R. et al. Multiple recombination events between two cytochrome P450 loci contribute to global pyrethroid resistance in Helicoverpa armigera. PLoS ONE 13(11), e0197760, https://doi.org/10.1371/journal.pone.0197760 (2018).
63. Yang, Y., Chen, H., Wu, Y., Xing, Y. & Qiu, S. Mutated cadherin alleles in resistant strains. Pesticide Biochemistry and Physiology 280, 20130409, https://doi.org/10.1098/rspb.2013.0409 (2013).
64. Walsh, T. K. Multiple recombination events between two cytochrome P450 loci contribute to global pyrethroid resistance in Helicoverpa armigera. PLoS ONE 13(11), e0197760, https://doi.org/10.1371/journal.pone.0197760 (2018).
65. Yang, Y., Chen, H., Wu, Y., Xing, Y. & Qiu, S. Mutated cadherin alleles in resistant strains. Pesticide Biochemistry and Physiology 280, 20130409, https://doi.org/10.1098/rspb.2013.0409 (2013).
66. Zhang, H., Wu, S., Yang, Y., Tabashnik, B. E. & Wu, Y. Non-recessive B toxin resistance conferred by an intracellular cadherin mutation in field-selected populations of cotton bollworm. PLoS ONE 7, e53418, https://doi.org/10.1371/journal.pone.0053418 (2012).
67. Durigan, M. R. et al. High frequency of CYP337B3 gene associated with control failures of Helicoverpa armigera with pyrethroid insecticides in Brazil. Pesticide Biochemistry and Physiology, https://doi.org/10.1016/j.pestbp.2017.09.005 (2017).
68. Nair, R., Kalia, V., Aggarwal, K. K. & Gajar, G. T. Variation in the cadherin gene sequence of Cry1Ac susceptible and resistant Helicoverpa armigera (Lepidoptera: Noctuidae) and the identification of mutant alleles in resistant strains. Curr. Sci. 104, 215 (2013).
69. Vares, J. P., Hughes, A. R. & Grosberg, R. K. Mechanisms that drive evolutionary change: insights from species introductions and invasions in Species inspections: Insights into ecology, evolution, and biogeography (ed. Sax, D. F., Stachowicz, J. J., Gaines, S. D.) 229–257. Massachusetts: Sinauer Associates (2005).
70. Ellestrand, N. C. & Schierenbeck, K. A. Hybridization as a stimulus for the evolution of invasiveness in plants? Proc. Natl. Acad. Sci. USA 97, 7043–7050 (2000).
71. De Barro, P. J., Liu, S. S., Boykin, L. M. & Dinsdale, A. B. Bemisia tabaci: a statement of species status. Annu Rev Entomol 56, 1–19 (2011).
72. Tamura, K., Stetter, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729 (2013).
73. Librado, P. & Rozas, J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451–1452 (2009).
74. Gotelli, N. J. Null Model Analysis of species co-occurrence patterns. Ecology 81, 2606–2621 (2000).
75. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society B 57, 289–300 (1995).
76. McDonald, J. H. Handbook of Biological Statistics, 3rd ed. Sparky House Publishing, Baltimore, Maryland (2014).
77. Patfield, W. M. Algorithm AS 159: an efficient method of generating random RC tables with given row and column totals. Applied Statistics 30, 91–97 (1981).
78. Manly, F. J. B. Randomization, Bootstrap and Monte Carlo Methods in Biology, 2nd edn. London: Chapman and Hall/CRC (2001).

Acknowledgements
We thank Gustavo Ugalde and Elgion Loreto for laboratory assistance. We thank Claudia C. C. Antúnez and Enrique Castiglioni for providing Paraguay and Uruguay samples, respectively. Clerison R. Perini, Cristiano De Carli, Glauber R. Sturmer, Luis E. Curioletti and Rubens A. Fiorin assisted with Brazilian H. armigera sampling and DNA preparation. This study was funded by CAPES, UFSM, CSIRO Health and Biosecurity ‘Genes of Biosecurity Significance’ (R-8681-1) and Research Foundation – Flanders (FWO-Vlaanderen, Ghent University, Belgium. Portions of text/results in the manuscript have been previously published as part of PhD thesis of J.A.A.: Armemann, JA. (2015). Molecular tools to identify and study invasive pests in Brazil. PhD Thesis. Ghent University, Ghent, Belgium.
Author contributions
J.A.A., W.T.T., J.G., K.G., T.W. and G.S. conceived the ideas and designed methodology; J.A.A. and W.T.T. collected the data; J.A.A., W.T.T. and S.R. analysed the data; W.T.T., T.W., S.R., K.G. and J.A.A. led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-019-55919-9.

Correspondence and requests for materials should be addressed to J.A.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019