Received: 29 May 2024 | Revised: 5 July 2024 | Accepted: 6 August 2024

DOI: 10.1111/1755-0998.14007

PERSPECTIVE

MOLECULAR ECOLOGY RESOURCES WILEY

Leveraging genomic load estimates to optimize captive breeding programmes

Evelyn L. Jensen[1] | Rachel Gray[1] | Joshua M. Miller[2]

[1]School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne, UK

[2]Department of Biological Sciences, MacEwan University, Edmonton, Alberta, Canada

Correspondence
Evelyn L. Jensen, School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne, UK.
Email: evelyn.jensen@newcastle.ac.uk

Handling Editor: Joanna Kelley

Rapid biodiversity loss threatens many species with extinction. Captive populations of species of conservation concern (such as those housed in zoos and dedicated breeding centres) act as an insurance should wild populations go extinct or need supplemental individuals to boost populations. Limited resources mean that captive populations are almost always small and started from few founding individuals. As a result, captive populations require careful management to minimize negative genetic impacts, with decisions about which individuals to breed together often guided by the principle of minimizing relatedness. Typically this strategy aims to retain 90% of genetic diversity over 200 years (Soulé et al., *Zoo Biology*, 1986, 5, 101), but it has a weakness in that it does not directly manage for genetic load. In this issue of Molecular Ecology Resources, Speak et al. (*Molecular Ecology Resources*, 2024, e13967) present a novel proof-of-concept study for taking this next step and incorporating estimates of individual genetic load into the planning of captive breeding, using an approach that is likely to be widely applicable to many captive populations.

KEYWORDS
CADD, captive breeding, conservation, genetic load, ultraconserved elements

The impact of genetic load on species survival and conservation is getting increased attention (van Oosterhout, 2020). As populations become smaller, inbreeding exposes the negative impacts of recessive deleterious mutations, thus converting what was 'masked load' into 'realized load' (Bertorelle et al., 2022). Over generations in captivity at small effective population sizes, offspring can end up with reduced fitness (i.e. inbreeding depression) as more and more deleterious mutations become homozygous through inbreeding and genetic drift, and thus unmasked. Understanding the masked load in the founders and managing breeding accordingly thus offers some critically needed hope that fitness could be maintained over the extended time horizons that the current biodiversity crisis necessitates. To put this into practice, Speak et al. (2024) illustrate how deleterious alleles that contribute to genetic load can be identified using genomic data that would be feasible to collect for wide range of endangered species.

Identifying the fitness impacts of particular alleles is a real challenge in non-model organisms. In many studies, researchers quantify load through summation of the number of derived alleles or non-synonymous variants under the assumption that derived alleles are more likely to be deleterious and contribute to load (reviewed in Bertorelle et al., 2022). For model species, other more sophisticated approaches are possible, such as combined annotation dependent depletion (CADD) scores that classify the deleteriousness of variants by combining and synthesizing annotations, functional information, and simulation data (Kircher et al., 2014). Rather than a binary classification of 'deleterious or not', CADD values are 'phred-like', being scaled based on the estimated values of deleteriousness, and can be derived for several types of variants (single nucleotide polymorphisms, indels, etc.) in both coding and noncoding regions of the genome. So far CADD scores have only been developed for humans (Rentzsch et al., 2018) and a handful of model species (e.g. mice, Groß et al., 2018, pigs, Groß, Derks, et al., 2020), but Speak et al. (2024) were able to transfer the scores from chickens (Groß, Bortoluzzi, et al., 2020) to the pink pigeon (Figure 1) despite a divergence of ~91 million years using a novel bioinformatic tool they present (LoadLift, https://github.com/saspeak/LoadLift).

The transfer of CADD scores over such a long divergence was facilitated by focusing on ultraconserved elements (UCEs). UCEs are defined as sequences that are shared among divergent animal genomes and maintained via purifying selection. Although the biological functions of these genomic regions are not fully clear (Harmston et al., 2013), mutations in UCEs are expected to be unconditionally deleterious. Thus, focusing on mutations in UCEs along with integration of CADD scores can give more robust estimates of both masked and realized load.

In addition to the ability to lift over CADD scores, UCEs are a pragmatic choice for genetic marker for investigating genetic load in conservation contexts as it should be possible to collect UCE data for almost any species. Speak et al. (2024) used the approach of collecting whole-genome resequencing data from pink pigeons and

FIGURE 1 A captive pink pigeon (*Nesoenas mayeri*), which narrowly escaped extinction in the 1990s thanks to a captive breeding and reintroduction programme. Image courtesy of Harriet Whitford.

subsampling down to UCE regions, but it is possible to sequence just UCE using commercially available bait sets developed for use across broad taxonomic groups (e.g. tetrapods, fish and invertebrates). There are also comprehensive bioinformatic tools (e.g. PHYLUCE; Faircloth, 2016) that facilitate the analysis of UCE data without the need for a reference genome, which is often a limiting factor for many non-model organisms.

The value of minimizing relatedness between breeding pairs in captivity has long been recognized to limit inbreeding and maintain diversity (Ballou & Lacy, 1995). However, despite decades of evidence that genetic markers can improve breeding decisions, many breeding programmes are still managed using pedigree analysis of studbook data. Termed the 'conservation genomics gap' (Shafer et al., 2015), the lack of uptake and application of even well-established methods by conservation practitioners has frustrated the academic conservation genomics community. Speak et al. (2024) show that directly accounting for genetic load provides an additional and substantial benefit over considering relatedness alone in their 50-generation long captive breeding simulations. In this way, the authors provide additional evidence that genetically informed captive breeding can maximize the chance of long-term population viability and provide a novel, potentially broadly applicable approach. The increased ability to assess genetic load also holds promise for genetic rescue and translocations and highlights the possibilities of managing not just for neutral genetic variation, but addressing the root causes of inbreeding depression both in captivity and the wild.

AUTHOR CONTRIBUTIONS

ELJ, RG and JMM conceived and wrote the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors affirm that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Evelyn L. Jensen https://orcid.org/0000-0002-1819-3117
Rachel Gray https://orcid.org/0000-0002-1192-1505
Joshua M. Miller https://orcid.org/0000-0002-4019-7675

REFERENCES

Ballou, J. D., & Lacy, R. C. (1995). Identifying genetically important individuals for management of genetic diversity in pedigreed populations. In J. D. Ballou, T. J. Foose, & M. Gilpin (Eds.), *Population management for survival and recovery* (pp. 76–111). Columbia Press.

Bertorelle, G., Raffini, F., Bosse, M., Bortoluzzi, C., Iannucci, A., Trucchi, E., Morales, H. E., & van Oosterhout, C. (2022). Genetic load: Genomic estimates and applications in non-model animals. *Nature Reviews Genetics*, *23*, 492–503.

Faircloth, B. C. (2016). PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, *32*, 786–788.

Groß, C., de Ridder, D., & Reinders, M. (2018). Predicting variant deleteriousness in non-human species: Applying the CADD approach in mouse. *BMC Bioinformatics*, *19*, 373.
Groß, C., Bortoluzzi, C., de Ridder, D., Megens, H. J., Groenen, M. A. M., Reinders, M., & Bosse, M. (2020). Prioritizing sequence variants in conserved non-coding elements in the chicken genome using chCADD. *PLoS Genetics*, *16*, e1009027.
Groß, C., Derks, M., Megens, H.-J., Bosse, M., Groenen, M. A. M., Reinders, M., & de Ridder, D. (2020). pCADD: SNV prioritisation in sus scrofa. *Genetics Selection Evolution*, *52*, 4.
Harmston, N., Baresic, A., & Lenhard, B. (2013). The mystery of extreme non-coding conservation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*, 20130021.
Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, *46*, 310–315.
Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., & Kircher, M. (2018). CADD: Predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Research*, *47*, D886–D894.
Shafer, A. B., Wolf, J. B., Alves, P. C., Bergstrom, L., Bruford, M. W., Brannstrom, I., Colling, G., Dalen, L., De Meester, L., Ekblom, R., Fawcett, K. D., Fior, S., Hajibabaei, M., Hill, J. A., Hoezel, A. R., Hoglund, J., Jensen, E. L., Krause, J., Kristensen, T. N., … Zielinski, P. (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology & Evolution*, *30*, 78–87.
Speak, S. A., Birley, T., Bortoluzzi, C., Clark, M. D., Percival-Alwyn, L., Morales, H. E., & van Oosterhout, C. (2024). Genomics-informed captive breeding can reduce inbreeding depression and the genetic load in zoo populations. *Molecular Ecology Resources*, *24*, e13967.
van Oosterhout, C. (2020). Mutation load is the spectre of species conservation. *Nature Ecology & Evolution*, *4*, 1004–1006.

How to cite this article: Jensen, E. L., Gray, R., & Miller, J. M. (2024). Leveraging genomic load estimates to optimize captive breeding programmes. *Molecular Ecology Resources*, *24*, e14007. https://doi.org/10.1111/1755-0998.14007