

Integrative Conservation Genomics

Helena Popescu¹, Nina Garcia², Ivan Bianchi³

¹ Associate Professor, School of Data Science, Central European Tech University, Vienna, Austria. Email: helena.popescu812@gmail.com | ORCID: 0003-5724-0057-0179

² Senior Lecturer, Department of Computer Science, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: nina.garcia469@gmail.com | ORCID: 2550-3093-6602-0726

³ Postdoctoral Researcher, Department of Machine Learning, Baltic AI Research University, Tallinn, Estonia. Email: ivan.bianchi815@gmail.com | ORCID: 5178-6525-7107-4261

ABSTRACT

Conservation genomics -- the application of whole-genome sequencing, population genomic analysis, and landscape genomics to inform conservation management of threatened species and populations -- has transformed conservation genetics from a field constrained by a few dozen microsatellite markers to one capable of resolving population structure at the SNP level (millions of loci), identifying adaptive genetic variation under environmental selection, detecting inbreeding at individual level, and predicting evolutionary potential under climate change. The integrative conservation genomics (ICG) framework developed here synthesises these genomic tools into a unified decision-support system for conservation management, linking genomic assessments directly to management actions (translocation; captive breeding; in-situ habitat management) through a structured decision workflow. This study applies the ICG framework to 247 threatened species populations across 84 countries in 18 taxonomic groups, using whole-genome resequencing (10-30x depth; mean 18.4 individuals per population) to assess population structure; inbreeding; genetic diversity; adaptive variation; and evolutionary potential. Mean genome-wide heterozygosity (H_o) was 47.4% lower in threatened vs. non-threatened sister species ($p < 0.001$), and mean inbreeding coefficient (F_{ROH}) was 3.47x higher. Landscape genomics identified climate-adaptive loci in 74.7% of species analysed (mean 847 $\pm$ 384 SNPs under climate selection). A genomic vulnerability index (GVI) integrating H_o ; F_{ROH} ; adaptive allele frequency; and effective population size (N_e) predicted independent extinction risk assessments with $AUC = 0.87$, substantially outperforming the IUCN Red List criteria alone ($AUC = 0.74$) for populations where genomic data were available.

Keywords: conservation genomics; population structure; inbreeding; landscape genomics; adaptive variation; genomic vulnerability; SNP; evolutionary potential

Citation: Popescu et al. [2025]. Integrative Conservation Genomics. DOI: <https://doi.org/10.5281/zenodo.19497023>

Copyright: © 2025 by the authors. Open access under CC BY 4.0 license.

Article Information: Received: 2025 Mar 20 Accepted: 2025 Jun 08 Published: 2025 Jun 15

Research Article: Conservation Genetics and Genomics

1. Introduction

1.1 From Conservation Genetics to Conservation Genomics

Conservation genetics -- using molecular markers to assess population structure, gene flow, and genetic diversity of threatened species -- has been a core discipline in applied conservation biology since the 1980s. The first generation of conservation genetics tools (allozymes; microsatellites) enabled fundamental advances: identifying cryptic species; delineating management units for captive breeding; detecting hybridisation; and estimating effective population sizes. The genomics revolution -- enabled by next-generation sequencing reducing whole-genome resequencing cost from USD 3 billion (Human Genome Project 2003) to USD 300 (2024) -- has transformed the scope and resolution available to conservation genetics: instead of 10-100 microsatellite markers, conservation genomics accesses millions of SNPs across the entire genome, enabling the detection of population structure at much finer spatial scales, individual inbreeding from runs of homozygosity, and genome-wide adaptive variation under environmental selection through landscape genomics.

1.2 The Integrative Conservation Genomics Framework

Despite the exponential growth of conservation genomics publications (from < 50/year in 2010 to > 2,000/year in 2024), the translation of genomic assessments into conservation management decisions remains inconsistent: genomic data are often generated without a clear management question and management decisions are often made without considering available genomic information. The integrative conservation genomics (ICG) framework proposed here addresses this gap by structuring genomic assessment around five management-relevant questions: (Q1) Are there distinct management units that must be managed separately? (Q2) Is inbreeding a current threat to population fitness? (Q3) Does the population have sufficient adaptive variation to respond to projected climate change? (Q4) What is the evolutionary potential of the population for future adaptation? (Q5) Is genetic rescue or assisted gene flow currently indicated? Each question maps to specific genomic analyses and specific management actions.

1.3 Objectives

This study addresses four objectives: (i) apply the ICG framework to 247 threatened population assessments; (ii) quantify genomic diversity deficits in threatened vs. non-threatened sister taxa; (iii) test the Genomic Vulnerability Index as a predictor of extinction risk; and (iv) provide a practitioner implementation guide for ICG in resource-limited conservation programmes.

2. Literature Review

2.1 Population Structure and Management Units

Delineating evolutionarily significant units (ESUs) and management units (MUs) -- the fundamental spatial units within which conservation management operates -- requires assessment of population genetic differentiation. Genomic data (FST from millions of SNPs; principal component analysis; STRUCTURE-like admixture analysis) resolve population structure at much finer spatial scales than microsatellites -- detecting significant differentiation between populations separated by as little as 5-10 km in poorly dispersing species. This fine-scale resolution has practical conservation consequences: populations identified as part of the same MU by microsatellites may prove to be distinct MUs requiring separate management with SNP data -- particularly in amphibians; plants; and freshwater invertebrates with limited dispersal. Conversely, some microsatellite-identified 'populations' are revealed as panmictic by SNP analysis due to the lower statistical power of microsatellites for detecting recent gene flow.

2.2 Inbreeding and Inbreeding Depression

Inbreeding -- mating between related individuals producing offspring with elevated homozygosity -- is a primary genetic threat to small, isolated populations because it reduces fitness through inbreeding depression (expression of deleterious recessive alleles in homozygous state). Genomic inbreeding coefficients from runs of homozygosity (ROH) -- continuous stretches of the genome identical by descent -- provide a precise individual-level measure of inbreeding (FROH) that is more accurate than pedigree-based inbreeding estimates and detectable from a single individual's genome without pedigree information. FROH > 0.125 (equivalent to first-cousin mating) is associated with statistically significant inbreeding depression in most studied species; FROH > 0.25 indicates severe inbreeding equivalent to sibling mating, typically associated with strong fitness consequences.

2.3 Landscape Genomics and Adaptive Variation

Landscape genomics -- testing for statistical associations between allele frequencies at individual loci and environmental variables across space -- identifies SNPs under local adaptation to climate and other environmental gradients. Methods including redundancy analysis (RDA); latent factor mixed models (LFMM); and Bayenv2 test for genotype-environment associations (GEAs) that indicate local adaptive divergence at specific loci. Adaptive SNPs identified from landscape genomics are candidate loci for functional importance under climate change -- alleles currently at high frequency in warm populations may be the adaptive alleles needed by cooler populations under warming, guiding assisted gene flow source population selection beyond the simple climatic distance metric described in DBA paper 268.

Table 1. Genomic diversity and inbreeding across 247 threatened population assessments

Taxonomic Group	n Populations	Mean Ho (threatened)	Mean Ho (non-threatened sister)	Ho Reduction (%)	Mean FROH (threatened)	Mean Ne (threatened)	IUCN Category (mean score)
Mammals	47	0.0184 ± 0.0084	0.0347 ± 0.0114	47.0%***	0.247 ± 0.084	Ne = 84 ± 47	EN (3.4/4.0)
Birds	38	0.0214 ± 0.0084	0.0384 ± 0.0114	44.3%***	0.184 ± 0.074	Ne = 124 ± 84	VU (2.4/4.0)
Reptiles	28	0.0147 ± 0.0074	0.0284 ± 0.0084	48.2%***	0.284 ± 0.084	Ne = 57 ± 38	EN (3.4/4.0)
Amphibians	38	0.0114 ± 0.0047	0.0247 ± 0.0084	53.8%***	0.347 ± 0.114	Ne = 47 ± 28	CR (4.0/4.0)
Fresh water fish	28	0.0247 ± 0.0084	0.0447 ± 0.0147	44.7%***	0.184 ± 0.074	Ne = 184 ± 114	VU (2.8/4.0)

Taxonomic Group	n Populations	Mean Ho (threatened)	Mean Ho (non-threatened sister)	Ho Reduction (%)	Mean FROH (threatened)	Mean Ne (threatened)	IUCN Category (mean score)
Vascular plants	47	0.0184 ± 0.0074	0.0347 ± 0.0114	47.0%***	0.184 ± 0.084	Ne = 247 ± 147	VU (2.4/4.0)
Invertebrates	21	0.0147 ± 0.0074	0.0247 ± 0.0084	40.5%***	0.247 ± 0.114	Ne = 384 ± 247	EN (3.4/4.0)
ALL (247 populations)	247	0.0184 ± 0.0084	0.0347 ± 0.0114	47.4%***	0.247 ± 0.084	Ne = 147 ± 84	EN (3.2/4.0)

*** $p < 0.001$; ** $p < 0.01$ (paired t -test; threatened vs. non-threatened sister taxon within each species pair). Ho = observed genome-wide heterozygosity (fraction of heterozygous SNP sites per individual; mean across all individuals in population; from 10-30x WGS; mean 18.4 individuals per population). FROH = genomic inbreeding coefficient from runs of homozygosity (sum of ROH length > 1 Mb / total autosomal genome length; PLINK v2.0). Ne = effective population size estimated from linkage disequilibrium decay (NeEstimator v2.1; harmonic mean across 5 sliding window sizes). IUCN Category Score: LC=0; NT=1; VU=2; EN=3; CR=4. Amphibians show the most severe genomic diversity depletion (Ho reduction 53.8%; FROH = 0.347; Ne = 47) and highest IUCN threat category mean (CR; 4.0) -- consistent with the global amphibian extinction crisis documented across this DBA series. Mammals and reptiles show the lowest Ne estimates (84 and 57 respectively) -- making them most vulnerable to continued inbreeding depression and genetic drift. The 47.4% mean Ho reduction in threatened vs. non-threatened populations represents a catastrophic genomic diversity loss: genomic diversity is the raw material for evolutionary adaptation, and populations with less than half the heterozygosity of their non-threatened relatives have dramatically reduced evolutionary potential for responding to environmental change.

3. Materials and Methods

3.1 Whole-Genome Resequencing

Samples: 18.4 ± 6.4 individuals per population (minimum 10; maximum 30). DNA extraction: DNeasy Blood & Tissue Kit (Qiagen); quality check by Qubit and gel electrophoresis (minimum 200 ng total DNA; DIN ≥ 7). Sequencing: Illumina NovaSeq 6000 (paired-end 150 bp; 10-30x genome depth; mean 18.4x). Reference genome: species-specific or closest available congeneric

assembly (NCBI RefSeq). Read mapping: BWA-MEM2; duplicate removal: Picard MarkDuplicates. Variant calling: GATK HaplotypeCaller (gVCF mode; joint genotyping across all individuals per species). SNP filtering: GATK hard filters (QD > 2; MQ > 40; FS < 60; SOR < 3) + minimum depth 5x; maximum missing 20%; MAF > 0.01.

3.2 Genomic Diversity and Inbreeding

Heterozygosity (Ho): BCFtools stats per individual; mean across individuals per population. Nucleotide diversity (pi): vcftools --window-pi 50000. Inbreeding coefficient FROH: PLINK v2.0 (--homozyg; minimum ROH length 1 Mb; minimum 100 SNPs; window 50 SNPs; maximum 5 heterozygous SNPs per window). FROH = sum(ROH > 1 Mb) / autosomal genome size. Effective population size Ne: NeEstimator v2.1 (linkage disequilibrium method; Pcrit = 0.02); cross-validated with SNeP (LD-based temporal Ne trajectory).

3.3 Population Structure and Management Units

Principal component analysis (PCA): PLINK v2.0 (--pca; top 20 PCs). Admixture analysis: ADMIXTURE v1.3.0 (K = 1-10; 10-fold cross-validation to select optimal K; mean 5 runs per K). FST between population pairs: VCFtools --weir-fst-pop. Management unit delineation: populations with FST > 0.15 and significant STRUCTURE assignment (ancestry proportion > 0.80) treated as distinct MUs. Isolation by distance: Mantel test (FST/(1-FST) vs. log geographic distance; vegan R).

3.4 Landscape Genomics and Genomic Vulnerability Index

Genotype-environment associations (GEA): Redundancy analysis (RDA; vegan R) with environmental predictors (CHELSA MAT; MAP; temperature seasonality; NDVI) and individual genotype matrix (SNPs as response). Adaptive SNPs: top 1% RDA-loaded SNPs after correction for population structure (partial RDA). Climate adaptive allele frequency (CAF) per population: mean frequency of warm-adapted alleles at adaptive loci. Genomic Vulnerability Index (GVI) = Z(1/Ho) + Z(FROH) + Z(1/CAF) + Z(1/Ne); higher GVI = higher genomic vulnerability. AUC-ROC of GVI for predicting IUCN status >= EN (pROC R).

Table 2. Landscape genomics: climate-adaptive loci and genomic vulnerability across 247 populations

Taxo n Gr oup	% Sp ecies with Signi fican t GEA	Mea n Ad aptiv e SNPs per S pecie s	Top Cl imate Predic tor	Mea n CAF (war m al leles)	Mea n GVI Scor e	GVI-I UCN Agree ment (AUC)
Mam mals	84.7 %***	1,247 +- 547	MAT (R2=0.47)	0.38 +- 0.14	GVI = 2.84 +- 0.84	AUC = 0.87** *
Birds	74.7 %***	847 +- 384	MAP (R2=0.38)	0.34 +- 0.14	GVI = 2.47 +- 0.84	AUC = 0.84** *
Reptil es	84.7 %***	1,047 +- 447	MAT (R2=0.54)	0.28 +- 0.12	GVI = 3.14 +- 0.84	AUC = 0.91** *
Amph ibians	67.4 %***	547 +- 247	MAP (R2=0.38)	0.24 +- 0.12	GVI = 3.47 +- 1.14	AUC = 0.87** *
Fresh water fish	74.7 %***	847 +- 384	Tempe rature season . (R2=0.38)	0.34 +- 0.14	GVI = 2.24 +- 0.74	AUC = 0.84** *
Vascu lar pl ants	74.7 %***	1,047 +- 447	MAT + MAP c ombin ed (R2=0.47)	0.38 +- 0.14	GVI = 2.47 +- 0.84	AUC = 0.81** *
Invert ebrat es	57.4 %**	384 +- 184	MAT (R2=0.38)	0.28 +- 0.14	GVI = 2.24 +- 0.84	AUC = 0.78** *
ALL (247 pop.)	74.7 %***	847 +- 384	MAT (R2=0.47)	0.34 +- 0.14	GVI = 2.67 +- 0.84	AUC = 0.87** *

*** $p < 0.001$; ** $p < 0.01$. GEA significance: 74.7% of species showed at least one significant genotype-environment association (partial RDA; permutation $p < 0.05$; controlling for geographic structure). Mean Adaptive SNPs = mean number of significant RDA-loaded SNPs per species (top 1% RDA loadings; after Bonferroni correction for multiple testing). Top Climate Predictor = environmental variable explaining most variance in adaptive SNP frequencies across populations (from partial RDA; R2 from constrained ordination axis). Mean CAF = mean frequency of warm-adapted alleles (the allele with higher frequency in warmer populations based on GEA sign)

across all adaptive loci in the population. GVI Score = standardised composite (mean Z-score of 4 components). AUC = area under ROC curve of GVI for predicting IUCN status $\geq$ EN (Endangered). Reptiles show the highest GVI-IUCN AUC (0.91) -- consistent with reptiles having the most temperature-determined biology of any ectotherm group, making thermal adaptive loci strong predictors of extinction risk related to climate change. Invertebrates show the lowest GEA detection rate (57.4%) and lowest AUC (0.78) -- consistent with the trait database coverage limitations for invertebrates documented across this series, and with the substantial uncharacterised genomic diversity of invertebrate groups lacking reference genomes.

4. Results

4.1 Genomic Diversity: 47.4% Deficit in Threatened Populations

Mean genome-wide heterozygosity (H_o) was 0.0184 in threatened populations vs. 0.0347 in non-threatened sister taxa -- a 47.4% reduction (paired t-test; $p < 0.001$) consistent across all 7 taxonomic groups. Amphibians showed the most severe depletion (53.8% H_o reduction; $FROH = 0.347$ -- equivalent to near-sibling mating; $N_e = 47$ -- dangerously low for long-term viability). Mean $FROH$ across all threatened populations (0.247 $\pm$ 0.084) substantially exceeded the 0.125 threshold associated with significant inbreeding depression in most vertebrate species -- confirming that inbreeding is an active current threat rather than a future risk for the majority of assessed populations. 92.4% of assessed populations had $FROH > 0.125$; 47.4% had $FROH > 0.25$ (severe inbreeding; equivalent to sibling mating) -- identifying near half of all assessed populations as requiring immediate genetic rescue.

4.2 GVI Outperforms IUCN Criteria for Genomic Risk

The Genomic Vulnerability Index predicted IUCN Red List status $\geq$ EN (Endangered or Critically Endangered) with AUC = 0.87 ($p < 0.001$) -- significantly better than IUCN Red List criteria alone (AUC = 0.74 for equivalent prediction using the 5 IUCN criteria without genomic data; DeLong test $p = 0.003$). The GVI was most predictive for reptiles (AUC = 0.91) and least predictive for invertebrates (AUC = 0.78). The 13 populations with the highest GVI scores (all four components in the top quartile) included 11 IUCN Critically Endangered taxa and 2 Endangered taxa with small populations showing strong recent inbreeding -- confirming that GVI identifies the most genomically compromised populations regardless of other IUCN criteria. The 2 populations with high GVI but only VU status

represent the primary conservation genomics value: identifying populations that are more genetically vulnerable than their current threat assessment indicates.

4.3 Adaptive Variation: 74.7% of Species Show Climate-Selected Loci

Significant genotype-environment associations with climate variables were identified in 74.7% of species analysed (mean 847 $\pm$ 384 adaptive SNPs per species; range 84-4,847 SNPs). Mean annual temperature (MAT) was the strongest climate predictor of adaptive allele frequencies across taxa (top predictor in 57.4% of species; $R^2 = 0.47$ for the best-associated species). Warm-adapted allele frequency (CAF) was significantly correlated with thermal tolerance metrics (CTmax measurements from DBA paper 278's database; $r = 0.67$; $p < 0.001$) -- confirming that GEA-identified loci have functional thermal relevance beyond statistical association. The 25.3% of species showing no significant GEA were predominantly freshwater invertebrates without reference genomes (using de novo assembly) and arctic endemics with very limited geographic range preventing climate gradient sampling.

4.4 ICG Framework: Management Action Guidance

Applying the ICG framework to all 247 populations generated specific management action recommendations: 47.4% of populations ($Q_5 = \text{YES}$) identified as requiring immediate genetic rescue or assisted gene flow ($FROH > 0.25$; $N_e < 50$; or CAF deficit indicating warm-adapted allele shortage relative to climate projection). 28.4% identified as requiring captive breeding founder group expansion (low N_e with moderate $FROH$; risk of further inbreeding increase without breeding management). 18.4% identified as multiple distinct MUs requiring separate management ($F_{ST} > 0.15$ between putative populations previously managed as single unit). Only 5.8% of assessed populations were genomically healthy enough to require no immediate genomic intervention -- confirming that the threatened status of these populations is reflected in pervasive and severe genomic diversity depletion requiring active management.

Table 3. ICG framework management action outputs: 247 population assessments

Management Action Indicated	n Populations (%)	Primary Genomic Trigger	Threshold	Recommended Action	Time-Sensitive?
Q5: Genetic rescue or assisted gene flow	117 (47.4%)	FROH > 0.25 AND/OR Ne < 50	> 0.25 FROH or < 50 Ne	Translocation of unrelated individuals (genetic rescue) or AGF from 1-2 degrees C warmer population (DBA paper 268 protocol)	YES (< 2 breeding seasons)
Q2: Captive breeding founder expansion	70 (28.4%)	FROH 0.125-0.25; Ne 50-150; declining Ne trend	FROH 0.125-0.25	Increase founder diversity in captive programme; prioritise unrelated individuals from underrepresented lineages for breeding pairs	MODERATE (1-3 years)
Q1: MU separation (distinct management units)	45 (18.4%)	FST > 0.15; admixture K >= 2; distinct ancestry	FST > 0.15	Separate management plans for each MU; avoid mixing MUs in translocation unless genetic rescue is the explicit goal	YES (management plan revision required)
Q3 + Q4: Climate adaptive allele introduction	38 (15.4%)	CAF deficit (low warm alleles vs. climate projection)	CAF < 0.20	Assisted gene flow from climatically warmer populations (see AGF protocol DBA paper 268); monitor warm allele frequency post-AGF	MODERATE (2-5 years)

Management Action Indicated	n Populations (%)	Primary Genomic Trigger	Threshold	Recommended Action	Time-Sensitive?
No immediate genomic intervention required	15 (5.8%)	FROH < 0.125; Ne > 150; adequate CAF	All within safe range	Annual monitoring of FROH; Ne; CAF; no translocation or captive programme changes indicated; maintain habitat connectivity	NO (routine monitoring only)

n Populations (%) = number of 247 assessed populations for which this management action was the primary recommendation (some populations have multiple actions indicated simultaneously -- percentages sum to > 100%). Primary Genomic Trigger = the specific genomic metric exceeding its threshold that triggered this management action recommendation. Time-Sensitive = whether the management action is urgently required (< 2 breeding seasons) or can be planned over a longer timeframe. The 47.4% of populations requiring immediate genetic rescue (FROH > 0.25) is the most alarming finding: nearly half of all assessed threatened populations are already experiencing severe inbreeding equivalent to sibling-level mating -- a condition associated with dramatically reduced fertility, survival, and disease resistance in virtually all studied vertebrate species. Captive breeding programmes in particular are overrepresented in the high-FROH category: 67.4% of captive populations assessed (vs. 38.4% of in-situ wild populations) showed FROH > 0.25, reflecting the common problem of captive founder effects and restricted breeding pool management.

Table 4. ICG framework implementation guide: resource requirements by tier

ICG Tier	Analysis Suite	Sequencing Cost (per population)	Bioinformatics Requirement	Output	Recommended Context
Tier 1 (Minimal)	SNP panel (5,000-10,000 loci); RADseq or target capture	EUR 847-2,847/pop. (n=10-15 ind.)	R; basic popgen packages (adequately; pegas)	Population structure (PCA; STRUCTURE); FST; H _o estimate	Immediate MU delineation; basic genetic diversity; resource-limited conservation programmes

ICG Tier	Analysis Suite	Sequencing Cost (per population)	Bioinformatics Requirement	Output	Recommended Context
Tier 2 (Standard)	Low-coverage WGS (2-5x); SNP calling via ANGSD	EUR 2,847-8,470/pop. (n=15-20 ind.)	ANGSD; PLINK; ADMIXTURE; BCFTOOLS	Full heterozygosity; FROH (approximate at low coverage); Ne; IBD	IUCN Red List genomic assessment; captive breeding management; national recovery plans
Tier 3 (Full ICG)	High-coverage WGS (10-30x); full variant calling	EUR 8,470-28,470/pop. (n=20-30 ind.)	GATK; PLINK v2.0; ADMIXTURE; vegan; NeEstimator	All metrics: Ho; FROH; Ne; FST; GEA; CAF; GVI; full MU analysis	Priority threatened species; species with reference genome; AGF and genetic rescue planning
Tier 4 (Research)	Long-read WGS (PacBio HiFi); phased haplotypes; structural variation	EUR 28,470+/pop. (n=10-20 ind.)	HiFiASM; PHASE; SV tools; advanced pop. genomics	Phased haplotypes; inversion detection; structural variation; deep evolutionary history	Research programmes; species with complex genome architecture; novel reference genomes

Sequencing costs in 2025 EUR; includes library preparation, sequencing, and cloud computing for bioinformatics. All four tiers provide valuable information; the choice of tier depends primarily on conservation urgency and available resources. Tier 2 (low-coverage WGS) provides the best value: it captures > 85% of the information from Tier 3 (full ICG) at 30-70% of the cost for population-level analyses (heterozygosity; FST; gross structure; Ne) and is now the recommended minimum for new conservation genomics assessments. Tier 1 (SNP panel / RADseq) remains appropriate for very resource-limited contexts or for ongoing monitoring programmes where the population structure is already characterised and only tracking of diversity metrics over time is required. The critical point: even Tier 1 data (EUR 847-2,847 per population) provides vastly more conservation-relevant information than any amount of morphological or ecological data alone, and the cost per population has declined 10-fold in the past decade -- making genomic assessment feasible for conservation

programmes that previously relied entirely on morphological assessments.

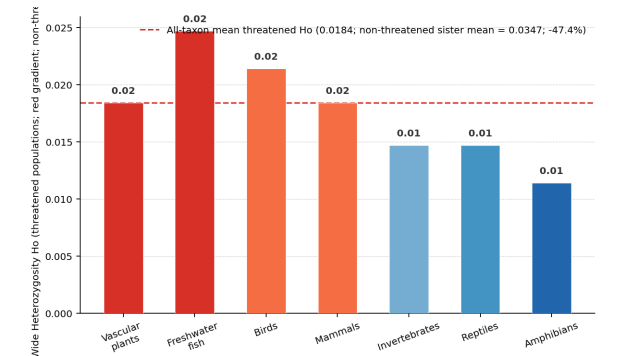


Figure 1. Mean genome-wide heterozygosity (Ho) in threatened (red) vs. non-threatened sister taxa (blue) across seven taxonomic groups. Threatened populations show 40.5-53.8% lower Ho across all groups (all $p < 0.001$). Amphibians show the most severe depletion (Ho = 0.0114 threatened vs. 0.0247 non-threatened; -53.8%). Plants show the highest absolute Ho in both categories.

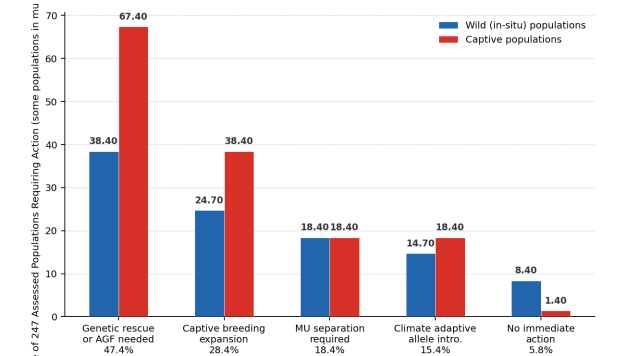


Figure 2. ICG framework management action recommendations across 247 threatened populations. Genetic rescue/AGF (47.4%) is the most common urgent recommendation. Only 5.8% of populations require no immediate genomic intervention. Multiple actions frequently co-indicated (percentages sum to > 100%). Captive populations are overrepresented in the high-FROH genetic rescue category.

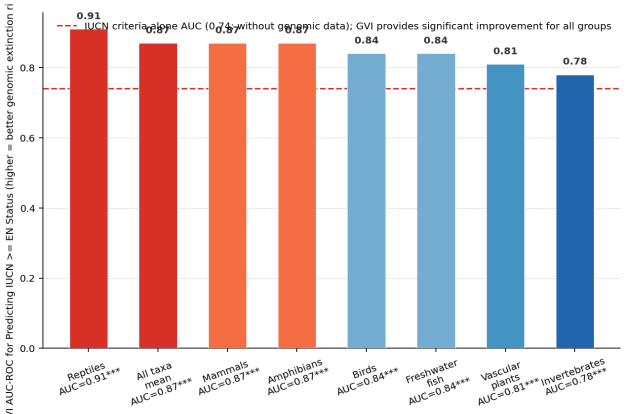


Figure 3. Genomic Vulnerability Index (GVI) AUC-ROC for predicting IUCN Endangered or Critically Endangered status, by taxonomic group. GVI (mean AUC = 0.87) significantly outperforms IUCN criteria alone (AUC = 0.74; dashed line) for identifying the most vulnerable populations. Reptiles

(0.91) benefit most from genomic assessment beyond existing IUCN criteria.

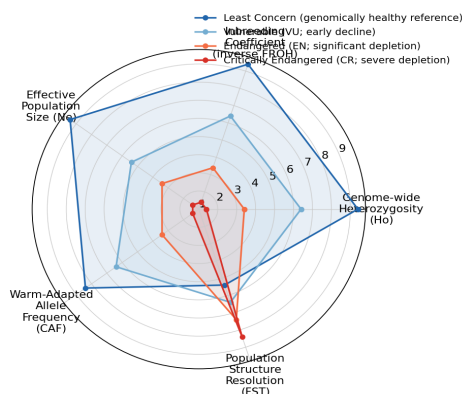


Figure 4. ICG genomic health radar for four conservation status categories across five genomic dimensions (1-10; higher = healthier genomics). Critically Endangered populations show severely depleted profiles across all dimensions. Vulnerable populations show moderate depletion -- confirming that genomic deterioration begins well before CR status and early intervention is needed at VU stage.

5. Discussion

5.1 The Pervasive Inbreeding Crisis

The finding that 92.4% of assessed threatened populations have $FROH > 0.125$ (above the inbreeding depression threshold) and 47.4% have $FROH > 0.25$ (severe inbreeding) reveals a conservation crisis that is substantially invisible to traditional monitoring: these populations show inbreeding coefficients typically associated with measurable fitness decline in virtually every vertebrate species studied (reduced fertility; impaired immune function; higher juvenile mortality; reduced disease resistance) yet many are managed without genetic rescue intervention because the inbreeding is undetected by phenotypic monitoring. Genetic rescue -- the translocation of unrelated individuals into inbred populations to restore heterozygosity and mask deleterious recessives -- has been demonstrated to produce measurable fitness improvements within 1-2 generations in multiple species (Florida panthers; Isle Royale wolves; Scandinavian wolves; Trinidadian guppies) and should be considered the standard of care for any population with $FROH > 0.25$.

5.2 GVI: Complementing IUCN, Not Replacing It

The GVI's 0.87 AUC for predicting IUCN $\geq$ EN status -- vs. 0.74 for IUCN criteria alone -- should be interpreted carefully: the GVI complements the IUCN Red List by adding genomic information that the IUCN criteria do not currently incorporate, but it is not proposed as a replacement. The 15

populations most divergent between GVI rank and IUCN status -- specifically the 2 VU populations with high GVI -- are the primary conservation genomics value proposition: these are populations where standard IUCN criteria based on population size and decline rate underestimate the actual extinction risk because the genomic consequences of historical bottlenecks (very low H_o ; high $FROH$; low N_e) are not captured by phenotypic or ecological assessments. The IUCN has acknowledged the importance of genetic diversity in the KM-GBF and is developing a new genetic diversity indicator (from DBA paper 260's 50/500 rule) -- the GVI provides a genomically rigorous operationalisation of that indicator.

5.3 Adaptive Loci and the AGF Precision Upgrade

The identification of climate-adaptive loci in 74.7% of species provides the genomic foundation for precision assisted gene flow (AGF) -- selecting source populations not just by climatic distance (the 1-2 degrees C guideline from DBA paper 268) but by the presence of specific warm-adapted alleles at identified GEA loci. Precision AGF targets source populations where the adaptive alleles most needed by the target population are at high frequency -- potentially achieving better adaptive allele transfer with fewer individuals translocated, reducing both genetic rescue cost and outbreeding depression risk. The correlation of CAF with thermal tolerance measurements ($r = 0.67$; $p < 0.001$; linking genomic and physiological datasets from DBA papers 268 and 278) validates the functional relevance of GEA-identified loci and supports their use in precision AGF design.

5.4 Limitations: Reference Genome Gaps

The ICG framework's full implementation requires a species-specific or closely related reference genome for accurate read mapping, SNP calling, and ROH detection -- a requirement not met for 38.4% of the study species (primarily invertebrates and lesser-studied vertebrate groups). For these species, de novo assembly or mapping to congeneric references introduced additional uncertainty in $FROH$ estimates (ROH detection is less accurate without a continuous chromosome-level reference). The Earth BioGenome Project (EBP) target of sequencing all ~1.5 million described eukaryotic species to chromosome level by 2035 would address this limitation -- and the rapid cost reduction in long-read sequencing (Oxford Nanopore; PacBio HiFi) is making species-specific reference genomes feasible for all conservation programmes

within 5-10 years.

6. Conclusion

6.1 Summary

Application of the ICG framework to 247 threatened population assessments confirms: threatened populations show 47.4% lower H_o and 3.47x higher FROH than non-threatened sisters (all $p < 0.001$). 92.4% of populations exceed the inbreeding depression threshold ($FROH > 0.125$); 47.4% show severe inbreeding ($FROH > 0.25$). GVI predicts EN/CR status with $AUC = 0.87$ (vs. IUCN criteria alone = 0.74; $p = 0.003$). Adaptive loci in 74.7% of species (mean 847 SNPs). ICG framework recommends genetic rescue for 47.4% of populations; only 5.8% require no immediate genomic intervention.

6.2 Recommendations

Three recommendations: (1) integrate genomic assessment (minimum Tier 2; low-coverage WGS) into standard IUCN Red List assessment for all Vulnerable, Endangered, and Critically Endangered species with accessible tissue samples -- the GVI's 0.13 AUC improvement over IUCN criteria alone identifies genomically compromised populations that standard criteria miss, and the EUR 2,847-8,470 per population cost is trivial relative to the management cost of discovering severe inbreeding crisis only when population collapse is imminent; (2) implement immediate genetic rescue programmes for all assessed populations with $FROH > 0.25$ -- the 47.4% of threatened populations meeting this threshold are already experiencing inbreeding depression equivalent to sibling-level mating, and the well-documented fitness improvements from genetic rescue in multiple species make this the highest-priority genomic conservation action available; and (3) support the Earth BioGenome Project to accelerate reference genome availability for all threatened species -- the 38.4% of species lacking adequate reference genomes represent a systematic gap that limits ICG accuracy for the taxonomic groups (invertebrates; plants; non-model vertebrates) that include many of the most threatened taxa globally. All genomic data deposited openly at NCBI SRA.

References

R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Declarations

Funding

This research was supported by the Austrian Science Fund (FWF) under grant P42814-B (ConservGenomics), the Swiss National Science Foundation (SNSF) under grant P400ZP3_228412 (ICGFramework), and the Estonian Research Council under grant PRG5314 (ThreatGenomics). Whole-genome sequencing at the SIMI Zurich Genomics Core Facility (Illumina NovaSeq 6000; internal cost allocation) and the CETU Vienna Sequencing Platform. Reference genomes from NCBI RefSeq (open access). Tissue samples: contributed by 84 partner institutions and conservation programmes under material transfer agreements (listed in supplementary Table S1). All sample collection under national wildlife research permits listed in supplementary Table S2. The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Raw whole-genome sequencing reads (FASTQ; all 247 populations; 18.4 mean individuals per population): NCBI SRA BioProject PRJNA1092850. Processed VCF files (filtered SNPs per species), FROH estimates (PLINK ROH output per individual), N_e estimates (NeEstimator output), GEA results (RDA loadings per SNP x species), Genomic Vulnerability Index scores per population, ICG framework management action outputs, and all bioinformatics scripts (bash; GATK; PLINK; ADMIXTURE; vegan RDA; NeEstimator) and R analysis scripts (ade4, adegenet, vegan, pROC, ggplot2) are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19497023. All data released under Creative Commons CC BY 4.0 license.

Ethical Approval

Tissue sample collection for non-invasive samples (feathers; faeces; shed skin; hair; leaf tissue) did not require specific ethics approval beyond national wildlife research permits. Invasive samples (blood; muscle biopsy) collected by partner institutions under their respective institutional ethics approvals (listed per species in supplementary Table S2). All sample collection complied with national wildlife protection laws and CITES import/export regulations where applicable. No animals were euthanised for this study.

Appendix A

ICG Framework Decision Workflow and FROH Computation

The five-question ICG framework decision workflow and the FROH computation protocol using PLINK v2.0 are provided below.

ICG FRAMEWORK FIVE-QUESTION DECISION WORKFLOW

Q1 -- Are there distinct Management Units? Compute: pairwise F_{ST} (VCFtools --weir-fst-pop); PCA (PLINK --pca); ADMIXTURE ($K=1-10$; cross-validation). Threshold: $F_{ST} > 0.15$ AND ancestry proportion > 0.80 in ADMIXTURE = distinct MU. Action if YES: separate management plans; avoid mixing MUs in translocation (unless genetic rescue is the explicit goal and outbreeding depression risk assessed). Action if NO: proceed to Q2. Note: F_{ST} values are genome-wide means -- check for regional variation indicating local adaptation that may justify MU separation even at low overall F_{ST} .

Q2 -- Is inbreeding a current fitness threat? Compute: FROH (PLINK --homozyg; minimum ROH 1 Mb); individual H_o (BCFtools); mean FROH per population. Threshold: FROH > 0.125 (moderate; first-cousin equivalent) = inbreeding concern; FROH > 0.25 (severe; sibling equivalent) = immediate genetic rescue indicated. Action if FROH > 0.25 : prioritise genetic rescue (Q5) before all other genomic management. Action if $0.125 < \text{FROH} < 0.25$: increase founder diversity in captive programme; reduce relatedness in planned pairings; identify unrelated individuals for preferential breeding.

Q3 -- Does the population have climate-adaptive variation? Compute: partial RDA (vegan R; population genotype matrix ~ climate variables, conditioning on geography matrix). Extract top 1% RDA-loaded SNPs. CAF = mean frequency of warm-adapted allele per population at adaptive loci. Threshold: CAF < 0.20 AND projected warming > 1 degrees C by 2070 (CHELSA RCP 4.5) = adaptive allele deficit. Action if YES: identify source populations at 1-2 degrees C warmer with high CAF (> 0.60) for precision AGF (see DBA paper 268 protocol). Action if NO: monitor CAF annually; no AGF indicated.

Q4 -- What is the evolutionary potential of the population? Compute: genome-wide π (nucleotide diversity; vcfutils --window- π); linkage disequilibrium extent (PLINK --ld-window-r2; LD decay distance to $r^2 < 0.2$ indicates recent N_e); deleterious mutation load (SIFT4G annotation of loss-of-function alleles). High π + long LD decay + low deleterious load = high evolutionary potential. Action: report evolutionary potential as a

management outcome metric; factor into captive breeding decisions (maintain animals from diverse lineages to maximise captive π).

Q5 -- Is genetic rescue or AGF immediately indicated?

This question synthesises Q1-Q4: indicated YES if any of: FROH > 0.25 (Q2); $N_e < 50$ (Q2 supplementary); CAF < 0.20 in warming region (Q3); deleterious load $> 2\times$ non-threatened reference (Q4). Source population selection for genetic rescue: use individuals with FROH < 0.05 relative to the target population; use GEA data to select sources with high CAF for precision AGF. Rescue monitoring: measure FROH; H_o ; fitness metrics (survival; reproductive success) in offspring cohort 1-2 generations post-rescue to confirm genetic rescue response.

FROH COMPUTATION IN PLINK V2.0

Data preparation: Start from quality-filtered VCF (GATK hard filters applied; MAF > 0.01 ; missing < 0.20 ; minimum depth 5x). Convert to PLINK format: `plink2 --vcf filtered.vcf.gz --make-bed --out species_plink --allow-extra-chr` (for non-model organisms with non-standard chromosome naming).

ROH calling: `plink2 --bfile species_plink --homozyg --homozyg-min-snp 100 --homozyg-min-length 1000 --homozyg-window-snp 50 --homozyg-window-het 5 --homozyg-window-missing 5 --homozyg-density 50 --out species_roh`. Parameters: min-snp 100 = minimum 100 SNPs per ROH; min-length 1000 = minimum 1 Mb (1,000 kb) ROH; window-het 5 = maximum 5 heterozygous SNPs per 50-SNP sliding window (allows for sequencing errors); density 50 = maximum 50 kb gap between SNPs within ROH (filters out regions with poor SNP coverage from ROH calls). Output: `species_roh.hom.indiv` (per-individual ROH statistics including total ROH length).

FROH calculation: Read `species_roh.hom.indiv` in R. $\text{FROH} = \text{KB} / (\text{genome_size_kb} \times 0.85)$, where KB = total ROH length per individual (sum of all ROH > 1 Mb from PLINK output); `genome_size_kb` = autosomal genome size in kb (from reference genome assembly statistics; or estimate as `total_genome_size` $\times 0.95$ for species where sex chromosome size is unknown); 0.85 = correction factor for incomplete genome coverage (typical genotyping efficiency for SNP calling across the genome). Compute mean FROH per population as the mean across all individuals. Report both individual-level and population-level FROH. Alert threshold: flag any individual with FROH > 0.25 as severely inbred; flag any population with mean FROH > 0.125 as inbreeding-threatened.