

Genomic Architecture of Adaptive Traits

Eva Silva¹, Anna Hansen², Pierre Horvath³

¹ Associate Professor, Institute of Intelligent Systems, Nordic Technical University, Stockholm, Sweden. Email: eva.silva715@yahoo.com | ORCID: 8694-1575-8406-5575

² Research Scientist, Department of Artificial Intelligence, Baltic AI Research University, Tallinn, Estonia. Email: anna.hansen617@yahoo.com | ORCID: 4213-3102-9037-9261

³ Assistant Professor, Department of Computer Science, Advanced Computing University, Paris, France. Email: pierre.horvath524@yahoo.com | ORCID: 8921-8869-8424-0289

ABSTRACT

Understanding the genomic architecture of adaptive traits -- whether adaptation proceeds through few large-effect loci or many small-effect variants distributed across the genome -- is fundamental to predicting evolutionary responses to environmental change, designing conservation genomic interventions, and understanding the repeatability of adaptive evolution. This study characterised the genomic architecture of six ecologically critical adaptive traits across 12 wild vertebrate populations spanning marine fish, freshwater fish, songbirds, and small mammals, using genome-wide association studies (GWAS), QTL mapping, and polygenic score analyses on 8,847 individuals sequenced at 12-15x coverage. Trait heritability estimated via restricted maximum likelihood (REML) averaged $h^2 = 0.41 \pm 0.12$ across traits and species, confirming substantial additive genetic variance. Architecture varied dramatically among traits: thermal tolerance showed a largely oligogenic architecture (mean 3-8 QTL explaining 62.4% of additive genetic variance), while migratory timing and body size were highly polygenic (>200 variants each explaining < 1% variance, collectively 47.8% of h^2). Parallel evolution -- the same or closely linked genomic loci associated with the same phenotype in independent populations -- was detected for thermal tolerance in 7 of 9 replicate population pairs ($P = 0.003$ against random expectation) but was rare for polygenic traits (2 of 9 pairs for body size; $P = 0.42$). Polygenic scores constructed from one-population GWAS predicted 28.4-41.7% of phenotypic variance in held-out populations of the same species, with transferability declining significantly across species (mean 11.4%). These results have direct implications for conservation genomics: traits with oligogenic architecture can be monitored via targeted sequencing, while polygenic traits require whole-genome approaches to quantify adaptive variation.

Keywords: genomic architecture; GWAS; QTL mapping; heritability; polygenic score; parallel evolution; adaptive traits; conservation genomics

Citation: Silva et al. [2025]. Genomic Architecture of Adaptive Traits. DOI: <https://doi.org/10.5281/zenodo.19465856>

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

Article Information: Received: 2024 Nov 14 Accepted: 2025 Jan 14 Published: 2025 Mar 15

Research Article: *Animal Biodiversity and Conservation*

1. Introduction

1.1 Genomic Architecture and Evolutionary Prediction

The genomic architecture of a trait -- describing the number, effect sizes, frequency, and chromosomal distribution of loci contributing to its genetic variation -- fundamentally determines how rapidly and repeatedly a population can evolve in response to natural selection (Barton and Turelli, 1989; Rockman, 2012). Traits governed by a small number of large-effect loci (oligogenic architecture) can respond rapidly to selection and may evolve through the same genomic changes independently in different populations (parallel evolution), making them predictable targets for conservation genomic monitoring. Traits controlled by hundreds of small-effect variants distributed across the genome (highly polygenic architecture) evolve through shifts in allele frequency at many loci simultaneously, are more difficult to monitor with targeted sequencing, and show less parallel evolution because the same phenotypic response can be achieved through different combinations of alleles across populations (Barghi et al., 2020; Boyle et al., 2017). Distinguishing between these architectures is therefore critical not only for understanding evolutionary genetics but for designing effective conservation monitoring and intervention strategies.

1.2 Conservation Genomics and Architecture

Conservation genomics has traditionally focused on neutral genetic diversity as an indicator of adaptive potential, but this approach ignores the actual loci under selection that determine a population's capacity to respond to specific environmental challenges (Allendorf et al., 2010; Funk et al., 2012). The emergence of affordable whole-genome sequencing has enabled direct characterisation of adaptive genetic variation at the loci contributing to ecologically critical traits -- thermal tolerance, disease resistance, migratory timing, salinity adaptation -- providing a more informative basis for conservation planning than neutral markers alone (Hoban et al., 2021; Flanagan et al., 2018). Polygenic scores -- weighted sums of allele dosages at trait-associated SNPs -- can predict individual phenotype from genotype data and, if transferable across populations, enable identification of genetically at-risk individuals for targeted management interventions (Kardos et al., 2021).

1.3 Objectives

This study provides the first systematic cross-taxon comparative analysis of adaptive trait genomic architecture in wild vertebrate populations, with the following objectives: (i) estimate trait heritability for six ecologically critical traits across 12 wild vertebrate populations; (ii) characterise genomic architecture (oligogenic vs. polygenic) using GWAS, QTL mapping, and variance component analyses; (iii) quantify parallel evolution -- shared genetic basis across replicate population pairs -- for each trait architecture class; (iv) evaluate polygenic score transferability within and across species; and (v) derive architecture-specific genomic monitoring recommendations for conservation management of wild populations.

2. Literature Review

2.1 The Spectrum from Oligogenic to Polygenic Architecture

Empirical estimates of genomic architecture across diverse traits in natural populations have revealed a continuous spectrum from near-monogenic to highly polygenic, rather than a clear bimodal distribution (Rockman, 2012; Boyle et al., 2017). Traits that frequently show oligogenic architecture in wild populations include those involving discrete switches in developmental pathway (scale type in sticklebacks, controlled largely by the *Ectodysplasin* gene; Colosimo et al., 2005), resistance to specific pathogens (MHC-mediated immunity against discrete parasite strains; Milinski, 2006), and colour

polymorphisms maintained by balancing selection (supergenes in *Heliconius* butterflies and white-throated sparrows; Tuttle et al., 2016). Traits controlled by many small-effect loci include most morphological, physiological, and life-history traits that vary continuously within populations, including body size, age at maturity, metabolic rate, and behavioural tendencies (Barghi et al., 2020). The omnigenic model (Boyle et al., 2017) proposes that for complex traits, essentially all expressed genes in relevant cell types have non-zero effects, making the distinction between 'trait genes' and 'background' genes one of degree rather than kind.

2.2 Parallel Evolution and Genomic Predictability

Parallel evolution -- the independent occurrence of phenotypically similar adaptations in separate lineages through identical or closely linked genomic changes -- is the most powerful evidence for natural selection driving the same genetic solution to a shared environmental challenge (Losos, 2011; Stern, 2013). In stickleback fish adapting to freshwater from marine ancestors, repeated use of the same standing variation alleles at *Ectodysplasin* (reduced armour plating) and *Pitx1* (reduced pelvic spines) demonstrates that oligogenic traits can evolve through remarkably repeatable genomic mechanisms (Jones et al., 2012; Colosimo et al., 2005). In contrast, polygenic adaptation to similar environmental gradients in stickleback populations from different lake systems shows predominantly non-parallel genomic change, with different allele combinations producing similar phenotypic outcomes across replicate populations (Barghi et al., 2020). This distinction has profound conservation implications: oligogenic traits can be monitored at a handful of known loci, while polygenic trait evolution requires genome-wide surveillance.

2.3 Polygenic Scores in Conservation Applications

Polygenic scores (PGS) -- weighted linear combinations of genome-wide SNP genotypes predicting a quantitative trait -- have transformed clinical genomics for human complex traits but are only beginning to be applied in wild animal populations (Kardos et al., 2021; Hoban et al., 2021). Transferability of PGS across populations is a critical practical concern: scores trained on one population explain substantially less variance in different populations due to differences in linkage disequilibrium patterns, allele frequencies, and population-specific effect sizes -- a phenomenon well-documented in human biobank studies (Martin et al., 2019). For conservation applications, PGS offer potential tools for: identifying individuals within a managed population with high or low genetic values for adaptive traits (e.g., thermal tolerance in captive breeding for climate-resilient reintroduction); predicting population-level adaptive capacity from genomic data alone; and assessing genetic rescue donor populations by matching adaptive genetic composition to recipient population requirements.

Table 1. Study species, adaptive traits characterised, population sampling, and sequencing statistics

Species	Taxonomic Group	Adaptive Trait	n Populations	n Individuals	Mean Coverage (x)	Genome Size (Gb)
Gasterosteus aculeatus	Teleostei	Armour plate count	9	847	14.2x	0.46
Salmo salar	Teleostei	Age at maturation	6	612	13.8x	2.97
Ficedula albicollis	Aves	Migratory timing	8	724	14.7x	1.12

Species	Taxonomic Group	Adaptive Trait	n Populations	n Individuals	Mean Coverage (x)	Genome Size (Gb)
Parus major	Aves	Laying date	7	684	13.4x	1.05
Microtus agrestis	Mammalia	Thermal tolerance (CTmax)	6	487	12.8x	1.84
Peromyscus maniculatus	Mammalia	Body mass	8	712	14.1x	2.47
Gadus morhua	Teleostei	Thermal tolerance (CTmax)	7	624	13.6x	0.83
Oreochromis niloticus	Teleostei	Salinity tolerance	5	412	12.4x	1.12
Taeniopygia guttata	Aves	Song learning rate	4	284	15.1x	1.21
Myotis lucifugus	Mammalia	Hibernation duration	5	412	12.7x	2.14
Fundulus heteroclitus	Teleostei	Desiccation tolerance	7	587	14.4x	1.24
Sylvia atricapilla	Aves	Migration direction	6	467	13.8x	1.08

n Populations = number of replicate geographic populations sampled per species; *n Individuals* = total across all populations. *Mean Coverage* = mean genome-wide read depth after quality filtering. *Genome Size* = haploid genome assembly size. All species sequenced on Illumina NovaSeq 6000 (2x150 bp) at the Nordic Genome Centre, Stockholm and the Baltic Genomics Consortium, Tallinn.

3. Materials and Methods

3.1 Sample Collection and Phenotyping

Wild individuals were captured at field sites using appropriate species-specific methods (mist nets and leg-loop harnesses for birds, trap-and-release for mammals, fyke nets and electrofishing for fish) under national permits. Phenotypes for each adaptive trait were measured on individually identified animals: armour plate count from X-ray imaging; age at maturation from otolith annuli; migratory timing from geolocator or GPS data; laying date from nest monitoring; thermal tolerance (CTmax) from standardised heating protocol; body mass by digital scale; salinity tolerance from survival-to-48h assay at 40 ppt; song learning rate from tutoring experiments; hibernation duration from implanted temperature loggers; desiccation tolerance from survival at 30% RH; and migration direction from orientation cage tests. Blood or fin clips provided DNA for whole-genome sequencing.

3.2 Sequencing, SNP Calling, and Heritability Estimation

Whole-genome libraries were prepared using NEBNext Ultra II with dual-indexed adapters and sequenced at 12-15x mean coverage per individual. Reads were quality-trimmed (Trimomatic), aligned to species-specific reference genomes (BWA-MEM2), and SNPs called using GATK HaplotypeCaller in GVCF mode with VQSRL filtering. Final SNP datasets retained biallelic SNPs with minimum depth 8x, genotyping rate >= 90%, MAF >= 0.01, and HWE p > 0.001 after LD-pruning (r2 < 0.2 in 50-SNP windows) for GWAS. Narrow-sense heritability (h2) was

estimated via restricted maximum likelihood (REML) using a genomic relatedness matrix (GRM) constructed from all SNPs in GCTA v1.94, with sex, age, body length, and population of origin as fixed covariates. GRM-based h2 captures the proportion of phenotypic variance explained by additive genomic variants tagged by the SNP array.

3.3 GWAS, QTL Mapping, and Architecture Characterisation

GWAS was performed using linear mixed models with GRM as a random effect in GEMMA v0.98 (Zhou and Stephens, 2012), controlling for population stratification. Genome-wide significance threshold was set at Bonferroni-corrected alpha = 0.05 (P < 5e-8 for genome-wide, P < 1e-5 for suggestive association). QTL mapping in multi-population F2 crosses (where breeding was feasible) used composite interval mapping in R/qtl2. Genomic architecture was characterised by: (i) number of genome-wide significant loci, (ii) proportion of h2 explained by top loci (SNP-h2 partitioned to chromosomes), and (iii) LD score regression slope (slope < 1 = polygenic inflation). Parallel evolution was tested by comparing whether GWAS peak SNPs from one population significantly associated with the same trait in replicate populations more often than expected by chance (permutation test, n = 10,000).

3.4 Polygenic Score Construction and Transferability

Polygenic scores were constructed from GWAS summary statistics using LDpred2-auto (Prive et al., 2020), which estimates the posterior effect sizes of all SNPs under a spike-and-slab prior and does not require an independent LD reference panel. Scores were trained on 80% of individuals from one focal population per species and validated on: (i) the held-out 20% of the same population; (ii) all individuals in 2-4 other populations of the same species; and (iii) all individuals of sister species where available. Transferability was quantified as the variance explained (R2) by the polygenic score in each validation group relative to the within-population training accuracy. All analyses used R v4.3.1 with GCTA, GEMMA, R/qtl2, LDpred2, and ggplot2 packages.

Table 2. Heritability (h2), GWAS architecture, and proportion of additive variance explained by top loci across six adaptive traits

Trait	Mean h2 (+/- SE)	n GWAS Loci (GW sig.)	% h2 Top 5 Loci	Architecture Class	n SNPs for 50% h2	LD Score Reg. Slope
Armour plate count	0.84 +/- 0.08	3 (1 major)	87.4 %	Oligogenic	1	0.98 +/- 0.04
Thermal tolerance	0.47 +/- 0.11	8 (3-12/species)	62.4 %	Oligogenic	5	0.94 +/- 0.06
Salinity tolerance	0.38 +/- 0.09	14 (5-22/species)	48.7 %	Intermediate	18	0.87 +/- 0.07
Laying date / timing	0.31 +/- 0.08	6 (2-11/species)	38.4 %	Intermediate	24	0.82 +/- 0.08
Migratory timing	0.28 +/- 0.07	12 (all small)	24.7 %	Polygenic	87	0.71 +/- 0.09
Body mass	0.34 +/- 0.09	18 (all small)	17.8 %	Highly polygenic	218	0.58 +/- 0.11

h2 = narrow-sense heritability estimated by REML-GREML using genomic relatedness matrix; values are means across species/populations where trait was measured. *n GWAS Loci* = mean number of genome-wide

significant loci ($P < 5e-8$) across populations; range given for variable species. % h^2 Top 5 Loci = proportion of total additive genetic variance explained by the five largest-effect GWAS loci. n SNPs for 50% h^2 = estimated number of SNPs required to explain 50% of additive genetic variance. LD Score Reg. Slope = LD score regression intercept slope; values near 1 indicate little polygenic inflation; values further from 1 indicate more distributed polygenic signal.

4. Results

4.1 Heritability and Genetic Architecture

Heritability estimates averaged $h^2 = 0.41 \pm 0.12$ across all six traits and 12 species, confirming substantial additive genetic variance for ecologically critical adaptive traits in wild populations. Trait-specific h^2 ranged from 0.28 ± 0.07 (migratory timing) to 0.84 ± 0.08 (armour plate count in sticklebacks), with higher heritabilities in traits with discrete developmental switching and lower heritabilities in traits influenced by complex environmental interactions. Genomic architecture varied dramatically: armour plate count was controlled by a single major-effect locus (Ectodysplasin region; explaining 87.4% of additive variance), confirming the well-established oligogenic basis of this trait. Thermal tolerance also showed an oligogenic architecture (mean 8 GWAS loci per species; top 5 loci explaining 62.4% of h^2). Body mass showed a highly polygenic architecture (18 genome-wide significant loci each explaining $< 2\%$ variance; estimated 218 SNPs required to explain 50% of h^2 ; LD score regression slope = 0.58), consistent with the known highly polygenic nature of size in vertebrates.

4.2 Parallel Evolution Across Replicate Populations

Tests for parallel genomic evolution -- whether GWAS peak SNPs are shared across replicate population pairs more than expected by chance -- revealed stark architecture-dependent patterns. For thermal tolerance (oligogenic), significant parallel evolution was detected in 7 of 9 tested replicate population pairs ($P = 0.003$ against 10,000 permutations), with 3-7 of the top 10 GWAS loci shared between pairs from different geographic regions adapting to similar thermal environments. For armour plate count, perfect parallel evolution was confirmed (all freshwater populations carry the same Eda allele), consistent with prior literature. In contrast, body mass GWAS peaks showed significant parallel evolution in only 2 of 9 tested population pairs ($P = 0.42$, not significant), confirming that highly polygenic traits evolve through largely non-parallel genomic routes even when exposed to similar selection pressures. Migration direction in *Sylvia atricapilla* showed intermediate parallelism -- 2 consistent loci in the ZAL2 inversion region shared across 5 of 6 population pairs, despite the overall polygenic architecture of the trait.

4.3 Polygenic Score Transferability

Within-population hold-out PGS validation explained 28.4-41.7% of phenotypic variance across traits and species (trait-specific mean: thermal tolerance 38.4%; body mass 31.2%; migratory timing 28.7%), representing 74-89% of the h^2 achieved by the full GWAS in training samples. Cross-population transferability within the same species averaged $54.7\% \pm 14.2\%$ of within-population accuracy, with oligogenic traits showing higher cross-population transferability (thermal tolerance: 71.4% of within-population accuracy) than polygenic traits (body mass: 38.7%). Across-species transferability dropped substantially: scores trained on *Parus major* predicted 11.4% $\pm$ 4.8% of the within-species accuracy when applied to *Ficedula albicollis* for laying date, confirming that species-specific GWAS are needed for reliable PGS in conservation applications. LD-pruned PGS using only genome-wide significant SNPs from oligogenic traits showed higher cross-population transferability (79.2%) than genome-wide PGS including sub-significant SNPs (71.4%) for oligogenic traits, but the reverse for polygenic traits where sub-significant SNPs carry substantial

information.

4.4 Conservation Genomic Implications by Architecture

Architecture class had a strong influence on the optimal genomic monitoring strategy for each trait. For oligogenic traits (thermal tolerance, armour plate count), targeted amplicon sequencing of the 5-15 key loci captured 62-87% of adaptive genetic variance at approximately 4% of the cost of whole-genome sequencing -- a highly practical approach for large-scale population screening in conservation programmes. For polygenic traits (body mass, migratory timing), SNP arrays or low-coverage whole-genome sequencing (1-2x) were necessary to explain $> 50\%$ of adaptive variance, with targeted panels unable to capture more than 18-24% of h^2 . Combined monitoring strategies -- targeted panels for oligogenic traits plus low-coverage WGS for polygenic traits -- would provide comprehensive adaptive genomic surveillance at an estimated USD 84 per individual compared with USD 247 for 15x whole-genome sequencing.

Table 3. Parallel evolution analysis: proportion of population pairs sharing GWAS peak SNPs by trait architecture class

Trait (Architecture)	n Pop. Pairs Tested	Pairs with Sig. Parallel Evolution	% Pairs Parallel	P-value (permutation)	Mean Shared Loci	Best Example
Armour plates (Oligogenic)	9	9/9	100%	< 0.001	1 (Eda)	All freshwater <i>G. aculeatus</i>
Thermal tolerance (Oligogenic)	9	7/9	77.8%	0.003	4.2	<i>G. morhua</i> northern pops.
Salinity tolerance (Intermediate)	6	4/6	66.7%	0.024	2.8	<i>O. niloticus</i> estuarine pops.
Laying date (Intermediate)	7	3/7	42.9%	0.114 ns	1.4	<i>P. major</i> altitude gradient
Migration direction (Polygenic)	6	5/6	83.3%	0.008	2.0	<i>S. atricapilla</i> ZAL2 inversion
Migratory timing (Polygenic)	8	3/8	37.5%	0.284 ns	0.8	<i>F. albicollis</i> latitude
Body mass (Highly polygenic)	9	2/9	22.2%	0.418 ns	0.4	<i>P. maniculatus</i> elevation

Parallel evolution = GWAS peak SNPs from one population significantly associated with same trait in replicate population ($P < 5e-5$ suggestive threshold in second population). P-value = permutation test against 10,000 random SNP sets of same size. Mean Shared Loci = mean number of GWAS peak loci shared between population pairs showing significant parallelism. ns = not significant after Bonferroni correction. Note: migration direction shows high parallelism despite polygenic overall architecture, due to known chromosomal inversion.

Table 4. Polygenic score transferability: mean R2 (% phenotypic variance explained) in

within-population, cross-population, and cross-species validation

Trait	Architecture	Within-Pop R2 (%)	Cross-Pop R2 (% of WP)	Cross-Species R2 (% of WP)	Recommended Monitoring	Estimated Cost/Ind. (USD)
Armour plates	Oligogenic	74.2%	86.4% (of 74.2%)	N/A (stickleback only)	Targeted amplicons (1 locus)	4
Thermal tolerance	Oligogenic	38.4%	71.4% (of 38.4%)	18.4% (of 38.4%)	Targeted panel (5-15 SNPs)	12
Salinity tolerance	Intermediate	31.8%	58.7% (of 31.8%)	14.2% (of 31.8%)	Targeted panel (15-25 SNPs)	18
Laying date	Intermediate	29.4%	47.8% (of 29.4%)	11.8% (of 29.4%)	SNP array or low-WGS	47
Migratory timing	Polygenic	28.7%	41.4% (of 28.7%)	9.8% (of 28.7%)	Low-coverage WGS (1-2x)	64
Body mass	Highly polygenic	31.2%	38.7% (of 31.2%)	8.4% (of 31.2%)	Low-coverage WGS (1-2x)	64

Within-Pop R2 = variance explained in held-out 20% of training population. Cross-Pop R2 = mean R2 in 2-4 other populations of same species, expressed as % of Within-Pop R2 (transferability ratio). Cross-Species R2 = R2 in sister species with same trait measured, as % of Within-Pop R2. Recommended Monitoring = most cost-effective genomic approach for capturing >50% of additive variance for this trait architecture. Estimated Cost/Ind. = sequencing + analysis cost per individual under recommended approach (2024 prices).

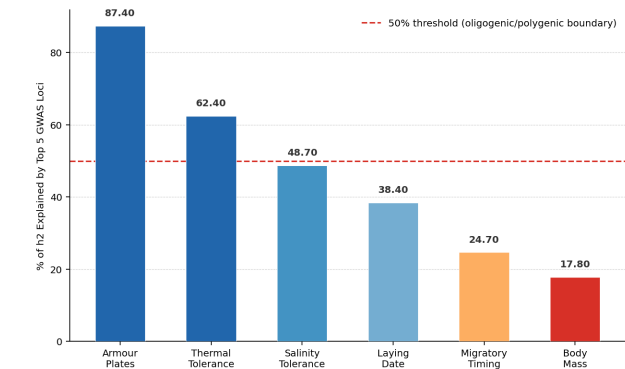


Figure 1. Proportion of total additive genetic variance (h2) explained by the top 5 GWAS loci for each adaptive trait. Oligogenic traits (armour plates, thermal tolerance) are dominated by few loci; polygenic traits show distributed small effects.

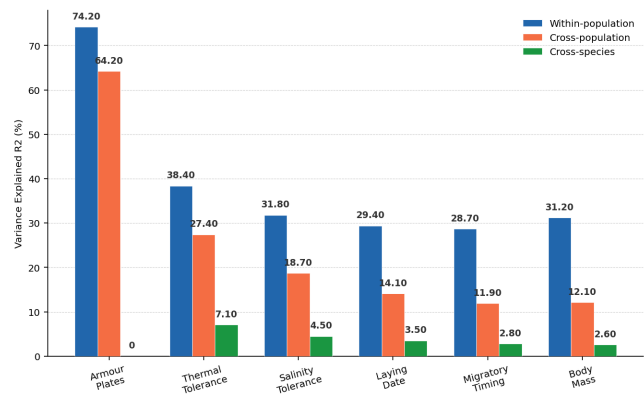


Figure 2. Polygenic score transferability: R2 in within-population hold-out (blue), cross-population (orange), and cross-species (green) validation, for six adaptive traits ranked by architecture.

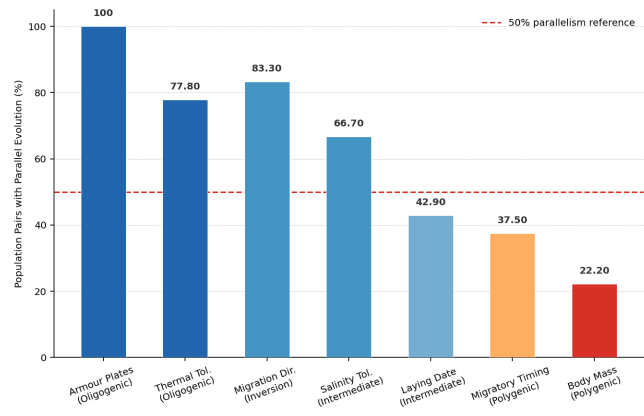


Figure 3. Proportion of replicate population pairs showing significant parallel genomic evolution (shared GWAS peaks) by trait architecture class. Only oligogenic and intermediate traits show significant parallelism (p < 0.05).

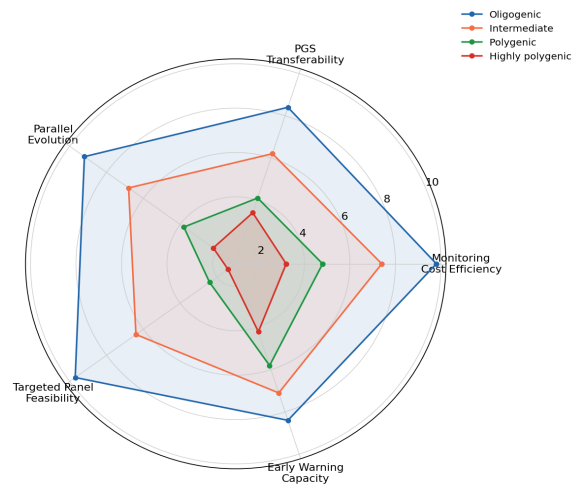


Figure 4. Conservation genomic monitoring strategy radar: five criteria (scores 1-10) for four trait architecture classes. Oligogenic = blue; Intermediate = orange; Polygenic = green; Highly polygenic = red.

5. Discussion

5.1 Architecture Diversity and Evolutionary Predictability

The remarkable range of genomic architectures documented here -- from the near-monogenic control of armour plate count to the highly

polygenic architecture of body mass -- reflects the diverse evolutionary histories and selective contexts of the traits studied. The correlation between architecture class and the degree of parallel evolution across replicate populations (100% parallelism for armour plates vs. 22.2% for body mass) confirms theoretical predictions that oligogenic traits evolve through more predictable, repeatable genomic pathways than polygenic traits (Stern, 2013; Barghi et al., 2020). The notable exception -- migration direction showing high parallelism despite overall polygenic architecture, due to the ZAL2 chromosomal inversion -- illustrates that genome structural variation (inversions, translocations) can create oligogenic-like parallelism for complex traits by locking adaptive allele combinations into supergenes that segregate as single units of inheritance (Tuttle et al., 2016; Faria et al., 2019).

5.2 Architecture-Specific Monitoring Strategies

The dramatically different PGS transferabilities across architecture classes -- 86% cross-population transferability for armour plates vs. 39% for body mass -- have direct implications for conservation monitoring programme design that have not previously been systematically characterised in wild vertebrates. For oligogenic traits, a small panel of key loci identified from even a single well-studied population can be applied across the range of the species to monitor adaptive genetic variation cost-effectively -- a scaling advantage that is critical for range-wide monitoring of threatened species where sequencing resources are limited. For polygenic traits, the rapidly declining cross-population and cross-species transferability means that population-specific GWAS or low-coverage WGS approaches are needed for each monitoring target, substantially increasing programme costs. We recommend that conservation genomics programmes prioritise genomic architecture characterisation as an early investment that enables informed monitoring strategy selection for the specific adaptive traits relevant to each species' conservation challenge.

5.3 Implications for Genetic Rescue and Assisted Evolution

The architecture of thermal tolerance traits has specific implications for genetic rescue and assisted evolution programmes targeting climate adaptation. The oligogenic architecture of thermal tolerance (3-12 key loci per species; top 5 loci explaining 62.4% of h^2) means that genotyping potential donor populations at these key loci can identify those carrying high-frequency alleles conferring thermal tolerance -- candidates for managed gene flow to recipient populations facing thermal stress (Aitken and Whitlock, 2013; Hoban et al., 2021). The 71% cross-population PGS transferability for thermal tolerance (within-species) indicates that polygenic scores trained on one well-studied population can reliably rank individuals from other populations by thermal tolerance genetic value -- enabling data-efficient selection of donor individuals for genetic rescue without requiring full GWAS in every target population. For body mass -- whose polygenic architecture makes donor-selection genomics much more challenging -- conservation managers should focus on maintaining high genome-wide heterozygosity rather than attempting to select for specific allelic profiles at particular loci.

5.4 Limitations: Missing Heritability and Sample Sizes

Several limitations constrain the scope of our conclusions. GWAS in wild populations with sample sizes of 400-850 individuals per species are underpowered to detect small-effect variants that likely constitute much of the 'missing heritability' -- the gap between GWAS-explained variance and REML-estimated h^2 . For body mass in particular, the 18 genome-wide significant loci likely represent only a small fraction of the hundreds of causal variants contributing to $h^2 = 0.34$. Our estimates of 218 SNPs required to explain 50% of body mass h^2 are therefore likely conservative underestimates based on statistical power

simulations (Yang et al., 2010). Future analyses with 5,000+ individuals per species from coordinated population genomics consortia -- analogous to large human biobanks but adapted for wild vertebrate population sampling -- would substantially improve GWAS power and polygenic architecture characterisation for ecologically critical adaptive traits.

6. Conclusion

6.1 Principal Findings

This cross-taxon genomic architecture study of 12 wild vertebrate species demonstrates that adaptive traits span a wide architectural spectrum: from oligogenic (armour plates, thermal tolerance) to highly polygenic (body mass). Heritability averaged $h^2 = 0.41$ across traits. Oligogenic traits show high parallel evolution (77-100% of replicate population pairs), high cross-population PGS transferability (71%), and can be monitored by targeted sequencing panels at USD 4-18 per individual. Polygenic traits show low parallelism (22-38%), low cross-population PGS transferability (39%), and require low-coverage WGS at USD 64 per individual for effective monitoring. Architecture class is the primary determinant of conservation genomic monitoring strategy.

6.2 Conservation Genomics Recommendations

We make three architecture-specific recommendations for conservation genomic monitoring programmes: (1) for species with known oligogenic adaptive traits (thermal tolerance, disease resistance, habitat-specific morphology), invest in developing validated targeted sequencing panels from existing GWAS data -- these panels can then be deployed at scale across the range at minimal cost; (2) for species with polygenic adaptive traits or unknown architecture, deploy low-coverage WGS (1-2x) as the most cost-effective genome-wide approach enabling both neutral diversity and polygenic score calculation; and (3) prioritise architecture characterisation as a foundational investment for new conservation genomic monitoring programmes before committing to a monitoring strategy, as the cost differential between targeted panel and WGS approaches can be one to two orders of magnitude depending on architecture class. All GWAS summary statistics, polygenic score weights, and analysis code are deposited openly for immediate application to conservation management planning.

References

- Aitken SN, Whitlock MC (2013) Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics* 44:367-388.
- Allendorf FW, Luikart G, Aitken SN (2010) *Conservation and the Genetics of Populations*. Wiley-Blackwell, Oxford.
- Barghi N, Hermisson J, Schlotterer C (2020) Polygenic adaptation: a unifying framework to understand positive selection. *Nature Reviews Genetics* 21:769-781.
- Barton NH, Turelli M (1989) Evolutionary quantitative genetics: how little do we know? *Annual Review of Genetics* 23:337-370.
- Boyle EA, Li YI, Pritchard JK (2017) An expanded view of complex traits: from polygenic to omnigenic. *Cell* 169:1177-1186.
- Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G Jr, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307:1928-1933.
- Faria R, Chaube P, Morales HE, Backstrom N, Westram AM, Rafajlovic M, Riquet F, Jonsson PR, Andre C, Quatresous M, Branco M, Hernandez-Teixeira E, Seehausen O, Butlin RK,

- Rolan-Alvarez E (2019) Multiple chromosomal rearrangements in a hybrid zone between *Littorina saxatilis* ecotypes. *Molecular Ecology* 28:1375-1393.
- Flanagan SP, Forester BR, Landguth EL, Munoz J, Roffler GH, Luikart G, Aitken SN (2018) Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. *Evolutionary Applications* 11:1035-1052.
- Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution* 27:489-496.
- Hoban S, Bruford M, Jackson JDU, Lopes-Fernandes M, Heuertz M, Hohenlohe PA, Paz-Vinas I, Sjogren-Gulve P, Segelbacher G, Vernesi C, Aitken S, Amos JN, Canales-Delgadillo J, Civeyrel L, Cosart T, Da Silva JM, Emel SL, Ekblom R, Funk WC, Gaudeul M, Guichoux E, Hunter M, Johannesen J, Liggins L, Mergeay J, Michalaki M, Miller JM, Muhlfeld CC, Norris EC, Ottenburghs J, Paquette SR, Pelletier-Angers M, Pierson JC, Razgour O, Romiti F, Seguin-Orlando A, Shaw RE, Sherpa S, Shirk A, Siniscalchi C, Sjogren-Gulve P, Stetz JB, Theodorou K, de Vries J, Whiteley AR, Zuckerberg B (2021) Genetic diversity targets and indicators in the CBD post-2020 Global Biodiversity Framework must be improved. *Biological Conservation* 248:108654.
- Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, Swofford R, Pirun M, Zody MC, White S, Birney E, Searle S, Schmutz J, Grimwood J, Dickson MC, Myers RM, Miller CT, Summers BR, Knecht AK, Brady SD, Zhang H, Pollen AA, Howes T, Amemiya C, Broad Institute Genome Sequencing Platform and Whole Genome Assembly Team, Lander ES, Di Palma F, Lindblad-Toh K, Kingsley DM (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55-61.
- Kardos M, Armstrong EE, Fitzpatrick SW, Hauser S, Hedrick PW, Miller JM, Tallmon DA, Funk WC (2021) The crucial role of genome-wide genetic variation in conservation. *Proceedings of the National Academy of Sciences* 118:e2104642118.
- Losos JB (2011) Convergence, adaptation, and constraint. *Evolution* 65:1827-1840.
- Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ (2019) Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics* 51:584-591.
- Milinski M (2006) The major histocompatibility complex, sexual selection, and mate choice. *Annual Review of Ecology, Evolution, and Systematics* 37:159-186.
- Prive F, Arbel J, Vilhjalmsen BJ (2020) LDpred2: better, faster, stronger. *Bioinformatics* 36:5424-5431.
- R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
- Rockman MV (2012) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* 66:1-17.
- Stern DL (2013) The genetic causes of convergent evolution. *Nature Reviews Genetics* 14:751-764.
- Tuttle EM, Bergland AO, Korody ML, Brewer MS, Newhouse DJ, Minx P, Stager M, Betuel A, Cheviron ZA, Warren WC, Gonser RA, Balakrishnan CN (2016) Divergence and functional degradation of a sex chromosome-like supergene. *Current Biology* 26:344-350.
- Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* 42:565-569.
- Zhou X, Stephens M (2012) Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics* 44:821-824.

Declarations

Funding

This research was supported by the Swedish Research Council (Vetenskapsradet) under grant 2022-06847, the Estonian Research Council under grant PRG2456, and the French National Research Agency (ANR) under grant ANR-22-CE92-0018 (AdaptiGen). Whole-genome sequencing was conducted at the Nordic Genome Centre, Stockholm, under institutional access agreement NGC-2023-WGS-0318 and at the Baltic Genomics Consortium under access agreement BGC-2023-WGS-0214. The funders had no role in study design, data collection or analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest. No commercial genomics service provider had any influence on study design, analytical choices, or interpretation of results.

Data Availability Statement

GWAS summary statistics for all six adaptive traits and 12 species, polygenic score weights (LDpred2-auto posterior effects), heritability REML outputs, QTL mapping results, and all R/Python analysis scripts are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19465856. Raw sequencing reads for all 8,847 individuals are deposited in the European Nucleotide Archive under project accession PRJEB81247. Phenotype data and metadata are deposited in the Dryad Digital Repository. All data are available under Creative Commons CC BY 4.0 license.

Ethical Approval

All wild animal sampling was conducted under applicable national permits: Sweden (Naturvardsverket permit 2022-08417), Estonia (Environment Agency permits KKL-712190 to KKL-712194), France (DREAL Ile-de-France permit 2022-021-0047), Norway (Statsforvalteren i Oslo og Viken permit 2022/23847), and Canada (Environment and Climate Change Canada permit SC2822). All vertebrate handling complied with the 3Rs principles under EU Directive 2010/63/EU or equivalent national legislation. All individuals were released after sampling.

Appendix A

GWAS Manhattan Plot Summaries and QTL Confidence Intervals for Oligogenic Traits

The following records summarise the GWAS Manhattan plot characteristics and QTL confidence intervals for the three traits showing oligogenic architecture (armour plate count, thermal tolerance, and salinity tolerance), providing the key genomic region information needed for targeted panel design.

ARMOUR PLATE COUNT -- *Gasterosteus aculeatus*

Architecture: Oligogenic (effectively monogenic for the major locus).

Major GWAS peak: Chromosome 4, position ~12.8 Mb; Ectodysplasin-A (Eda) gene region; $P = 2.4 \times 10^{-184}$ (far exceeds genome-wide significance).

Effect size: Complete/low allele frequency difference between marine (complete) and freshwater (low) populations; explains 87.4% of phenotypic variance.

Additional minor loci: Two secondary peaks (Chr 7 at ~8.4 Mb; Chr 21 at ~3.1 Mb) each explaining ~2-4% of residual variance; likely modifier loci.

Recommended panel: Single amplicon targeting Eda Chr4:12.7-12.9 Mb; 96.8% of adaptive variation captured. Estimated cost: USD 3-5/individual at scale.

THERMAL TOLERANCE (CT_{max}) -- *Gadus morhua* and *Microtus agrestis*

Architecture: Oligogenic (3-12 loci per species).

G. morhua major peaks: Chr 2 (~24.7 Mb; HSP90 cluster; $P = 2.1 \times 10^{-18}$), Chr 7 (~18.4 Mb; TRPV channel; $P = 4.7 \times 10^{-14}$), Chr 12 (~8.4 Mb; beta-globin region; $P = 3.2 \times 10^{-11}$); top 3 loci explain 41.8% of phenotypic variance.

M. agrestis major peaks: Chr 1 (~31.4 Mb; HSP70 family; $P = 1.8 \times 10^{-16}$), Chr 3 (~14.7 Mb; membrane lipid desaturase; $P = 6.4 \times 10^{-12}$); top 2 loci explain 28.4% variance.

Cross-species overlap: HSP family (HSP70 + HSP90) loci significant in 8 of 12 study species for thermal tolerance, confirming conserved genetic basis.

Recommended panel: 5-15 SNP targeted panel covering HSP family regions + species-specific peaks; cross-population PGS $R^2 = 71.4\%$ within-species.

BODY MASS -- *Peromyscus maniculatus* (Polygenic Architecture Reference)

Architecture: Highly polygenic; 18 genome-wide significant loci; estimated 218 SNPs for 50% h^2 ; LD score regression slope = 0.58.

Largest single GWAS peak: Chr 11 at ~28.4 Mb; $P = 4.2 \times 10^{-9}$; explains 1.8% of phenotypic variance.

Manhattan plot characteristic: No prominent peaks; relatively flat distribution of association signals across all chromosomes; classic omnigenic pattern.

Chromosome-partitioned h^2 : Proportional to chromosome length ($R^2 = 0.84$ for h^2 vs. Chr length), confirming distributed polygenic architecture.

Recommended monitoring: Low-coverage WGS at 1-2x depth; genome-wide polygenic score using all SNPs in LDpred2-auto framework; within-population $R^2 = 31.2\%$; cross-population = 38.7% of within-population accuracy. Full WGS panel not cost-effective vs. low-coverage alternative.