OKINAWA INSTITUTE OF SCIENCE AND TECHNOLOGY GRADUATE UNIVERSITY
沖縄科学技術大学院大学

Museum Genomics Confirms that the Lord Howe Island Stick Insect Survived Extinction

Author	Alexander S. Mikheyev, Andreas Zwick, Michael J.L. Magrath, Miguel L. Grau, Lijun Qiu, You Ning Su, David Yeates
journal or publication title	Current Biology
volume	27
number	20
page range	3157-3161
year	2017-10-23
Publisher	Elsevier Ltd.
Rights	(C) 2017 Elsevier Ltd.
Author's flag	publisher
URL	http://id.nii.ac.jp/1394/00000280/

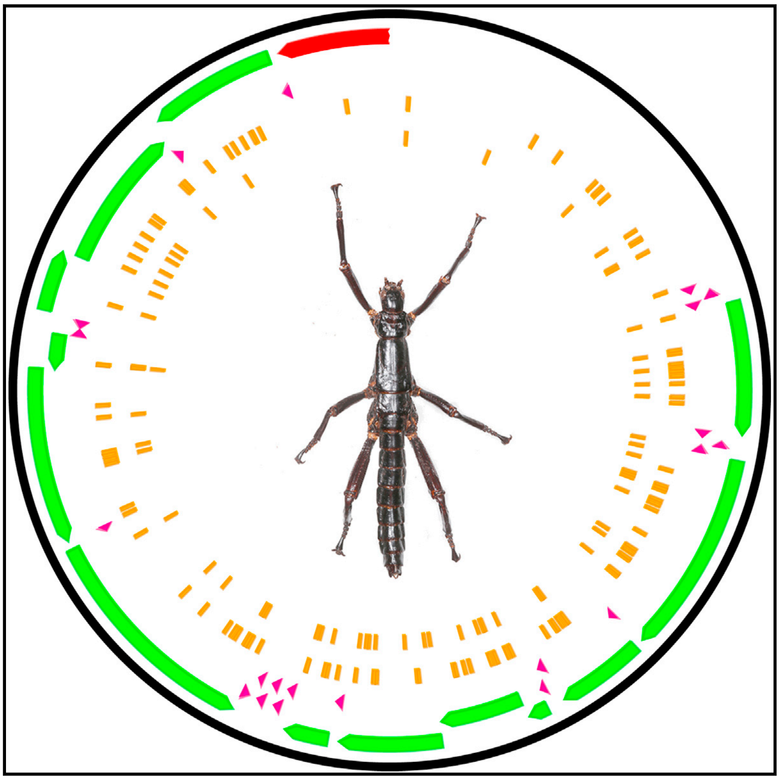
doi: info:doi/10.1016/j.cub.2017.08.058

Current Biology

Report

Museum Genomics Confirms that the Lord Howe Island Stick Insect Survived Extinction

Graphical Abstract



Authors

Alexander S. Mikheyev,
Andreas Zwick, Michael J.L. Magrath,
Miguel L. Grau, Lijun Qiu, You Ning Su,
David Yeates

Correspondence

sasha@homologo.us

In Brief

The largest insect driven to extinction in recent history was the flightless Lord Howe Island stick insect. In 2001, a small population of similar insects was discovered nearby, and preparations are already underway for reintroduction. Using DNA from museum collections, Mikheyev et al. show that the rediscovered population is indeed the same species.

Highlights

- Mitogenomic data confirm that the Lord Howe Island stick insect escaped extinction
- *D. australis* has a massive, most likely polyploid, genome over 4 Gb in size
- The recently discovered population should be suitable for reintroduction

Mikheyev et al., 2017, Current Biology *27*, 3157–3161

http://dx.doi.org/10.1016/j.cub.2017.08.058

Current Biology
Report

Museum Genomics Confirms that the Lord Howe Island Stick Insect Survived Extinction

Alexander S. Mikheyev,[1,2,5,*] Andreas Zwick,[3] Michael J.L. Magrath,[4] Miguel L. Grau,[1] Lijun Qiu,[1] You Ning Su,[3] and David Yeates[3]

[1]Ecology and Evolution Unit, Okinawa Institute of Science and Technology, Tancha 1919-1, Onna-son, Okinawa, 904-0495 Japan
[2]Research School of Biology, Australian National University, Canberra, ACT 2601, Australia
[3]Australian National Insect Collection, CSIRO, Clunies Ross St., Acton, ACT 2601, Australia
[4]Wildlife Conservation and Science, Zoos Victoria, Elliott Avenue, Parkville, VIC 3052, Australia
[5]Lead Contact
*Correspondence: sasha@homologo.us
http://dx.doi.org/10.1016/j.cub.2017.08.058

SUMMARY

The Lord Howe Island stick insect, *Dryococelus australis*, was once common on the island but was driven to extinction after the arrival of ship rats in the early 20th century [1, 2]. It was thought to be extinct for decades, until a tiny population of similar-looking stick insects was discovered 20 km away, on the islet of Ball's Pyramid, in 2001 [2]. Individuals from this population are currently being reared in Australia and elsewhere in the world, with the eventual goal of recolonizing Lord Howe Island [3]. Recent surveys of the wild population on Ball's Pyramid suggest that it is among the world's rarest species. However, there are significant morphological differences between Ball's Pyramid and museum specimens, and there has never been a genetic confirmation of the rediscovered population's species identity. Because *Dryococelus* is monotypic, there are also no known extant relatives for comparison. Using shotgun genomic data from the Ball's Pyramid population, we assembled a draft genome and the complete mitochondrial genome. We found that the genome is massive, over 4 Gb in size, and is most likely hexaploid. We re-sequenced mitochondrial genomes from historic museum specimens collected on Lord Howe Island before the extinction event. Sequence divergence between the two populations is less than 1% and is within the range of intraspecific differences between the museum specimens, suggesting that they are conspecific and that *D. australis* has successfully evaded extinction so far. This work highlights the importance of museum collections for taxonomic validation in the context of ongoing conservation efforts.

RESULTS AND DISCUSSION

Worldwide, only around 60 insect species have been recorded as recently extinct [4], but hundreds have disappeared, and populations of many more are in steep decline [5]. Reintroduction efforts aim to restore species to previously occupied habitat or to reinforce numbers in existing populations, and these have most commonly been applied to vertebrates and plants [6]. In these cases, part of the risk management strategy is to determine whether the introduced populations or species are appropriate for translocation. However, such efforts are frequently hindered by uncertain taxonomy, and novel methods relying on museum specimens can be necessary to establish species identity [7].

Although the discovery of giant stick insects on Ball's Pyramid made it probable that *D. australis* has indeed been rediscovered [2], there are several compelling reasons why the species identification needs to be verified genetically. First, in the paper announcing the rediscovery, Priddel et al. [2] acknowledged that Ball's Pyramid specimens were morphologically distinct from those found on Lord Howe Island (see Figure 1). A later morphometric analysis of captive-reared and museum insects likewise found significant differences, though it remained unclear whether they were due to genetics or environment [8]. Second, surprisingly, the giant "tree lobster" phenotype, characterized by flightlessness and a stocky dorsoventrally flattened body with square-edged thoracic segments, has evolved repeatedly, raising the possibility of morphological convergence [9]. Finally, the Ball's Pyramid population may have had a common origin with that on Lord Howe but isolated for an extended period of time. All of these scenarios could potentially complicate conservation efforts planned on Lord Howe Island.

Genome Sequencing and Resequencing

Unfortunately, no genetic resources existed for this species, and species verification was technically challenging, given the massive genome size of related stick insects [10] and the absence of closely related genomic reference material [9]. To remedy this situation, we used fresh material from the captive-bred population at the Melbourne Zoo to assemble nuclear and mitochondrial genomes for *D. australis*. The bioinformatic nuclear genome size estimate was indeed massive, at 4.2 Gb, and the assembly contained 3.4 Gb, with a contig N50 of 17,265 bp. The mitochondrial genome assembly was 16,604 bp long and was not substantially different from other phasmid genomes in length or composition (Figure 2) [11]. We

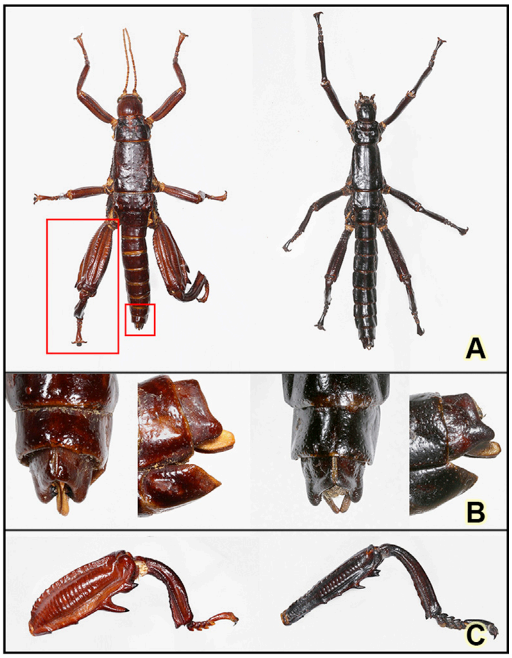


Figure 1. Morphological Differences between Males from Lord Howe Island and Ball's Pyramid

Museum specimens from Lord Howe Island are shown on the left, and captive-reared specimens originally from Ball's Pyramid are shown on the right. Specimens from Lord Howe Island tend to be more robust (A) and have larger femoral spines (C), as well as differences in cercal morphology (B). Although coloration of the specimens also differs, this is most likely a consequence of aging. Because no congeneric taxa exist, it is unclear whether these differences correspond to species-level differentiation or are merely a reflection of environmental or ontogenetic differences, necessitating a genetic investigation of how the two populations relate to each other.

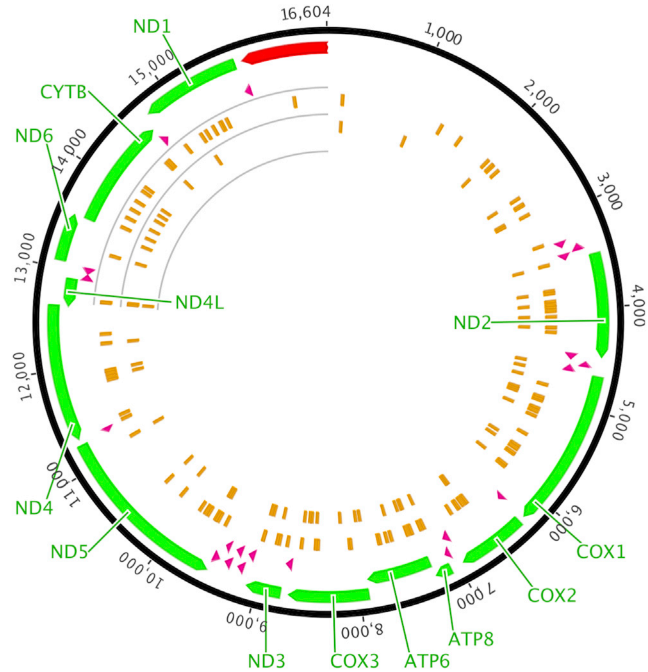


Figure 2. Mitochondrial Genome of *D. australis* and Variants Identified in the Historic Lord Howe Island Specimens

Variants that distinguish the Lord Howe Island specimens from the Ball's Pyramid mitochondrial genome sequence are shown by orange bars. The two specimens had 100 and 62 variants different from the 16,604 bp reference sequence from the Ball's Pyramid population, corresponding to 0.60% and 0.37% divergence, respectively. The two historic Lord Howe Island specimens differed by a comparable amount—92 variant sites (0.55% divergence). Major features are color coded as follows: green, genes; red, mitochondrial ribosomal RNA; and purple, tRNAs. The low mitochondrial divergence between Ball's Pyramid and museum samples collected on Lord Howe Island suggests that they are part of the same species, and, in fact, the two populations may not have diverged significantly.

also found that *D. australis* is most likely hexaploid, though this finding should be confirmed through future cytogenetic work (Figure 3). Polyploidy is not uncommon in stick insects, particularly in parthenogenetic lineages [12], and parthenogenesis may occur in *D. australis* as well [3]. Polyploidy does pose challenges for population genetic analysis, since many methods are specifically developed for diploid data. Furthermore, any investigations of nuclear genetic diversity should rely on methods capable of detecting allelic ratios, which will require high sequencing depths.

To mitigate difficulties posed by high ploidy and large nuclear genome size, we leveraged low-coverage resequencing of museum specimens to recover complete mitochondrial genomes in order to compare between- and within-island genetic diversity between Lord Howe Island and Ball's Pyramid. Mitochondrial genome coverage was high (at least 50×) for all specimens, allowing for accurate genotype calls and no missing data (Table 1). High coverage also eliminates errors introduced by DNA degradation [13], which can complicate population genetic analyses [14]. None of the four pinned zoo-bred specimens had detectable differences from the mitochondrial genome reference, which is consistent with their being descendants of the original female collected on Ball's Pyramid. By contrast, both museum specimens from Lord Howe Island had a number of genetic differences from the reference, though within typical range of variation expected within a species (Figure 2). These within- and between-island differences were of the same order, and less than 1% overall, suggesting that the two populations most likely diverged after the origin of this species and not long enough ago for speciation to have taken place.

Conservation Implications

Both Lord Howe Island and Ball's Pyramid were considerably larger during the Last Glacial Maximum, though apparently not in contact with each other [15], and could have harbored allopatric stick insect populations for an extended period of time. Estimating the actual age of separation would require more museum specimens from Lord Howe Island, and possibly more of those from Ball's Pyramid. In addition, accurate calibration of the mitochondrial clock would also be necessary. Current studies investigating phasmid phylogenies have confidence intervals at the tips that are far too wide to provide informative insights into population-level splits [16]. Alternatively, more sophisticated approaches using methods such as approximate Bayesian computation could be employed, but they would

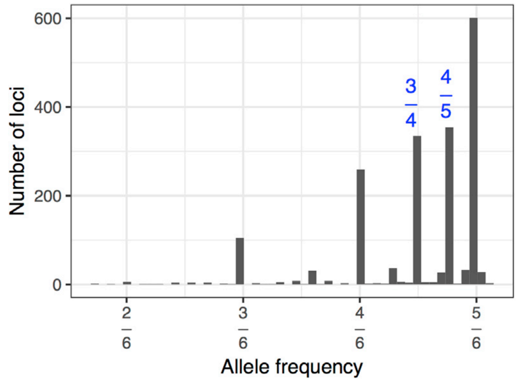


Figure 3. Allele Frequency Distribution in the *D. australis* Genome Suggests Hexaploidy

Most of the 1,900 loci analyzed were bi-allelic and had just two alleles, so only the frequency of one allele is shown. The ratios between alleles occur largely as fractions of six, consistent with there being this number gene copies present. The average coverage in these regions was low 6.5× ± 6.2 (SD), and peaks corresponding to fractions of four and five (highlighted with blue numbers) most likely correspond to lower-coverage regions where the total number of copies could not be correctly estimated. Indeed, the coverage at peaks 3/4 and 4/5 was significantly lower than that at peaks 4/6 and 5/6 (t test, n = 1,542, $p = 1.6 \times 10^{-6}$). The high copy number of most alleles can make population genomic analyses challenging, as it will require high read counts to estimate allelic frequencies with confidence.

Table 1. Summary Statistics from Sequencing and Re-mapping Reads to the Mitochondrial Reference Assembly

Sample ID	Type	Total Yield (Gb)	Read Length ± SD	Mapped Reads	Coverage ± SD
15_000002	museum	6.3	38 ± 8	6,842,938	412 ± 301
15_000003	museum	13	42 ± 5	2,801,149	167 ± 117
15_000004	zoo	12.5	44 ± 2	9,073,150	548 ± 421
15_000005	zoo	10.8	40 ± 7	3,216,964	194 ± 143
15_000006	zoo	1.5	40 ± 6	935,265	56 ± 28
15_000007	zoo	9.7	43 ± 5	3,711,997	223 ± 94

Museum samples were collected on Lord Howe Island, and zoo samples were related to the individual used to make the reference sequence, having descended from the pair captured for captive breeding [2]. Although the total yield was high, mapping rates to the genome reference were relatively low, resulting in low overall coverage. By contrast, mitochondrial coverage was high enough to overcome stochastic errors introduced by DNA degradation due to age [13] and to produce reliable genotypes.

require developing an array of markers for the study of museum specimens. The current reference genome and the captive-bred population can be used to select suitable loci and to test them for polymorphism. However, the observation that the difference between the two museum specimens is within the same range as the divergence between populations (Figure 1) suggests that Ball's Pyramid was most likely colonized relatively recently and that this population is suitable for reintroduction.

Potential Role of Polyploidy for *D. australis* Evolution and Conservation

Polyploidy is relatively rare in animals [17], though it has been detected in a variety of stick insects, particularly in parthenogenetic lineages (reviewed by Scali [18]). It difficult to state with any degree of certainty what role, if any, polyploidy may play in *D. australis*. However, extensive work on plants allows us to propose two effects: a resistance to inbreeding and an explanation for the large body size in this species.

Extensive research in plants has shown that polyploids are overrepresented among invasive species [19, 20], and polyploidy may help invasions by masking recessive mutations [21] or slowing progression toward full homozygosity in inbred populations [22]. Thus, polyploidy may be a reason for the successful survival of *D. australis* on the marginal habitat of Ball's Pyramid. It may also slow the onset of inbreeding depression in captive population, but not eliminate it. Consequently, we recommend that genetic diversity in these captive populations be assessed and actively maintained or enhanced through appropriate breeding strategies, including the collection and integration of new founders from Ball's Pyramid.

Another common outcome of polyploidy is an increase in body size [23]. Though the mechanism underlying this relationship is unclear and not always manifest, hexaploids are often the largest size class in an autopolyploid plant series (reviewed by [24]). Thus, polyploidy could be a mechanism underlying the repeated evolution of the giant "tree lobster" form in stick insects [9] and, even more generally, larger body sizes. If this is the case, we make the easily testable prediction that other tree lobster genera should have higher ploidy numbers.

Conclusions

Previous morphological assessment of the stick insects discovered on Ball's Pyramid suggested that they were a relic population of *D. australis* [2]. This current study provides genetic support for this conclusion and may help facilitate efforts toward the recovery of this species. Pending final approvals, a rodent eradication program is planned to commence on Lord Howe Island in mid-2018 [25]. Although ship rat eradication is challenging, and the resident human population on Lord Howe Island will make it even more so, historically >80% of tropical island eradication programs have succeeded [26, 27]. If successful, this will provide the first opportunity for the reintroduction of this species to its former range. However, had our findings demonstrated that the Ball's Pyramid population was a distinct species, the release of these insects on Lord Howe Island would be regarded as an introduction (the release of a species outside its indigenous range [28]). Introductions can be justified on the basis that a population or species faces a high risk of extinction in its current range or that it would most likely restore an important ecological function that has been lost by the extinction of another species. Nevertheless, introductions generally require a greater level of scrutiny and sometimes face legislative barriers because they can result in negative ecological and/or economic impacts that may often be difficult to foresee [28]. Consequently, the greater certainty provided by this study that the insects from Ball's Pyramid are indeed surviving representatives of *D. australis* can only strengthen the argument for reintroduction to Lord Howe Island should the opportunity arise.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Library preparation
  - Sequencing and reference nuclear and mitochondrial genome assemblies
  - Re-mapping museum specimens
  - Assessment of "index hopping"
  - Ploidy estimation
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

AUTHOR CONTRIBUTIONS

Conceptualization, A.S.M., D.Y., and A.Z.; Investigation, A.S.M., M.L.G., and L.Q.; Resources, M.J.L.M.; Data Curation, M.L.G.; Writing – Original Draft, A.S.M., M.J.L.M., and D.Y.; Visualization, Y.N.S. and A.S.M.; Funding Acquisition, A.S.M. and D.Y.

ACKNOWLEDGMENTS

We are grateful to the Okinawa Institute of Science and Technology DNA Sequencing Section (SQC) for help with sequencing. Funding for this project was provided by the Okinawa Institute of Science and Technology Graduate University to A.S.M. D.K.Y.'s research program is supported by the Schlinger Foundation Trust to the Australian National Insect Collection.

Received: July 30, 2017
Revised: August 22, 2017
Accepted: August 23, 2017
Published: October 5, 2017

REFERENCES

1. Paramonov, S.J. (1963). Lord Howe Island, a riddle of the Pacific, part III. Pac. Sci. *17*, 361–373.
2. Priddel, D., Carlile, N., Humphrey, M., Fellenberg, S., and Hiscox, D. (2003). Rediscovery of the "extinct" Lord Howe Island stick-insect (*Dryococelus australis* (Montrouzier)) (Phasmatodea) and recommendations for its conservation. Biodivers. Conserv. *12*, 1391–1403.
3. Honan, P. (2008). Notes on the biology, captive management and conservation status of the Lord Howe Island stick insect (*Dryococelus australis*) (Phasmatodea). J. Insect Conserv. *12*, 399–413.
4. IUCN (2017). The IUCN Red List of Threatened Species. http://www.iucnredlist.org.
5. Dirzo, R., Young, H.S., Galetti, M., Ceballos, G., Isaac, N.J.B., and Collen, B. (2014). Defaunation in the Anthropocene. Science *345*, 401–406.
6. Seddon, P.J., Griffiths, C.J., Soorae, P.S., and Armstrong, D.P. (2014). Reversing defaunation: restoring species in a changing world. Science *345*, 406–412.
7. Yeates, D.K., Zwick, A., and Mikheyev, A.S. (2016). Museums are biobanks: unlocking the genetic potential of the three billion specimens in the world's biological collections. Curr. Opin. Insect Sci. *18*, 83–88.
8. Silcocks, S., and Magrath, M.J.L. (2014). Morphological comparison of Lord Howe Island Stick Insects originating from the Ball's Pyramid population and the extinct Lord Howe Island population. https://www.aesconferences.com.au/wp-content/uploads/2012/10/Poster-program-with-numbers-2014-9-Sept.pdf.
9. Buckley, T.R., Attanayake, D., and Bradler, S. (2009). Extreme convergence in stick insect evolution: phylogenetic placement of the Lord Howe Island tree lobster. Proc. Biol. Sci. *276*, 1055–1062.
10. Nagl, W., and Schaffner, K.H. (1981). High 2C DNA content and endopolyploidy of ganglia in the leaf insect, Extatosoma tiaratum (Phasmida). Cell and Chromosome Newsletter *4*, 10–13. http://agris.fao.org/agris-search/search.do?recordID=US201301982132.
11. Tomita, S., Yukuhiro, K., and Kômoto, N. (2011). The mitochondrial genome of a stick insect *Extatosoma tiaratum* (Phasmatodea) and the phylogeny of polyneopteran insects. J. Insect Biotechnol. Sericology *80*, 79–88.
12. Morgan-Richards, M., Trewick, S.A., and Stringer, I.A.N. (2010). Geographic parthenogenesis and the common tea-tree stick insect of New Zealand. Mol. Ecol. *19*, 1227–1238.
13. Pääbo, S., Poinar, H., Serre, D., Jaenicke-Després, V., Hebler, J., Rohland, N., Kuch, M., Krause, J., Vigilant, L., and Hofreiter, M. (2004). Genetic analyses from ancient DNA. Annu. Rev. Genet. *38*, 645–679.
14. Morozova, I., Flegontov, P., Mikheyev, A.S., Bruskin, S., Asgharian, H., Ponomarenko, P., Klyuchnikov, V., ArunKumar, G., Prokhortchouk, E., Gankin, Y., et al. (2016). Toward high-resolution population genomics using archaeological samples. DNA Res. *23*, 295–310.
15. Papadopulos, A.S.T., Baker, W.J., Crayn, D., Butlin, R.K., Kynast, R.G., Hutton, I., and Savolainen, V. (2011). Speciation with gene flow on Lord Howe Island. Proc. Natl. Acad. Sci. USA *108*, 13188–13193.
16. Bradler, S., Cliquennois, N., and Buckley, T.R. (2015). Single origin of the Mascarene stick insects: ancient radiation on sunken islands? BMC Evol. Biol. *15*, 196.
17. White, M.J.D. (1977). Animal Cytology and Evolution (CUP Archive).
18. Scali, V. (2009). Stick insects: parthenogenesis, polyploidy and beyond. In Life and Time: The Evolution of Life and its History, V. Scali, P. Burighel, and A. Minelli, eds. (Cleup), pp. 171–192.
19. Pandit, M.K., Pocock, M.J.O., and Kunin, W.E. (2011). Ploidy influences rarity and invasiveness in plants. J. Ecol. *99*, 1108–1115.
20. Pandit, M.K., White, S.M., and Pocock, M.J.O. (2014). The contrasting effects of genome size, chromosome number and ploidy level on plant invasiveness: a global analysis. New Phytol. *203*, 697–703.
21. Rosche, C., Hensen, I., Mráz, P., Durka, W., Hartmann, M., and Lachmuth, S. (2017). Invasion success in polyploids: the role of inbreeding in the contrasting colonization abilities of diploid versus tetraploid populations of Centaurea stoebe sl. J. Ecol. *105*, 425–435.
22. te Beest, M., Le Roux, J.J., Richardson, D.M., Brysting, A.K., Suda, J., Kubesová, M., and Pysek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. Ann. Bot. (Lond.) *109*, 19–45.
23. Gregory, T.R., and Mable, B.K. (2005). Polyploidy in animals. In The Evolution of the Genome, T.R. Gregory, ed. (Elsevier Academic Press), pp. 427–517.
24. Tsukaya, H. (2013). Does ploidy level directly control cell size? Counterevidence from Arabidopsis genetics. PLoS ONE *8*, e83729.
25. Lord Howe Island Board (2017). Lord Howe Island rodent eradication project: NSW species impact statement. Report of the Lord Howe Island Board, February 15, 2007. http://www.lhib.nsw.gov.au/sites/lordhowe/files/public/images/documents/lhib/Environment/Rodent%20Eradication/2_LHIREP_NSW_SpeciesImpactStatement_15Feb17.pdf.
26. Keitt, B., Griffiths, R., Boudjelas, S., Broome, K., Cranwell, S., Millett, J., Pitt, W., and Samaniego-Herrera, A. (2015). Best practice guidelines for rat eradication on tropical islands. Biol. Conserv. *185*, 17–26.
27. Holmes, N.D., Griffiths, R., Pott, M., Alifano, A., Will, D., Wegmann, A.S., and Russell, J.C. (2015). Factors associated with rodent eradication failure. Biol. Conserv. *185*, 8–16.
28. IUCN/SSC (2013). Guidelines for reintroductions and other conservation translocations. Version 1.0. Report of the IUCN Species Survival Commission. https://portals.iucn.org/library/sites/library/files/documents/2013-009.pdf.

29. Mikheyev, A.S., Tin, M.M.Y., Arora, J., and Seeley, T.D. (2015). Museum samples reveal rapid evolution by wild honey bees exposed to a novel parasite. Nat. Commun. *6*, 7991.

30. Langmead, B. (2010). Aligning short sequencing reads with Bowtie. Curr. Protoc. Bioinformatics *Chapter 11*, 7.

31. Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv, arXiv:1207.3907, https://arxiv.org/abs/1207.3907.

32. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics *28*, 1647–1649.

33. Hahn, C., Bachmann, L., and Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads–a baiting and iterative mapping approach. Nucleic Acids Res. *41*, e129.

34. Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J., Chen, Z., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature *437*, 376–380.

35. Sedlazeck, F.J., Rescheneder, P., and von Haeseler, A. (2013). NextGenMap: fast and accurate read mapping in highly polymorphic genomes. Bioinformatics *29*, 2790–2791.

36. González-Domínguez, J., and Schmidt, B. (2016). ParDRe: faster parallel duplicated reads removal tool for sequencing studies. Bioinformatics *32*, 1562–1564.

37. Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics *30*, 614–620.

38. Dos Santos, R.A.C., Goldman, G.H., and Riaño-Pachón, D.M. (2017). ploidyNGS: visually exploring ploidy with next generation sequencing data. Bioinformatics *33*, 2575–2576.

39. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

40. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al.; 1000 Genomes Project Analysis Group (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

41. Tin, M.M.-Y., Economo, E.P., and Mikheyev, A.S. (2014). Sequencing degraded DNA from non-destructively sampled museum specimens for RAD-tagging and low-coverage shotgun phylogenetics. PLoS ONE *9*, e96793.

42. Tin, M.M.Y., Rheindt, F.E., Cros, E., and Mikheyev, A.S. (2015). Degenerate adaptor sequences for detecting PCR duplicates in reduced representation sequencing data improve genotype calling accuracy. Mol. Ecol. Resour. *15*, 329–336.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
D. australis specimens	This study	ANIC Database # 15-000002 to 15-000007
Chemicals, Peptides, and Recombinant Proteins		
SiMAG/FK-Silanol (100 mg/mL, Ø 1.0 μm)	Chemicell	Cat#1101
Dynabeads MyOneTM Carboxylic Acid	ThermoFisher Scientific	Cat#65012
Buffer PE	QIAGEN	Cat#19065
Buffer EB	QIAGEN	Cat#19086
UltraPure guanidine isothiocyanate	Invitrogen	Cat#15535016; CAS: 593-84-0
Trizma hydrochloride	Sigma-Aldrich	Cat#T5941; CAS: 1185-53-1
Ethylenediaminetetraacetic Acid	Nacalai Tesque	Cat#15105-35; CAS: 60-00-4
N-Lauroylsarcosine sodium salt	Sigma-Aldrich	Cat#L9150; CAS: 137-16-6
2-mercaptoethanol	Sigma-Aldrich	Cat#M6250; CAS: 60-24-2
Ethanol	Nacalai Tesque	Cat#14712-05; CAS: 64-17-5
Agarose -LE, Classic Type	Nacalai Tesque	Cat#01157-95; CAS: 9012-36-6
Polyethylene Glyco #6000	Nacalai Tesque	Cat#28254-85; CAS: 25322-68-3
5mol/l-Sodium Chloride Solution	Nacalai Tesque	Cat#31334-51; CAS: 7647-14-5
FastAP Alkaline phosphatase	ThermoFisher Scientific	Cat#EF0651
NEBuffer 4	New England Biolabs	Cat#B7004S
2.5mM $CoCl_2$	New England Biolabs	Cat#B0252S
Terminal Transferase	New England Biolabs	Cat#M0315L
100 mM GTP	Takara	Cat#4042
klenow fragment (3′→5′ exo-)	New England Biolabs	Cat#M0212M
10 mM dNTP mix	Promega	Cat#U151A
BSA	New England Biolabs	Cat#B9000S
T4 DNA polymerase	New England Biolabs	Cat#M0203S
Adenosine 5′-Triphosphate (ATP)	New England Biolabs	Cat# P0756S
Calf Intestinal Alkaline Phosphatase	ThermoFisher Scientific	Cat# 18009-019
2x Quick Ligase Reaction Buffer	New England Biolabs	Cat#B2200S
T4 DNA ligase	New England Biolabs	Cat#M0202L, M0202M
5x Phusion HF Buffer	ThermoFisher Scientific	Cat#F-518
Phusion DNA polymerase	ThermoFisher Scientific	Cat#F-530L
2-Log DNA Ladder	New England Biolabs	Cat#N3200L
EZ-Vision one	AMRESCO	Cat#N472-KIT
Critical Commercial Assays		
Truseq DNA LT sample Prep KIT	Illumina	Cat#FC-121-2002
MinElute Reaction Cleanup Kit	QIAGEN	Cat#28206
Quant-iT PicoGreen dsDNA Assay Kit	ThermoFisher Scientific	Cat#P7589
High Sensitivity DNA Kit	Agilent	Cat#5067-4626
KAPA SYBR FAST Universal qPCR kit	KAPA Biosystems	Cat#KK4601
Illumina DNA Standards and Primer Premix Kit	KAPA Biosystems	Cat#KK4808
Sequence assembly and raw data	DDBJ/ENA/GenBank	GenBank: PRJNA387351
Oligonucleotides		
CCCCC_TS: GTGACTGGAGTTCAGACGTGTGCTCTTCC GATCCCCC	This study	N/A
Truseq PCR1: AATGATACGGCGACCACCGAGATCTACA	This study	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
I7_D701 primer_long: 5′-CAAGCAGAAGACGGCATACGAGAT CGAGTAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGAT	This study	N/A
I7_D702 primer_long: 5′-CAAGCAGAAGACGGCATACGAGAT TCTCCGGAGTGACTGGAGTTCAGACGTGTGCTCTTCCGAT	This study	N/A
I7_D705 primer_long: 5′-CAAGCAGAAGACGGCATACGAGATT TCTGAATGTGACTGGAGTTCAGACGTGTGCTCTTCCGAT	This study	N/A
I7 truseq upper:5′-p-ATCGGAAGAGCACACGTCTGAACT CCAGT*/ddC/ (* = phosphorothioate linkage)	[29]	N/A
I7-D701 FL truseq lower:5′-CAAGCAGAAGACGGCATACGA GATCGAGTAATGTGACTGGAGTTCAGACGTGTGCTC TTCCGAT*C*C*CCC (* = phosphorothioate linkage)	[29]	N/A
I7-D703 FL truseq lower:5′-CAAGCAGAAGACGGCATACGA GATAATGAGCGGTGACTGGAGTTCAGACGTGTGCTCTTC CGAT*C*C*CCC (* = phosphorothioate linkage)	[29]	N/A
I7-D704 FL truseq lower:5′-CAAGCAGAAGACGGCATACGAGA TGGAATCTCGTGACTGGAGTTCAGACGTGTGCTCTTCCG AT*C*C*CCC (* = phosphorothioate linkage)	[29]	N/A
I5 truseq upper: CCCTACACGACGCTCTTCCGATCT/ddC/	[29]	N/A
I5-D501 FL truseq lower: 5′-p-GAGATCGGAAGAGCGTCGTGT AGGGAAAGAGTGTAGGCTATAGTGTAGATCTCGGTGGTC GCCGTATCATT	[29]	N/A
I5-D502 FL truseq lower: 5′-p-GAGATCGGAAGAGCGTCGTG TAGGGAAAGAGTGTGCCTCTATGTGTAGATCTCGGTGGTC GCCGTATCATT	[29]	N/A
I5-D503 FL truseq lower: 5′-p-GAGATCGGAAGAGCGTCGTG TAGGGAAAGAGTGTAGGATAGGGTGTAGATCTCGGTGGT CGCCGTATCATT	[29]	N/A
I5-D504 FL truseq lower: 5′-p-GAGATCGGAAGAGCGTCGTGT AGGGAAAGAGTGTTCAGAGCCGTGTAGATCTCGGTGGTC GCCGTATCATT	[29]	N/A
I5-D506 FL truseq lower: 5′-p-GAGATCGGAAGAGCGTCGTGT AGGGAAAGAGTGTTAAGATTAGTGTAGATCTCGGTGGTCG CCGTATCATT	[29]	N/A
Software and Algorithms		
Bowtie	[30]	http://bowtie-bio.sourceforge.net/index.shtml
FreeBayes	[31]	https://github.com/ekg/freebayes
Geneious (R8.1)	[32]	https://www.geneious.com/
MITOBim	[33]	https://github.com/chrishah/MITObim
Newbler	[34]	http://www.roche.com/
NextGenMap	[35]	http://cibiv.github.io/NextGenMap/
ParDRe	[36]	https://sourceforge.net/projects/pardre/
PEAR	[37]	https://sco.h-its.org/exelixis/web/software/pear/
ploidyNGS	[38]	https://github.com/diriano/ploidyNGS
Samtools	[39]	http://samtools.sourceforge.net/
VCFTools	[40]	http://vcftools.sourceforge.net/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests may be directed to and will be fulfilled by the Lead Contact, Alexander S. Mikheyev (sasha@homologo.us).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

A male (ANIC Database # 15-000002) and a female specimen (# 15-000003) of *D. australis* in the Australian National Insect Collection (ANIC) were sampled. These two specimens originated from Lord Howe Island proper and are part of the W. W. Froggatt collection; the collecting dates are unknown. In addition to historical museum specimens, we included four more recent control individuals from the Ball's Pyramid, which were captive-reared specimens from the Melbourne Zoo deposited in the ANIC (# 15-000004 to 15-000007).

METHOD DETAILS

Library preparation

A library from an ethanol-preserved zoo-bred specimen from Ball's Pyramid was constructed using a Truseq DNA LT Sample Prep Kit. The library was size-selected so that many of the forward and reverse reads would overlap to facilitate assembly.

Most libraries for the museum specimens were prepared using a PCR-free approach as described previously, while libraries for samples 15-000002 and 15-000004 were amplified with a limited number of PCR cycles (less than eight cycles) [29, 41]. Libraries were purified with 17% PEG-6000 [42], analyzed using a Bioanalyzer High Sensitivity Kit and then pooled together in equal amounts.

Sequencing and reference nuclear and mitochondrial genome assemblies

Reads for the genomic reference made from the ethanol-preserved Ball's Pyramid specimen were sequenced on an Illumina Hi-Seq2500 in PE250 mode, producing 97 million read pairs. PCR duplicates were removed using ParDRe [36], and were then stitched together using PEAR with parameters–min-overlap 10–memory 48G–threads 10 -n 200 -m 600 -p 0.0001 [37]. This resulted in 18.9 Gb of data with an average read length of 428 ± 51 (SD). The resulting super-reads were assembled using Newbler with parameters -large -m -cpu 10 -mi 95 -siom 450 -l 1000 -a 500 -urt -novs -a 1000 -minlen 45 -het [34]. The genome size estimate and was reported by Newbler. Mitochondrial DNA was assembled separately using MITObim [33] after aligning reads to the *Extatosoma tiaratum* mitochondrial genome (GenBank: NC_017748.1) [11] using NextGenMap [35]. MITObim had difficulties assembling the repetitive mitochondrial control region, resulting in a gap that was filled in manually by identifying super-reads that mapped to both sides of the gap. The *D. australis* mitochondrial assembly was annotated using the *E. tiaratum* sequence as a template in Geneious (r8.1), which was also used for interactive exploration and visualization of the data [32].

Re-mapping museum specimens

Libraries were sequenced using a HiSeq 2000 in SE50 mode. Reads were trimmed to remove 3′ adaptor sequences and mapped to the *D. australis* mitochondrial genome assembly using bowtie [30]. Variants were then called using freebayes in diploid mode [31]. Diploid mode was used because a small number of sites in the alignment had apparent mixtures of genotypes, possibly from nuclear copies or repetitive regions, making them difficult to resolve using short-read data from museum specimens. Therefore, 30 sites with heterozygous genotype calls were filtered from the variant call file, along with sites having quality scores less than 40. The final analysis retained 161 total variant sites, all of them single nucleotide polymorphisms. Because divergence from the reference sequence can cause poor mapping and variants to be missed, the mitochondrial read alignments were visually inspected for uniform coverage and other abnormalities.

Assessment of "index hopping"

Each sample was barcoded for sequencing using a unique combination Illumina adaptors, and sequenced in the same lanes. We quantified the extent of molecular recombination between the indexes by dividing the total number of invalid index combinations by the total number of valid index combinations (omitting indexes with any ambiguous calls). There was a low rate of index hopping (0.361%). It was a substantially smaller value than the average percentage of mapped reads assigned to mitochondrial DNA (2.1% ± 1.4% (SD)), and should not have influenced the result.

Ploidy estimation

Analysis of several candidate microsatellite loci developed from the reference genome suggested the presence of more than two alleles (L.Q. and A.S.M., unpublished data). However, since genetic variability in the captive population is limited, and microsatellites can give additional peaks as a result of PCR artifacts, these data were inconclusive, and we decided to further investigate ploidy of *D. australis* across multiple loci. We used ploidyNGS, a model-free approach for estimating ploidy [38], to calculate the frequencies of alleles in the ten longest scaffolds (2,617,109 bp) of the nuclear genome assembly, using super-reads remapped to the reference. This approach counts the relative frequencies of encountered alleles, which occur in predictable ratios that correspond to organismal ploidy. For example, diploid individuals will have biallelic sites predominantly at frequencies of 0, ½ and 1, whereas triploids will have 0, $^1/_3$, $^2/_3$ and 1, etc.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses were conducted in R (version 3.4) (https://www.R-project.org/), using built-in functions (*mean* and *sd)* to compute coverage means and standard deviations, respectively. Statistical differences in coverage were tested using the *t.test* function (Figure 3).

DATA AND SOFTWARE AVAILABILITY

The accession number for the sequences reported in this paper is GenBank: PRJNA387351.