



## OPEN

## SUBJECT AREAS:

HAPLOTYPES  
INVASIVE SPECIES  
URBAN ECOLOGYInterbreeding among deeply divergent mitochondrial lineages in the American cockroach (*Periplaneta americana*)Christoph von Beeren<sup>1</sup>, Mark Y. Stoeckle<sup>2</sup>, Joyce Xia<sup>1,3</sup>, Griffin Burke<sup>1,4</sup> & Daniel J. C. Kronauer<sup>1</sup>Received  
7 October 2014Accepted  
14 January 2015Published  
6 February 2015

Correspondence and requests for materials should be addressed to C.v.B. (vonbeeren@rockefeller.edu) or M.Y.S. (mark.stoeckle@rockefeller.edu)

<sup>1</sup>Laboratory of Insect Social Evolution, The Rockefeller University, New York, NY, USA, <sup>2</sup>Program for the Human Environment, The Rockefeller University, New York, NY, USA, <sup>3</sup>Hunter College High School, New York, NY, USA, <sup>4</sup>Bard College, Annandale-on-Hudson, NY, USA.

DNA barcoding promises to be a useful tool to identify pest species assuming adequate representation of genetic variants in a reference library. Here we examined mitochondrial DNA barcodes in a global urban pest, the American cockroach (*Periplaneta americana*). Our sampling effort generated 284 cockroach specimens, most from New York City, plus 15 additional U.S. states and six other countries, enabling the first large-scale survey of *P. americana* barcode variation. *Periplaneta americana* barcode sequences (n = 247, including 24 GenBank records) formed a monophyletic lineage separate from other *Periplaneta* species. We found three distinct *P. americana* haplogroups with relatively small differences within ( $\leq 0.6\%$ ) and larger differences among groups (2.4%–4.7%). This could be interpreted as indicative of multiple cryptic species. However, nuclear DNA sequences (n = 77 specimens) revealed extensive gene flow among mitochondrial haplogroups, confirming a single species. This unusual genetic pattern likely reflects multiple introductions from genetically divergent source populations, followed by interbreeding in the invasive range. Our findings highlight the need for comprehensive reference databases in DNA barcoding studies, especially when dealing with invasive populations that might be derived from multiple genetically distinct source populations.

Globalization facilitates the introduction of invasive species that can damage native ecosystems, cause severe economic losses, and threaten human health<sup>1,2</sup>. Early detection and rapid response are cornerstones of successful management strategies. However, identification can be problematic, first and foremost due to a lack in taxonomic expertise<sup>3,4</sup>. DNA-based approaches such as ‘DNA barcoding’ potentially provide relatively rapid and inexpensive species identifications<sup>4</sup>. Although the reliability and usefulness of DNA barcoding are subject of extensive debate<sup>5–8</sup>, this method has been shown to effectively distinguish species in many animal groups<sup>9–13</sup>. By comparing a short, standardized fragment of the mitochondrial gene *cytochrome c oxidase I* (COI) to a reference DNA barcode library, animal specimens can usually be assigned to species, as long as the database contains relevant reference sequences<sup>14,15</sup>. This approach has been applied to native populations of diverse animals<sup>16–21</sup> and a variety of invasive taxa<sup>3,22–25</sup>. Successful invaders are generally presumed to represent expansions of a small founder population and are therefore likely genetically uniform<sup>24,26–29</sup>. For example, the invasive gall wasp *Quadrastichus erythrinae* shows a complete lack of mitochondrial as well as nuclear diversity across the Pacific, including Japan, Hawaii, Guam, and Samoa<sup>30</sup>, presumably reflecting a single outbreak starting with a small number of individuals. On the other hand, invasive populations may also represent multiple independent introductions, potentially from genetically distinct source populations<sup>31,32</sup>. Such a pattern of introduction may complicate the identification of invasive species via DNA barcodes, and may require more extensive sampling to establish a reliable reference library.

Here we assess DNA barcode variation in invasive populations of the American cockroach, *Periplaneta americana* (Linnaeus), one of the most abundant, widely distributed, and hated urban pests<sup>33,34</sup>. Though the native range of *P. americana* is unknown (possibly tropical Africa or South Asia<sup>34–36</sup>), all urban populations will be considered invasive in this study as it seems likely that they were established via human-aided dispersal. According to the World Health Organization, cockroaches are highly damaging pests worldwide in terms of potential health problems (allergies, asthma, and transmission of pathogens by contaminating food) and costs for pest control<sup>37</sup>. Furthermore, cockroaches are uniquely unpopular as most people find them disgusting, associating their presence with uncleanliness and disease<sup>38</sup>.



Although abundant and living in close proximity to humans, surprisingly little is known about *P. americana* biology outside the laboratory. Aspects of infestation control have been examined<sup>39–42</sup>, but few studies have investigated the biology of *P. americana* in urban settings<sup>33,43</sup>. Particularly strikingly, data on genetic variation is all but absent. The goal of this study was twofold: (1) to assess the genetic diversity of *P. americana* in urban populations, thus creating a valuable reference DNA barcode library; (2) to test the potential of DNA barcoding for quick and accurate molecular identification of this urban pest species. To facilitate specimen collection, we set up a citizen science project based in New York City (NYC), a center of world commerce and home to an abundance of cockroaches as well as more than eight million people<sup>44–46</sup>. We received several hundred specimens from scores of locations in NYC and beyond, enabling the first large-scale study of genetic diversity and providing a reference DNA barcode library for this important urban pest. We found deeply divergent mitochondrial lineages that could naively be interpreted as indicative of multiple cryptic species. However, nuclear DNA sequence data revealed extensive gene flow between mitochondrial clades, consistent with a single biological species. Our findings highlight the value of expanded sampling to accurately delineate species boundaries via DNA barcoding, in particular for invasive species.

## Results

The collection effort of 85 participants generated 284 specimens from 16 U.S. states and Argentina, Australia, Belize, Guyana, Spain, and Venezuela (Supplementary Tab. S1). In addition, 24 *P. americana* *COI* GenBank records were included in genetic analyses representing samples from Iran, China, and Korea (Supplementary Tab. S2). Specimen conditions ranged from well preserved to substantially damaged (Fig. 1). *COI* barcodes were successfully recovered from 238 cockroach specimens (including 223 *P. americana* specimens), while either PCR or sequencing failed repeatedly for 46 specimens (Supplementary Tab. S1). Seven sequences were trimmed at one end (range: 3–30 base pairs (bp)) due to low quality signals. The remainder were high-quality, full-length reads of 658 bp ( $N = 231$ ) that contained 43 variable and parsimony-informative sites, as well as three singletons. Among the failed 46 specimens were eight specimens morphologically identified as the German cockroach (*Blattella germanica*). These specimens yielded a short PCR product (approximately 400 bp). Sequencing of one of these short PCR products revealed a putative nuclear pseudogene with a 285 bp internal deletion and multiple amino acid substitutions compared to available *B. germanica* GenBank records, including a complete mitochondrial genome (GenBank EU854321). This aberrant sequence was identical to *B. germanica* *COI* GenBank records KC473901 and KC473904, which thus presumably represent the same pseudogene. Except for that short sequence, no stop codons, unusual amino acid substitutions, or internal sequence deletions were found in any other sequence, indicating that all other sequences were likely functional mitochondrial sequences and not nuclear pseudogenes. In addition to *COI* barcodes, we analyzed a portion of the nuclear gene *wingless* (*wg*) for a subset of specimens. All *wg* amplifications yielded a full length, high quality sequence ( $N = 80$ ; 77 *P. americana* and three *P. fuliginosa* specimens). The *wg* alignment for *P. americana* contained 8 variable and parsimony-informative sites, as well as one singleton.

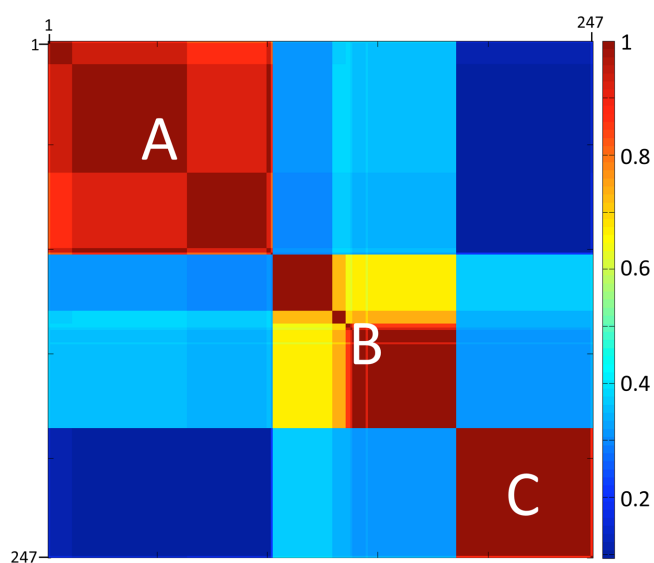


**Figure 1** | Representative *P. americana* specimens, ranging from well preserved (left) to substantially damaged (right). Scale bars are 5 mm.

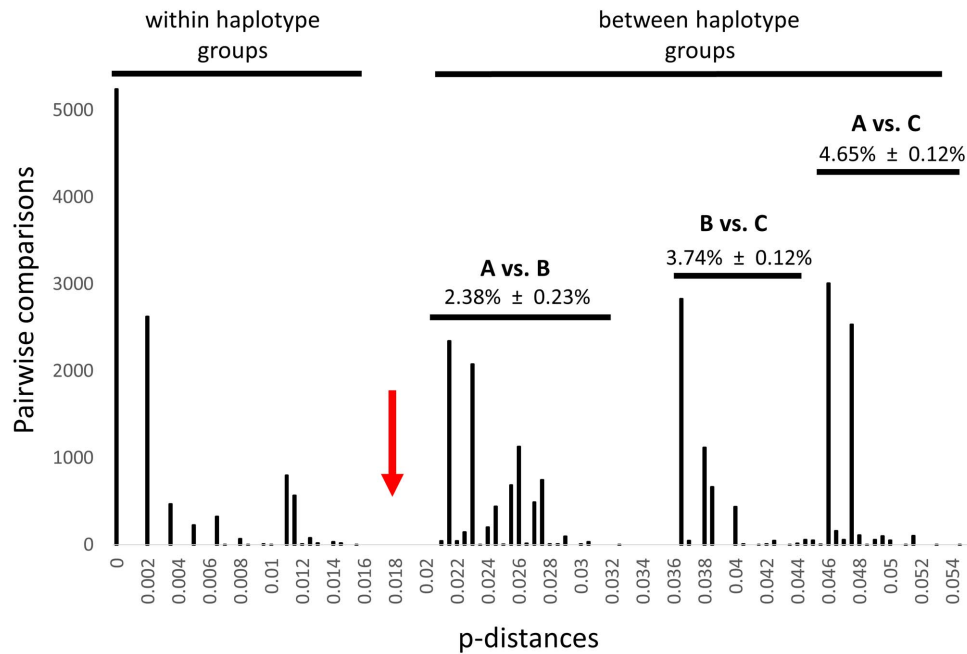
The American cockroach showed an unusual pattern of genetic diversity with three major mitochondrial haplogroups (A, B, C) (Figs. 2, 3). Including the 24 GenBank sequences, the three groups comprised 15 haplotypes, six of which were represented in NYC (Supplementary Tables S1, S2). Mean p-distances between groups were larger (ranging from 2.38% to 4.65%) than within groups (mean p-distance  $\pm$  SD, A: 0.12%  $\pm$  0.12%, B: 0.57%  $\pm$  0.52%, C: 0.03%  $\pm$  0.14%;  $N = 247$  specimens) (Fig. 3). Among NYC specimens, p-distances were very similar to those in the complete dataset (mean p-distances between groups: 2.34%–4.63%; variation within groups, A: 0.10%  $\pm$  0.09%, B: 0.44%  $\pm$  0.52%, C: 0.00%  $\pm$  0%;  $N = 161$  specimens). Thus, apparent ‘barcoding gaps’ existed between maximum intra-haplogroup and minimum inter-haplogroup p-distances (Fig. 3).

Mitochondrial (*COI*) and nuclear (*wg*) phylogenetic trees were highly discordant, i.e., nuclear gene analysis did not recover the three major clades (Fig. 4). Specimens from different *COI* haplogroups shared the same *wg* alleles, and specimens from the same haplogroup carried different *wg* alleles. Overall, we found no genetic differentiation at the *wg* locus among *COI* haplogroups, indicating interbreeding (within NYC population: overall  $F_{ST} = 0.041$ ,  $P = 0.100$ , individuals:  $N_{\text{haplogroup A}} = 20$ ,  $N_{\text{haplogroup B}} = 26$ ,  $N_{\text{haplogroup C}} = 14$ , alleles:  $N_{\text{haplogroup A}} = 7$ ,  $N_{\text{haplogroup B}} = 6$ ,  $N_{\text{haplogroup C}} = 6$ ; pairwise comparisons: haplogroup A vs. B:  $F_{ST} = 0.069$ ,  $P = 0.053$ ; haplogroup A vs. C:  $F_{ST} = 0.045$ ,  $P = 0.13$ ; haplogroup B vs. C:  $F_{ST} = -0.007$ ,  $P = 0.462$ ). The three haplogroups had broadly overlapping geographic distributions, including within NYC, across the U.S., and elsewhere in the world (Fig. 5).

All *P. americana* *COI* haplotypes formed a monophyletic lineage separate from other *Periplaneta* species obtained in this study or represented in GenBank (Supplementary Fig. S1). The closest p-distance was found with *Periplaneta australasiae* (minimum sequence divergence 7.4%). In addition to *P. americana*, four other cockroach species were identified by *COI* barcode in this study, comprising about 6% of our specimens, including Smokybrown cockroach (*P. fuliginosa*) ( $N = 9$ ), Brown-banded cockroach (*Supella*



**Figure 2** | Klee diagram of *COI* barcodes. A Klee diagram is a colored matrix of indicator vector correlations; identical sequences have a correlation of 1<sup>87</sup>. The dataset comprises 247 *P. americana* sequences (including 24 from GenBank). The matrix is arranged according to a NJ tree. Labeled blocks of high correlation along the diagonal correspond to *P. americana* haplogroups A, B, and C. The 15 individual haplotypes appear as sub-boxes. The color code for correlation coefficients is given on the right.

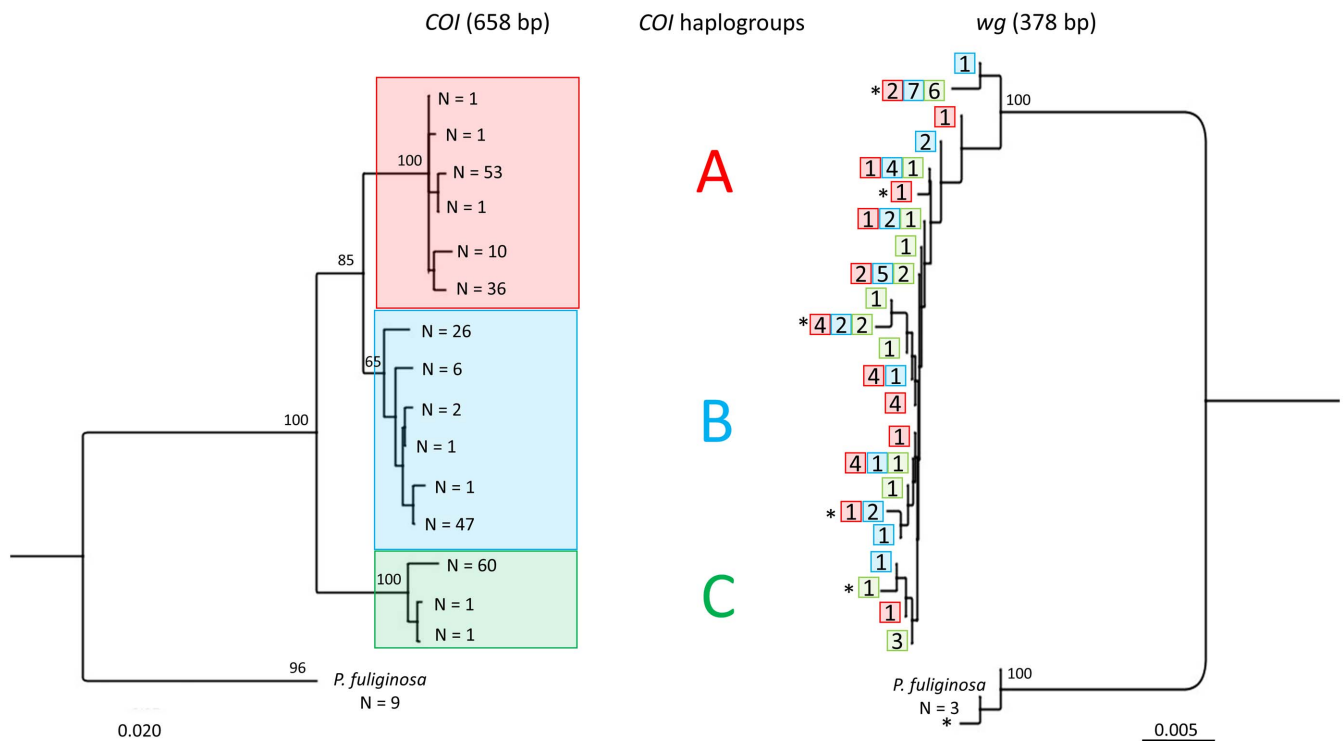


**Figure 3** | Histogram showing intra- and inter-haplogroup p-distances between *P. americana* mitochondrial *COI* sequences (N = 247). An apparent barcode gap (red arrow) separates maximum distances within and minimum distances between haplogroups. Mean  $\pm$  SD p-distances between groups are shown.

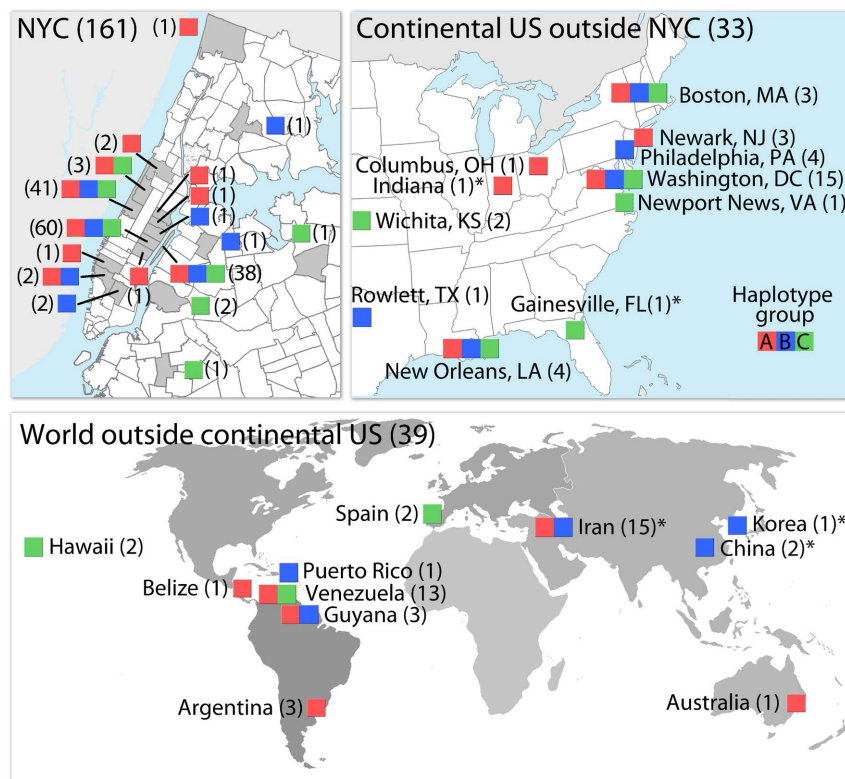
*longipalpa* (N = 2), Turkestan cockroach (*Shelfordella lateralis*) (N = 1), and Madagascar hissing cockroach (*Gromphadorhina portentosa*) (N = 1) (Supplementary Tab. S1). Finally, two specimens, one from the southern U.S. (Georgia) and one from Venezuela, could not be identified by barcode (Supplementary Tab. S1).

## Discussion

This is the first large-scale study of mitochondrial diversity in the American cockroach (*Periplaneta americana*). Surprisingly we found three deeply-divergent, widely-distributed *P. americana* *COI* haplogroups. In a limited geographic area such as NYC, deeply diver-



**Figure 4** | Lack of congruence between mitochondrial and nuclear data. Comparison of NJ trees based on Tamura-Nei distances for mitochondrial (*COI*, N = 256) and nuclear (*wg*, N = 80) DNA sequences of *P. americana* and *P. fuliginosa*. Colors depict different *COI* haplogroups. Numbers on the tree give the sample sizes (number of individuals) and boxes in one row represent individuals with the same allele combination. *Periplaneta fuliginosa* was used as outgroup in both analyses. Bootstrap support for major nodes is shown. Abbreviations: N = number of individuals, \* = homozygous individuals.



**Figure 5 | Distribution of *P. americana* COI haplogroups in NYC, continental U.S., and the world.** Sample sizes per site are given in parentheses, haplogroup colors and designations correspond to those in Figure 4. NYC zip codes and U.S. states are outlined on the maps in the upper left and upper right, respectively. Sequences retrieved from GenBank are marked with asterisks. Maps were created with Adobe Photoshop. Map templates are from d-maps.com (US, [http://d-maps.com/carte.php?num\\_car=5222&lang=en](http://d-maps.com/carte.php?num_car=5222&lang=en); world, [http://d-maps.com/carte.php?num\\_car=3267&lang=en](http://d-maps.com/carte.php?num_car=3267&lang=en)) and U.S. Census Bureau (NYC zip codes, [https://www.census.gov/geo/maps-data/data/cbf/cbf\\_zcta.html](https://www.census.gov/geo/maps-data/data/cbf/cbf_zcta.html)).

gent mtDNA lineages within a single species are unusual. Several scenarios can explain such a genetic pattern, which are not mutually exclusive: historical introgression between species<sup>47–53</sup>, manipulation by endosymbiotic bacteria<sup>54,55</sup>, secondary contact of formerly isolated populations<sup>56,57</sup>, or reproductive barriers among sympatric cryptic species<sup>17,19</sup>. The latter scenario does not apply in our case. The presence of a “barcode gap” between the maximum distances within and minimum distances between haplogroups has often been proposed to indicate cryptic species in native populations<sup>7,16,18,19,58,59</sup> (however, for critical views see refs. 8, 55, 60, 61). Although we observed clear barcode gaps in *P. americana*, the nuclear data are indicative of a single biological species. Cryptic speciation should be reflected in a detectable differentiation at nuclear markers, in particular in such deeply divergent mtDNA clades (up to 4.6% sequence divergence). This was not the case, and our analysis of nuclear sequences revealed extensive gene flow among COI haplogroups.

Discordant phylogenetic signals between maternally and biparentally inherited markers are sometimes due to infection with endosymbiotic bacteria<sup>54,55</sup>. *Wolbachia* bacteria, for example, transmit maternally and manipulate host reproduction in favor of infected females<sup>62,63</sup>. As mitochondria are maternally transmitted as well, selection favors those mtDNA types that are associated with *Wolbachia* infections, which can create unexpected mitochondrial population structures<sup>55</sup>. Vaishampayan and colleagues<sup>64</sup> detected *Wolbachia* infections in cockroaches of the genera *Blattella* and *Supella*, but not in *P. americana*, suggesting that *Wolbachia* infections in *P. americana* might be rare or absent. In general, *Wolbachia* infections will substantially reduce mtDNA diversity in a given population and skew the frequency distribution of alleles towards a single or very few variants (the latter in cases of multiple infections)<sup>55,65,66</sup>. The pattern of mtDNA

diversity detected in this study does, however, not reflect the typical pattern of reduced haplotype diversity found in *Wolbachia*-infected populations. We detected six different, divergent haplotypes in a single population of *P. americana* in NYC, and this general pattern seems to hold for other populations around the globe. Thus, it seems unlikely that *Wolbachia* infections have played a major role in creating the unusual pattern of mtDNA variation detected in this study. However, additional work screening specifically for *Wolbachia* infections will be required to conclusively rule out this possibility. Likewise, there is currently no evidence for historical introgression from other *Periplaneta* species. All available *P. americana* COI barcodes formed a monophyletic lineage clearly separated from congeneric species (Supplementary Fig. S1). Future studies may, however, uncover overlap in mitochondrial haplotypes with other species not yet represented in databases.

Currently, the most likely explanation for the detected genetic pattern is multiple human-mediated introductions from allopatric source populations followed by global dispersal among commercial centers. In fact, the different haplogroups must have diverged long before human-aided dispersal, even if the highest mutation rate estimates of insect mtDNA are applied (10–20% sequence divergence per million years; see Papadopoulou et al. 2010 for a review on mtDNA clocks in insects<sup>67</sup>). *Periplaneta americana*'s ability to inhabit human-built structures<sup>33,34</sup> has probably facilitated its introduction to new areas. In general, human-mediated transport creates many opportunities for introduction and interbreeding of previously isolated species or populations<sup>1,2</sup>. For example, Ruddy Ducks artificially introduced to the UK from North America hybridize freely with the indigenous White-headed Duck, effectively threatening extinction of the native form<sup>68</sup>. In the present case, it appears that



*P. americana* individuals from three or more historically isolated geographic populations are now effectively merged into a single global gene pool.

Invasive species often represent multiple introductions from genetically distinct source populations, and interbreeding may both be common and critical for long-term invasion success<sup>1,26,31,32,69–72</sup>. For example, interbreeding between distinct genetic lineages in the multicolored ladybird beetle (*Harmonia axyridis*) led to shifts in key life history traits enabling invasion success<sup>73,74</sup>. Conversely, invasive populations are often derived from small founding populations, and this genetic bottleneck can entail a substantial reduction in genetic diversity and lead to inbreeding depression<sup>1</sup>. The success of invasive species in their non-native range might be temporary in many cases, because prolonged inbreeding generally leads to a decrease in fitness<sup>75–77</sup> and, possibly, population extinctions<sup>78,79</sup>. For example, inbreeding depression has been suggested as one mechanism underlying the collapse of New Zealand populations of one of the world's worst invasive pests, the Argentine ant (*Linepithema humile*)<sup>80</sup>. Empirical research on the influence of multiple introduction events on invasion success has just begun<sup>1</sup>, but examples suggest that preventing repeated introductions may reduce adaptive potential in some cases<sup>32</sup> and thus may facilitate the long-term control of seemingly well-established invasive pests.

The spread of invasive species and thus of pest species like *P. americana* is expected to be facilitated and intensified by globalization. Our case study shows that species delimitation using a classical DNA barcoding approach can be particularly problematic when studying populations of invasive species, because existing populations might represent a mix of individuals from genetically distinct source populations. In such cases, it will be especially important to supplement mitochondrial DNA barcoding with nuclear genetic and/or morphological data. Nonetheless, with the comparatively large reference library provided in this study, DNA barcoding now has great potential to be employed as a quick and reliable tool to identify urban populations of the American cockroach.

## Methods

**Specimen collection and processing.** To obtain specimens of the American cockroach, we set up a citizen science project based in New York City (NYC). The project was launched between November 2012 and January 2014 starting with a website including participant instructions (<http://phe.rockefeller.edu/barcode/cockroachproject.html>). We sought publicity via word of mouth and social media, as well as traditional outlets including television, newspapers, and radio. Contributors were asked to provide specimen collection date and location, and their name and contact information, and to send unpreserved dead specimens via regular mail.

Each specimen was assigned a code, transferred to an individual 50 mL Falcon vial, and stored at  $-30^{\circ}\text{C}$  until further processing. Collection information was recorded in a spreadsheet (Supplementary Tab. S1). Six well-preserved *P. americana* specimens from each haplogroup (A, B, and C) were morphologically identified using the key of Helfer<sup>81</sup>. These samples are stored at the American Museum of Natural History (see Supplementary Tab. S1 for unique specimen identifiers). Other cockroach specimens were identified to species via DNA barcodes. *COI* barcode sequences were available in GenBank for most of the common urban roaches (see Supplementary Tab. S2 for reference sequences), allowing us to assign species names even to morphologically unrecognizable specimens.

**Genetic analysis.** Genetic analysis was performed on 283 specimens (Supplementary Tab. S1). DNA was extracted from cockroach leg fragments using the QIAGEN® DNeasy® kit and stored at  $-30^{\circ}\text{C}$ . The mitochondrial *COI* barcode region (658 base pairs) and a portion of the nuclear gene *wingless* (*wg*; 378 bp) were amplified in standard polymerase chain reactions (PCRs) using the primers LCO1490/HCO2198<sup>82</sup> and *wg*550F<sup>83</sup>/*wg*cockR, respectively. The primer *wg*cockR (sequence: 5' AACATGCACGCACACCTCTGCACCACGGACACC 3') was designed specifically for this study, because longer fragments using primers *wg*550F/*wg*AbrZ<sup>83</sup> did not amplify consistently. PCRs were set up in 25  $\mu\text{l}$  reaction volumes containing 14.3  $\mu\text{l}$  AccuGENE® water, 0.2  $\mu\text{l}$  AmpliTaq Gold®, 2.5  $\mu\text{l}$  10 $\times$  PCR buffer, 2.5  $\mu\text{l}$  MgCl<sub>2</sub> (25 mM), 2.5  $\mu\text{l}$  dNTPs (2 mM each), 2  $\mu\text{l}$  DNA template and 0.5  $\mu\text{l}$  of each of the respective primers (10  $\mu\text{M}$  each). An initial denaturation step of 5 min at  $95^{\circ}\text{C}$  was followed by 40 cycles ( $95^{\circ}\text{C}$  for 40 s;  $55^{\circ}\text{C}$  (*COI*) or  $64^{\circ}\text{C}$  (*wg*) for 40 s;  $72^{\circ}\text{C}$  for 40 s) and a final extension of 15 min at  $72^{\circ}\text{C}$  using an Eppendorf Mastercycler® proS. Purification and sequencing of PCR products was performed by commercial facilities (Macrogen USA or Eton Bioscience). All PCR products were sequenced in both directions. Sequences are deposited under accession numbers KM576918–

KM577157 (*COI*) and KM591602–KM591681 (*wg*) in GenBank (Supplementary Tab. S1).

**Data management and analysis.** The laboratory information management system (LIMS) implemented in the software Geneious® (version 7.1.7) with the biocode plugin (version 2.8.0) was used to track all workflows including collection data, extractions, PCRs, and cycle sequencing<sup>84</sup>. Sequences were aligned and trimmed in Geneious®, Neighbor-joining trees based on Tamura-Nei distances with bootstrap support (1,000 replicates) were created using Geneious® Tree Builder and a Randomized Axelerated Maximum Likelihood<sup>85</sup> tree (RAXML; version 8.1.X) was created using the GTR + gamma model. MEGA 6<sup>86</sup> was used to assess number of variable, parsimony-informative sites and singletons and to calculate p-distances with gaps deleted in pairwise comparisons. Tree-Parser-aided Klee diagrams were created as described in Stoeckle and Coffran (2013)<sup>87</sup>. The level of genetic differentiation for the nuclear gene *wg* among different *COI* haplogroups was calculated as pairwise  $F_{ST}$  in the software FSTAT 2.9.3.2<sup>88</sup> following Weir and Cockerham (1984)<sup>89</sup>. Statistical significance assuming Hardy-Weinberg equilibrium was assessed using randomization tests with 10,000 iterations.

Five *wg* alleles segregating in the population were directly observed in homozygous individuals. Three additional alleles were inferred from heterozygous individuals by subtracting one of the five alleles observed in the homozygous state. In the vast majority of heterozygous individuals, the chromatogram could only be explained by a single combination of two of these eight alleles. Only in two heterozygous individuals was there ambiguity over the respective allele combination, and therefore we omitted those individuals from the  $F_{ST}$  analysis.

1. Simberloff, D. *Invasive Species: What everyone needs to know*. (Oxford University Press, New York, 2013).
2. Hulme, P. E. Trade, transport and trouble: managing invasive species pathways in an era of globalization. *J. Appl. Ecol.* **46**, 10–18, doi:10.1111/j.1365-2664.2008.01600.x (2009).
3. Armstrong, K. F. & Ball, S. L. DNA barcodes for biosecurity: invasive species identification. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **360**, 1813–1823, doi:10.1098/rstb.2005.1713 (2005).
4. Darling, J. A. & Blum, M. J. DNA-based methods for monitoring invasive species: a review and prospectus. *Biol. Invasions* **9**, 751–765, doi:10.1007/s10530-006-9079-4 (2007).
5. Rubinoff, D. Utility of mitochondrial DNA barcodes in species conservation. *Conserv. Biol.* **20**, 1026–1033, doi:10.1111/j.1523-1739.2006.00372.x (2006).
6. Krishnamurthy, P. K. & Francis, R. A. A critical review on the utility of DNA barcoding in biodiversity conservation. *Biodivers. Conserv.* **21**, 1901–1919, doi:10.1007/s10531-012-0306-2 (2012).
7. Meier, R., Zhang, G. & Ali, F. The use of mean instead of smallest interspecific distances exaggerates the size of the “Barcoding Gap” and leads to misidentification. *Syst. Biol.* **57**, 809–813, doi:10.1080/10635150802406343 (2008).
8. Wiemers, M. & Fiedler, K. Does the DNA barcoding gap exist? - a case study in blue butterflies (Lepidoptera: Lycaenidae). *Front. Zool.* **4**, 8–8, doi:10.1186/1742-9994-4-8 (2007).
9. Waugh, J. DNA barcoding in animal species: progress, potential and pitfalls. *Bioessays* **29**, 188–197, doi:10.1002/bies.20529 (2007).
10. Hebert, P. D. N., Ratnasingham, S. & deWaard, J. R. Barcoding animal life: *cytochrome c oxidase* subunit 1 divergences among closely related species. *Proc. Roy. Soc. B* **270**, S96–S99, doi:10.1098/rsbl.2003.0025 (2003).
11. Hajibabaei, M., Singer, G. A. C., Hebert, P. D. N. & Hickey, D. A. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *Trends Genet.* **23**, 167–172 (2007).
12. Goldstein, P. Z. & DeSalle, R. Integrating DNA barcode data and taxonomic practice: Determination, discovery, and description. *Bioessays* **33**, 135–147, doi:10.1002/bies.201000036 (2011).
13. Jinbo, U., Kato, T. & Ito, M. Current progress in DNA barcoding and future implications for entomology. *Entomol. Sci.* **14**, 107–124, doi:10.1111/j.1479-8298.2011.00449.x (2011).
14. Ratnasingham, S. & Hebert, P. D. N. BOLD: The Barcode of Life Data System ([www.barcodinglife.org](http://www.barcodinglife.org)) *Mol. Ecol. Notes* **7**, 355–364, doi:10.1111/j.1471-8286.2006.01678.x (2007).
15. Ratnasingham, S. & Hebert, P. D. N. BOLD's role in barcode data management and analysis: a response. *Mol. Ecol. Res.* **11**, 941–942, doi:10.1111/j.1755-0998.2011.03067.x (2011).
16. Hebert, P. D. N., Stoeckle, M. Y., Zemlak, T. S. & Francis, C. M. Identification of birds through DNA barcodes. *PLoS Biol.* **2**, 1657–1663, doi:10.1371/journal.pbio.0020312 (2004).
17. Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. & Hallwachs, W. Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astrartes fulgurator*. *PNAS* **101**, 14812–14817, doi:10.1073/pnas.0406166101 (2004).
18. Hajibabaei, M., Janzen, D. H., Burns, J. M., Hallwachs, W. & Hebert, P. D. N. DNA barcodes distinguish species of tropical Lepidoptera. *PNAS* **103**, 968–971, doi:10.1073/pnas.0510466103 (2006).



19. Smith, M. A. *et al.* Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *PNAS* **105**, 12359–12364, doi:10.1073/pnas.0805319105 (2008).
20. Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R. & Hebert, P. D. N. DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **360**, 1847–1857, doi:10.1098/rstb.2005.1716 (2005).
21. Resch, M. C. *et al.* Where taxonomy based on subtle morphological differences is perfectly mirrored by huge genetic distances: DNA Barcoding in protura (Hexapoda). *PLoS One* **9**, doi:10.1371/journal.pone.0090653 (2014).
22. Scheffer, S. J., Lewis, M. L. & Joshi, R. C. DNA barcoding applied to invasive leafminers (Diptera: Agromyzidae) in the Philippines. *Ann. Entomol. Soc. Am.* **99**, 204–210, doi:10.1603/0013-8746(2006)099[0204:dbatil]2.0.co;2 (2006).
23. Smith, P. J., Webber, W. R., McVeagh, S. M., Inglis, G. J. & Gust, N. DNA and morphological identification of an invasive swimming crab, *Charybdis japonica*, in New Zealand waters. *New Zeal. J. Mar. Fresh.* **37**, 753–762 (2003).
24. Ficetola, G. F., Bonin, A. & Miaud, C. Population genetics reveals origin and number of founders in a biological invasion. *Mol. Ecol.* **17**, 773–782 (2008).
25. Saunders, G. W. Routine DNA barcoding of Canadian Gracilariiales (Rhodophyta) reveals the invasive species *Gracilaria vermiculophylla* in British Columbia. *Mol. Ecol. Res.* **9**, 140–150, doi:10.1111/j.1755-0998.2009.02639.x (2009).
26. Facon, B. *et al.* A molecular phylogeography approach to biological invasions of the New World by parthenogenetic Thiarid snails. *Mol. Ecol.* **12**, 3027–3039, doi:10.1046/j.1365-294X.2003.01972.x (2003).
27. Marescaux, J. & Van Doninck, K. Using DNA barcoding to differentiate invasive *Dreissena* species (Mollusca, Bivalvia). *Zookeys*, 235–244, doi:10.3897/zookeys.365.5905 (2013).
28. Simonsen, T. J., Brown, R. L. & Sperling, F. A. H. Tracing an invasion: Phylogeography of *Cactoblastis cactorum* (Lepidoptera: Pyralidae) in the United States based on mitochondrial DNA. *Ann. Entomol. Soc. Am.* **101**, 899–905, doi:10.1603/0013-8746(2008)101[899:taipoc]2.0.co;2 (2008).
29. Nadel, R. L. *et al.* DNA bar-coding reveals source and patterns of *Thaumastocoris peregrinus* invasions in South Africa and South America. *Biol. Invasions* **12**, 1067–1077, doi:10.1007/s10530-009-9524-2 (2010).
30. Rubinoff, D., Holland, B. S., Shibata, A., Messing, R. H. & Wright, M. G. Rapid invasion despite lack of genetic variation in the erythrina gall wasp (*Quadrastichus erythrinae* Kim). *Pac. Sci.* **64**, 23–31, doi:10.2984/64.1.023 (2010).
31. Roman, J. & Darling, J. A. Paradox lost: genetic diversity and the success of aquatic invasions. *Tr. Ecol. Evol.* **22**, 454–464, doi:10.1016/j.tree.2007.07.002 (2007).
32. Dlugosch, K. M. & Parker, I. M. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol. Ecol.* **17**, 431–449, doi:10.1111/j.1365-294X.2007.03538.x (2008).
33. Bell, W., Roth, L. & Nalepa, C. *Cockroaches: Ecology, behavior, and natural history*. (The Johns Hopkins University Press, Baltimore, 2007).
34. Guthrie, D. & Tindall, A. *The biology of the cockroach*. (St. Martin's Press, New York, 1968).
35. Bell, W. J. & Adiyodi, K. G. *The American cockroach*. (Chapman and Hall Ltd, New York 1981).
36. Rehn, J. A. G. Man's uninvited fellow travellers – the cockroach. *Sci. Mon.* **61**, 265–276 (1945).
37. Bonnefoy, X., Kampen, H. & Sweeney, K. *Public Health Significance Of Urban Pests*. (World Health Organization, Copenhagen, 2008).
38. Appel, A. G. & Smith, L. M. Biology and management of the smokybrown cockroach. *Ann. Rev. Entomol.* **47**, 33–55, doi:10.1146/annurev.ento.47.091201.145106 (2002).
39. Iwuala, M. O. E., Osisiogu, I. U. W. & Agbakwuru, E. O. P. Dennetia oil, a potential new insecticide: tests with adults and nymphs of *Periplaneta americana* and *Zonocerus variegatus*. *J. Econ. Entomol.* **74**, 249–52 (1981).
40. Schal, C. & Hamilton, R. L. Integrated suppression of synanthropic cockroaches. *Ann. Rev. Entomol.* **35**, 521–551 (1990).
41. Shahraki, G., Parhizkar, S. & Nejad, A. Cockroach infestation and factors affecting the estimation of cockroach population in urban communities. *Internat. J. Zool.* doi:10.1155/2013/649089 (2013).
42. Gade, G. & Goldsworthy, G. J. Insect peptide hormones: a selective review of their physiology and potential application for pest control. *Pest Manag. Sci.* **59**, 1063–1075, doi:10.1002/ps.755 (2003).
43. Lihoreau, M., Costa, J. T. & Rivault, C. The social biology of domiciliary cockroaches: colony structure, kin recognition and collective decisions. *Insect. Soc.* **59**, 445–452, doi:10.1007/s00040-012-0234-x (2012).
44. Whyatt, R. M. *et al.* Residential pesticide use during pregnancy among a cohort of urban minority women. *Environ. Health Perspect.* **110**, 507–514 (2002).
45. Evangelista, D., Buss, L. & Ware, J. Using DNA barcodes to confirm the presence of a new invasive cockroach pest in New York City. *J. Eco. Entomol.* **106**, 2275–2279 (2013).
46. Brenner, B. L. *et al.* Integrated pest management in an urban community: A successful partnership for prevention. *Environ. Health Perspect.* **111**, 1649–1653, doi:10.1289/ehp.6069 (2003).
47. Marshall, D. C., Hill, K. B. R., Cooley, J. R. & Simon, C. Hybridization, mitochondrial DNA phylogeography, and prediction of the early stages of reproductive isolation: Lessons from New Zealand cicadas (genus *Kikihia*). *Syst. Biol.* **60**, 482–502, doi:10.1093/sysbio/syr017 (2011).
48. Zhang, A.-B. & Sota, T. Nuclear gene sequences resolve species phylogeny and mitochondrial introgression in *Leptocarabus* beetles showing trans-species polymorphisms. *Mol. Phyl. Evol.* **45**, 534–546, doi:10.1016/j.ympev.2007.07.003 (2007).
49. McGuire, J. A. *et al.* Mitochondrial introgression and incomplete lineage sorting through space and time: Phylogenetics of crotaphytid lizards. *Evolution* **61**, 2879–2897, doi:10.1111/j.1558-5646.2007.00239.x (2007).
50. Whitworth, T. L., Dawson, R. D., Magalon, H. & Baudry, E. DNA barcoding cannot reliably identify species of the blowfly genus *Procalliphora* (Diptera: Calliphoridae). *Proc. Roy. Soc. B* **274**, 1731–1739, doi:10.1098/rspb.2007.0062 (2007).
51. Vences, M., Thomas, M., Bonett, R. M. & Vieites, D. R. Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Phil. Trans. Roy. Soc. B* **360**, 1859–1868, doi:10.1098/rstb.2005.1717 (2005).
52. Moritz, C., Dowling, T. E. & Brown, W. M. Evolution of animal mitochondrial-DNA - Relevance for population biology and systematics. *Ann. Rev. Ecol. Syst.* **18**, 269–292, doi:10.1146/annurev.es.18.110187.001413 (1987).
53. Ballard, J. W. O. & Whitlock, M. C. The incomplete natural history of mitochondria. *Mol. Ecol.* **13**, 729–744, doi:10.1046/j.1365-294X.2003.02063.x (2004).
54. Ritter, S. *et al.* *Wolbachia* infections mimic cryptic speciation in two parasitic butterfly species, *Phengaris teleius* and *P. nausithous* (Lepidoptera: Lycaenidae). *PLoS One* **8**, doi:10.1371/journal.pone.0078107 (2013).
55. Hurst, G. D. D. & Jiggins, F. M. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc. Roy. Soc. B* **272**, 1525–1534, doi:10.1098/rspb.2005.3056 (2005).
56. Munoz, A. G., Baxter, S. W., Linares, M. & Jiggins, C. D. Deep mitochondrial divergence within a *Heliconius* butterfly species is not explained by cryptic speciation or endosymbiotic bacteria. *Bmc Evol. Biol.* **11**, doi:10.1186/1471-2148-11-358 (2011).
57. Zamudio, K. R. & Savage, W. K. Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution* **57**, 1631–1652 (2003).
58. Candek, K. & Kuntner, M. DNA barcoding gap: reliable species identification over morphological and geographical scales. *Mol. Ecol. Res.* doi:doi: 10.1111/1755-0998.12304 (2014).
59. Pereira, L. H. G., Hanner, R., Foresti, F. & Oliveira, C. Can DNA barcoding accurately discriminate megadiverse Neotropical freshwater fish fauna? *Bmc Genet.* **14**, doi:10.1186/1471-2156-14-20 (2013).
60. Burns, J. M., Janzen, D. H., Hajibabaei, M., Hallwachs, W. & Hebert, P. D. N. DNA barcodes of closely related (but morphologically and ecologically distinct) species of skipper butterflies (Hesperiidae) can differ by only one to three nucleotides. *J. Lepidopt. Soc.* **61**, 138–153 (2007).
61. Elias, M. *et al.* Limited performance of DNA barcoding in a diverse community of tropical butterflies. *Proc. Roy. Soc. B* **274**, 2881–2889, doi:10.1098/rspb.2007.1035 (2007).
62. Serbus, L. R., Casper-Lindley, C., Landmann, F. & Sullivan, W. The genetics and cell biology of *Wolbachia*-host interactions. *Ann. Rev. Genet.* **42**, 683–707, doi:10.1146/annurev.genet.41.110306.130354 (2008).
63. Russell, J. A. The ants (Hymenoptera: Formicidae) are unique and enigmatic hosts of prevalent *Wolbachia* (Alphaproteobacteria) symbionts. *Myrmecol. News* **16**, 7–23 (2012).
64. Vaishampayan, P. A. *et al.* Molecular evidence and phylogenetic affiliations of *Wolbachia* in cockroaches. *Mol. Phylogenet. Evol.* **44**, 1346–1351, doi:10.1016/j.ympev.2007.01.003 (2007).
65. Jiggins, F. M. Male-killing *Wolbachia* and mitochondrial DNA: Selective sweeps, hybrid introgression and parasite population dynamics. *Genetics* **164**, 5–12 (2003).
66. Xiao, J.-H. *et al.* *Wolbachia* infection and dramatic intraspecific mitochondrial DNA divergence in a fig wasp. *Evolution* **66**, 1907–1916, doi:10.1111/j.1558-5646.2011.01561.x (2012).
67. Papadopoulou, A., Anastasiou, I. & Vogler, A. P. Revisiting the insect mitochondrial molecular clock: The mid-Aegean trench calibration. *Mol. Biol. Evol.* **27**, 1659–1672, doi:10.1093/molbev/msq051 (2010).
68. Henderson, I. & Robertson, P. Control and eradication of the North American ruddy duck in Europe. *Manag. Vert. Invas. Sp.* paper 16, <http://digitalcommons.unl.edu/nwrcinvasive/16> (2007).
69. Frankham, R. Invasion biology - Resolving the genetic paradox in invasive species. *Hereditas* **94**, 385–385, doi:10.1038/sj.hdy.6800634 (2005).
70. Simberloff, D. The role of propagule pressure in biological invasions. *Ann. Rev. Ecol. Evol. Syst.* **40**, 81–102, doi:10.1146/annurev.ecolsys.110308.120304 (2009).
71. Lavergne, S. & Molofsky, J. Increased genetic variation and evolutionary potential drive the success of an invasive grass. *PNAS* **104**, 3883–3888, doi:10.1073/pnas.0607324104 (2007).
72. Verhoeven, K. J. F., Macel, M., Wolfe, L. M. & Biere, A. Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proc. Roy. Soc. B* **278**, 2–8, doi:10.1098/rspb.2010.1272 (2011).
73. Facon, B. *et al.* Can things get worse when an invasive species hybridizes? The harlequin ladybird *Harmonia axyridis* in France as a case study. *Evol. Appl.* **4**, 71–88, doi:10.1111/j.1752-4571.2010.00134.x (2011).
74. Turgeon, J. *et al.* Experimental evidence for the phenotypic impact of admixture between wild and biocontrol Asian ladybird (*Harmonia axyridis*) involved in the European invasion. *J. Evol. Biol.* **24**, 1044–1052, doi:10.1111/j.1420-9101.2011.02234.x (2011).



75. Keller, L. F. & Waller, D. M. Inbreeding effects in wild populations. *Tr. Ecol Evol.* **17**, 230–241, doi:10.1016/s0169-5347(02)02489-8 (2002).
76. Crnokrak, P. & Roff, D. A. Inbreeding depression in the wild. *Heredity* **83**, 260–270, doi:10.1038/sj.hdy.6885530 (1999).
77. Charlesworth, D. & Charlesworth, B. Inbreeding depression and its evolutionary consequences. *Ann. Rev. Ecol. Syst.* **18**, 237–268, doi:10.1146/annurev.ecolsys.18.1.237 (1987).
78. Frankham, R. & Ralls, K. Conservation biology - Inbreeding leads to extinction. *Nature* **392**, 441–442, doi:10.1038/33022 (1998).
79. Saccheri, I. *et al.* Inbreeding and extinction in a butterfly metapopulation. *Nature* **392**, 491–494, doi:10.1038/33136 (1998).
80. Cooling, M., Hartley, S., Sim, D. A. & Lester, P. J. The widespread collapse of an invasive species: Argentine ants (*Linepithema humile*) in New Zealand. *Biol. Lett.* **8**(3), 430–433, doi:10.1098/rsbl.2011.1014 (2012).
81. Helfer, J. R. *How to Know the Grasshoppers, Crickets, Cockroaches and Their Allies*. [35–59] (Dover Publications, New York, 1987).
82. Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotech.* **3**, 294–299 (1994).
83. Wild, A. L. & Maddison, D. R. Evaluating nuclear protein-coding genes for phylogenetic utility in beetles. *Mol. Phylogen. Evol.* **48**, 877–891, doi:10.1016/j.ympev.2008.05.023 (2008).
84. Parker, M., Stones-Havas, S., Starger, C. & Meyer, C. Laboratory information management systems for DNA barcoding. *Meth. Mol. Biol.* **858**, 269–310, doi:10.1007/978-1-61779-591-6\_13 (2012).
85. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313, doi:10.1093/bioinformatics/btu033 (2014).
86. Tamura, K., Stecher, G., Peterson, D., Filipski, A. & Kumar, S. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729, doi:10.1093/molbev/mst197 (2013).
87. Stoeckle, M. Y. & Coffran, C. TreeParser-aided Klee diagrams display taxonomic clusters in DNA barcode and nuclear gene datasets. *Sci. Rep.* **3**, doi:10.1038/srep02635 (2013).
88. Goudet, J. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3.2) Available at: <http://www2.unil.ch/popgen/softwares/fstat.htm>(2001). Date of access: 05/05/2014.
89. Weir, B. S. & Cockerham, C. C. Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370, doi:10.2307/2408641 (1984).

## Acknowledgments

We are grateful for the contributions of citizen scientists. We thank Jesse Ausubel for comments on an earlier version of this manuscript, Sean McKenzie for help with phylogenetic analyses, Dominic Evangelista for help with identification keys, and Randall Schuh for storage of voucher specimens at the American Museum of Natural History. Special thanks go to our youngest citizen scientist Angel.

## Author contributions

C.v.B., M.Y.S. and D.J.C.K. designed research; C.v.B., J.X. and G.B. performed research; C.v.B., M.Y.S. and D.J.C.K. analyzed data and wrote the paper.

## Additional information

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** von Beeren, C., Stoeckle, M.Y., Xia, J., Burke, G. & Kronauer, D.J.C. Interbreeding among deeply divergent mitochondrial lineages in the American cockroach (*Periplaneta americana*). *Sci. Rep.* **5**, 8297; DOI:10.1038/srep08297 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>