Zoological Archives: An International Journal
ISSN: 3117-695X
https://stanfordgroup.org/index.php/ZAIJ/issue/archive

Phylogeographic Patterns in Riverine Crustaceans

Sofia Bianchi[1], Andreas Rossi[2]

[1] Senior Lecturer, Department of Machine Learning, Nordic Technical University, Stockholm, Sweden. Email: sofia.bianchi372@gmail.com | ORCID: 6336-2401-3860-4145
[2] Research Scientist, Department of Computer Science, Central European Tech University, Vienna, Austria. Email: andreas.rossi625@gmail.com | ORCID: 5608-6042-7677-2995

ABSTRACT

Riverine crustaceans -- freshwater crayfish, river crabs, freshwater shrimps, and amphipods -- are among the most phylogeographically informative organisms in continental biogeography: obligately aquatic and largely unable to disperse overland between river drainages, their population genetic structure directly reflects historical river capture events, glacial refuge locations, drainage divide crossings, and post-glacial recolonisation pathways in a way that vertebrates -- with their greater dispersal capacity -- cannot. This study presents the most comprehensive multi-taxon phylogeographic analysis of European riverine crustaceans yet conducted, sequencing 1,847 individuals from 247 sites across 84 river drainages in 18 countries at three molecular markers (COI + 16S rRNA + one nuclear marker per taxon) for 18 focal crustacean species from four orders. Phylogeographic structure was strongly drainage-concordant for most species: mean FST between drainages = 0.487 +- 0.084 vs. within-drainage FST = 0.084 +- 0.028 (t = 18.4; p < 0.001). Bayesian phylogenetic analysis resolved 47 independent clades not currently recognised in taxonomic classifications, with 18 clades showing divergences > 3% COI consistent with cryptic species-level differentiation. Coalescent demographic modelling identified two major glacial refugia for European freshwater crustaceans -- the Danube basin and the Iberian drainages -- as primary sources of post-glacial recolonisation, consistent with the stream capture hypothesis for European freshwater biodiversity patterns. Drainage history reconstruction identified 8 likely historical river capture events explaining anomalous cross-drainage gene flow signatures.

Keywords: phylogeography; freshwater crustaceans; river drainages; cryptic species; COI; glacial refugia; stream capture; amphipods

Citation: Bianchi and Rossi [2024]. Phylogeographic Patterns in Riverine Crustaceans. DOI: https://doi.org/10.5281/zenodo.19489621

Article Information: *Received:* 2023 Dec 01 *Accepted:* 2024 Mar 01 *Published:* 2024 Mar 15

Research Article: *Freshwater Biogeography and Crustacean Systematics*

1. Introduction

1.1 Riverine Crustaceans as Biogeographic Proxies

Freshwater organisms present a fundamental paradox in biogeography: they inhabit a medium that is globally connected through the hydrological cycle yet are effectively isolated within individual drainage basins by the terrestrial barriers between them -- continental divides, watershed ridges, and the absence of direct hydrological connection between adjacent drainages. For obligately aquatic invertebrates with no aerial dispersal capacity -- particularly freshwater crayfish (Decapoda: Astacoidea), freshwater crabs (Decapoda: Potamoidea), freshwater shrimps (Atyidae; Palaemonidae), and amphipods (Gammarus and relatives) -- drainage boundaries function as absolute dispersal barriers over ecological timescales. Population genetic structure in these organisms therefore directly reflects the history of drainage connectivity and isolation over geological timescales: populations in the same drainage are genetically connected; populations in different drainages are genetically isolated. The exceptions -- where populations in different drainages share haplotypes or low FST -- provide direct evidence of historical stream capture events (where one drainage captured the headwaters of an adjacent drainage through erosional headward advance) that allowed temporary faunal exchange between previously isolated systems.

1.2 European Freshwater Biogeography: Glacial History

European freshwater biodiversity distribution was profoundly shaped by the Quaternary glaciations: ice sheets covering northern and central Europe during the Last Glacial Maximum (LGM; approximately 21,000 years ago) extirpated freshwater fauna from most of the continent, leaving populations only in southern refugia -- the Iberian Peninsula, the Italian peninsula, the Balkans, and the Danube basin south of the ice margin. Post-glacial recolonisation from these refugia produced the current distribution of freshwater invertebrates -- but also created a complex mosaic of overlapping phylogeographic contact zones where post-glacial colonists from different refugia have met and occasionally hybridised. Identifying the location and relative contribution of glacial refugia and post-glacial recolonisation routes is a central question in European freshwater biogeography.

1.3 Objectives

This study addresses four objectives: (i) characterise phylogeographic structure in 18 European riverine crustacean species across 84 river drainages; (ii) identify clades with cryptic species-level divergence; (iii) reconstruct glacial refugia and post-glacial recolonisation routes using Bayesian phylogeographic analysis; and (iv) identify historical stream capture events from anomalous cross-drainage gene flow signatures.

2. Literature Review

2.1 Cryptic Species in Freshwater Invertebrates

Freshwater invertebrates harbour among the highest rates of cryptic species -- morphologically indistinguishable taxa revealed as genetically distinct by molecular analysis -- of any invertebrate group. The genus Gammarus (freshwater amphipod) alone contains approximately 60 described European species and is estimated to harbour a similar number of undescribed cryptic taxa based on COI barcode analyses showing divergences of 3-15% between morphologically identical populations from adjacent drainages. Freshwater crayfish in Europe -- currently classified as 7 native species in 3 genera -- have been shown by multilocus analyses to contain 12-18 genetically distinct clades whose taxonomic status remains contested. The high rate of cryptic species in freshwater crustaceans reflects both their long drainage isolation (allowing allopatric divergence over millions of years without morphological change) and the relatively few taxonomic characters available in these groups for species-level discrimination.

2.2 Stream Capture as a Biogeographic Process

Stream capture -- the process by which a vigorously eroding drainage headward-advances to intercept and divert the upper reaches of an adjacent drainage -- is one of the primary mechanisms of freshwater faunal exchange between otherwise isolated drainage basins. After capture, the biota of the captured headwater can colonise the capturing drainage while biota from the capturing drainage colonise the captured system. The genetic signature of stream capture is cross-drainage sharing of haplotypes or low FST between populations in currently separate drainages -- detectable as anomalous clustering in haplotype networks or as unexpectedly low FST pairs in isolation-by-distance analyses that control for current hydrological distance.

2.3 Bayesian Phylogeographic Methods

Bayesian phylogeographic inference using BEAST2 with the ancestral area reconstruction model enables the explicit statistical testing of alternative biogeographic hypotheses -- glacial refugia locations; number of recolonisation lineages; direction of post-glacial expansion -- against the molecular clock-calibrated phylogenetic tree. The discrete phylogeographic model (diffusion model in geographic space; Lemey et al., 2009) allows statistical inference of ancestral geographic locations and migration rates between geographic areas as parameters of the phylogenetic model -- providing a more rigorous biogeographic reconstruction than post-hoc ancestral area interpretation of static phylogenies.

Table 1. Study species: taxonomy, sampling coverage, and phylogeographic structure overview

Species	Order	n Sites	n Individuals	Mean Within-drainage FST	Mean Between-drainage FST	n Cryptic Clades (COI > 3%)
CRAYFISH (5 spp.):	---	---	---	---	---	---
Austropotamobius pallipes (white-clawed crayfish)	Decapoda	38	247	0.054 +- 0.018	0.547 +- 0.084	8 (highly structured)
A. torrentium (stone crayfish)	Decapoda	28	184	0.047 +- 0.018	0.484 +- 0.074	6
Astacus astacus (noble crayfish)	Decapoda	28	184	0.074 +- 0.024	0.384 +- 0.064	3
[2 more crayfish spp.]	---	---	---	---	---	---
AMPHIPODS (7 spp.):	---	---	---	---	---	---
Gammarus fossarum (fossil freshwater amphipod)	Amphipoda	47	384	0.084 +- 0.024	0.547 +- 0.084	12 (most diverse)
Gammarus roeselii	Amphipoda	28	184	0.074 +- 0.024	0.484 +- 0.074	6
[5 more amphipod spp.]	---	---	---	---	---	---
FRESHWATER SHRIMPS (4 spp.):	---	---	---	---	---	---
Atyaephyra desmaresti	Decapoda	18	124	0.064 +- 0.024	0.414 +- 0.064	4
[3 more shrimp spp.]	---	---	---	---	---	---
ISOPODS (2 spp.):	---	---	---	---	---	---
Asellus aquaticus	Isopoda	38	247	0.047 +- 0.018	0.347 +- 0.054	3 (lowest; wider dispersal)
[1 more isopod spp.]	---	---	---	---	---	---
ALL (18 spp.; mean)	4 orders	247	1,847	0.084 +- 0.028	0.487 +- 0.084	47 total (mean 2.6/spp.)

n Sites = sites with >= 5 sampled individuals for this species. n Individuals = total successfully sequenced individuals. Within-drainage FST = mean pairwise FST between populations within the same river drainage (Weir and Cockerham 1984). Between-drainage FST = mean pairwise FST between populations from different drainages. n Cryptic Clades = number of COI lineages with > 3% mean pairwise divergence that are not represented in current taxonomic classifications (i.e., clades that would qualify as separate species under the 3% COI species threshold commonly applied in freshwater invertebrates). Gammarus fossarum (12 cryptic clades; highest) shows the most extreme example of cryptic diversity -- a species currently treated as a single taxon but likely comprising > 12 distinct species when COI divergence patterns are combined with morphometric and ecological data. Asellus aquaticus (3 cryptic clades; lowest between-drainage FST) shows the most connected phylogeographic structure -- consistent with this species' known ability to colonise new water bodies via waterfowl gut passage (passive dispersal vector absent in crayfish and most amphipods).

3. Materials and Methods

3.1 Specimen Collection and DNA Extraction

Crustacean specimens were collected from 247 sites in 84 river drainages across 18 European countries (2018-2022; national collection permits listed in supplementary Table S1). Crayfish and shrimps: hand collection and baited traps; muscle tissue from pereopod stored in 95% ethanol. Amphipods and isopods: Surber sampler and kick net; whole individuals preserved in 95% ethanol. DNA extracted using CTAB protocol for crayfish; DNeasy Blood and Tissue Kit (Qiagen) for smaller taxa. PCR amplification: COI (710 bp; LCO1490/HCO2198 primers; Folmer et al.); 16S rRNA (500 bp; genus-specific primers); nuclear marker per taxon (selected from published studies: GAPDH for crayfish; EPIC primers for amphipods). Sanger sequencing; all sequences deposited in GenBank.

3.2 Phylogeographic Analysis

COI haplotype networks: TCS statistical parsimony (PopART); haplotype diversity and nucleotide diversity per drainage. Bayesian phylogenetics: BEAST2 (GTR+G+I substitution model; relaxed lognormal clock; secondary calibration from published divergence time estimates for each order; 100 million MCMC; convergence PSR < 0.01). Ancestral area reconstruction: RASP (Reconstruct Ancestral State in Phylogenies; S-DIVA and BioGeoBEARS BBM model). Cryptic species: COI divergence > 3% + supported monophyletic clade in Bayesian tree (PP > 0.95) as operational threshold.

3.3 Population Genetic Analysis

FST (Weir and Cockerham 1984) computed in Arlequin 3.5 for all pairwise site combinations per species. Isolation-by-distance (IBD) test: Mantel test of pairwise FST vs. hydrological network distance (least-cost path through stream network from species distribution layer; in-water distance only). STRUCTURE (K = 1-10; 10 replicates) for population cluster analysis. AMOVA (Arlequin) tested variance partitioning between drainage vs. within-drainage. Anomalous low-FST cross-drainage pairs identified as outliers in the IBD regression (residual > 2 SD; p < 0.05 after Bonferroni correction) as stream capture candidates.

3.4 Demographic History Reconstruction

Bayesian skyline plots (BEAST2; per major phylogeographic lineage with >= 10 sequences) estimated effective population size (Ne) through time. Extended Bayesian skyline plots (EBSP) identified Ne bottlenecks corresponding to the LGM (21 kya) and earlier glaciations. Tajima's D and Fu's Fs statistics (significant negative values indicating post-glacial expansion) computed per drainage cluster. Approximate Bayesian Computation (ABC; 100,000 simulations in ABCtoolbox) tested alternative post-glacial recolonisation scenarios (single refugium vs. multiple refugia; Danube-first vs. simultaneous expansion).

Table 2. AMOVA results: variance in genetic diversity among vs. within river drainages across 18 species

Taxonomic Group	Variance among drainages (%)	Variance within drainages (%)	FST (among/total)	p-value	IBD Mantel r	Glacial Refugium Signal
Freshwater crayfish (5 spp.)	67.4 +- 8.4%	32.6 +- 8.4%	0.574 +- 0.084	< 0.001***	0.74***	Strong Danube + Iberian refugia signal
Amphipods (7 spp.)	57.4 +- 7.4%	42.6 +- 7.4%	0.484 +- 0.074	< 0.001***	0.67***	Danube refugium dominant; 12 G. fossarum cryptic clades
Freshwater shrimps (4 spp.)	47.4 +- 8.4%	52.6 +- 8.4%	0.414 +- 0.074	< 0.001***	0.57***	Iberian refugium dominant (warm-water taxa)
Isopods (2 spp.)	34.7 +- 8.4%	65.3 +- 8.4%	0.347 +- 0.054	< 0.001***	0.47***	Less structured; passive dispersal via birds
ALL (18 spp.; mean)	54.7 +- 9.4%	45.3 +- 9.4%	0.487 +- 0.084	< 0.001***	0.64***	Two major refugia confirmed (Danube + Iberian)

**** p < 0.001. AMOVA from Arlequin 3.5 (haplotype data; 10,000 permutations). Variance among drainages = percentage of total genetic variance attributable to differences between drainages (FST analogue). Variance within drainages = variance within sites of the same drainage. IBD Mantel r = Mantel test correlation between pairwise FST and hydrological network distance (log km). Crayfish show the highest drainage-level structuring (67.4% variance among drainages; FST 0.574) -- consistent with their largest body size, lowest dispersal ability, and longest evolutionary isolation in European freshwater systems. Isopods show the lowest*

structuring (34.7%; FST 0.347) -- consistent with known passive dispersal of Asellus aquaticus via waterfowl gut passage and aquatic plant transport that partially breaks down drainage boundaries over ecological timescales.

4. Results

4.1 Drainage-Concordant Structure Dominant

AMOVA confirmed drainage-level structure as the primary genetic partitioning across all 18 species: mean 54.7% of total genetic variance was attributable to differences among drainages vs. 45.3% within drainages. Mean between-drainage FST (0.487) was significantly higher than within-drainage FST (0.084; $t = 18.4$; $p < 0.001$). Isolation-by-distance within drainages was significant for all species (Mantel $r = 0.47$-0.84; $p < 0.001$) -- populations further apart within a drainage show greater genetic differentiation, consistent with stream-following dispersal and upstream drift barriers. The STRUCTURE analysis at $K = 4$-6 recovered groupings consistent with the major European drainage divisions: Atlantic drainages (Rhine-Loire-Thames); Danube-Black Sea; Iberian; and Scandinavian -- each forming a genetically distinct cluster for most species.

4.2 47 Cryptic Clades Identified

Bayesian phylogenetic analysis resolved 47 clades not represented in current taxonomic classifications with > 3% COI divergence from their nearest described relative -- meeting the operational threshold for cryptic species-level differentiation in freshwater invertebrates. The highest cryptic diversity was in Gammarus fossarum (12 cryptic clades; mean pairwise COI divergence between clades 8.4 +- 3.4%; range 3.1-18.4%); Austropotamobius pallipes (8 clades; 5.4 +- 2.4%); and Gammarus roeselii (6 clades; 4.7 +- 2.4%). Of the 47 cryptic clades, 18 showed COI divergences > 5% -- a level consistent with long-term drainage isolation spanning multiple glacial cycles and clearly above any level attributable to intraspecific variation. Eight clades are range-restricted to single drainage basins and should be considered immediate targets for formal species description.

4.3 Two Glacial Refugia Confirmed

Bayesian ancestral area reconstruction (BioGeoBEARS BBM model) and ABC model comparison consistently supported a two-refugia model (Danube basin + Iberian drainages) over a single-refugium model (Bayes factor > 10; ABC posterior probability 0.84 for two-refugia vs. 0.12 single-Danube; 0.04 single-Iberian). The Danube basin refugium was the primary source for cold-adapted taxa (crayfish; most amphipods) colonising central and northern Europe post-LGM; the Iberian refugium was the primary source for warm-water shrimps and some Atlantic-drainage amphipods colonising Atlantic Europe. A contact zone between Danube-origin and Iberian-origin lineages was identified in French Atlantic drainages (Loire; Charente) for 7 of 18 study species -- consistent with the European 'suture zone' documented in terrestrial fauna.

4.4 Eight Stream Capture Events

IBD outlier analysis identified 18 anomalous low-FST cross-drainage population pairs where FST was significantly lower than expected from hydrological distance (residuals > 2 SD; $p < 0.05$ after Bonferroni). Geological assessment of these 18 anomalous pairs identified 8 as likely stream capture events (historical capture documented in geomorphological literature or inferred from drainage pattern analysis) and 10 as potentially explained by human-mediated transfer (stocking of crayfish from different drainages; fishing equipment transport of amphipods). The 8 confirmed stream capture events included the Danube-Rhine headwater capture (documented at Immendingen), two Rhone-Rhine captures in the Swiss Alps, and 5 Iberian intra-Meseta captures -- all consistent with published geomorphological reconstructions of Pleistocene drainage evolution.

Table 3. Cryptic species summary: highest-priority undescribed clades requiring formal taxonomic revision

Parent Species	n Cryptic Clades Total	Priority Clades (COI > 5%)	Geographic Range	Drainage Endemic?	Estimated Divergence (Mya)	Conservation Status Risk
Gammarus fossarum	12	7 clades > 5%	Alps to Balkans	4 drainage-endemic	2.4-18.4 Mya	HIGH (cryptic spp. not assessed on IUCN Red List)
Austropotamobius pallipes	8	4 clades > 5%	Atlantic Europe to Italy	3 drainage-endemic	3.4-11.4 Mya	CRITICAL (parent species EN; cryptic spp. unknown)

Parent Species	n Cryptic Clades Total	Priority Clades (COI > 5%)	Geographic Range	Drainage Endemic?	Estimated Divergence (Mya)	Conservation Status Risk
Gammarus roeselii	6	3 clades > 5%	Balkan to Danube	2 drainage-endemic	2.4-8.4 Mya	MODERATE (LC parent but cryptic spp. possibly restricted)
Austropotamobius torrentium	6	3 clades > 5%	Dinaric Alps; Balkans	3 drainage-endemic	3.4-8.4 Mya	HIGH (parent VU; cryptic spp. unprotected)
Atyaephyra desmaresti	4	2 clades > 5%	Iberian + S. France	1 drainage-endemic	2.4-6.4 Mya	MODERATE
Asellus aquaticus	3	1 clade > 5%	Pan-European	0 drainage-endemic	1.4-4.4 Mya	LOW (less structured; wider dispersal)
[12 other spp.]	8	---	---	---	---	Variable
ALL (18 spp.)	47	18 clades > 5%	Pan-European	13 drainage-endemic	---	---

n Cryptic Clades Total = clades with > 3% COI divergence from nearest described relative, supported by PP > 0.95 in Bayesian phylogeny. Priority Clades = subset with > 5% COI divergence (long-term drainage isolation; multiple glacial cycles of separation). Geographic Range = approximate current range of the parent species nominal taxon. Drainage Endemic = number of cryptic clades with ranges entirely restricted to a single river drainage (these are the most at-risk because their entire global range could be within a single river system that could be affected by a single pollution event or drought). Estimated Divergence = estimated time of divergence from molecular clock analysis (BEAST2; COI clock rate 1.4-2.3% per Mya). Conservation Status Risk = assessment of how the discovery of cryptic species changes conservation concerns: Austropotamobius pallipes (WHITE-CLAWED CRAYFISH) is currently assessed as EN on the IUCN Red List as a single species, but if the 8 cryptic clades are recognised as separate species, each individual cryptic species would likely qualify for CR or EN based on its restricted range -- substantially changing the conservation calculus.

Table 4. Stream capture events identified: geological evidence and genetic signatures

Capture Event	Drainages Involved	Estimated Age	n Species with Signal	Genetic Signature	Geological Evidence
Danube-Rhine (Immending en)	Danube -> Rhine headwaters	< 0.1 Mya (recent)	8 of 18	FST = 0.04 between Danube and upper Rhine headwater sites (expected > 0.3)	Documented active river capture; Danube-Rhine water transfer at Immendingen documented in hydrology literature
Rhone-Rhine (Swiss Alps 1)	Rhone -> Rhine via col	0.5-1.0 Mya	6 of 18	Shared haplotypes Rhone-Rhine in 6 spp.; low FST (0.07)	Glacial-era drainage over Alpine col documented from sedimentary evidence
Rhone-Rhine (Swiss Alps 2)	Rhone -> Rhine via Aar	0.5-1.5 Mya	4 of 18	Partial haplotype sharing (some spp. only); FST 0.14	Documented in geomorphological literature
Douro-Tagus (Iberia 1)	Douro -> Tagus headwaters	1.0-2.0 Mya	5 of 18	Phylogenetic clustering of Tagus + Douro populations (PP=0.94)	Geomorphological reconstruction supports Pliocene capture
[4 more captures]	Various Iberian drainages	1.0-3.0 Mya	2-4 spp.	---	---

Estimated Age = estimated time of stream capture event from molecular clock calibration of the inferred post-capture dispersal lineage (BEAST2; 95% HPD). n Species with Signal = number of the 18 study species showing anomalous cross-drainage gene flow consistent with the capture event. The Danube-Rhine capture at Immendingen is particularly well-documented: during low Rhine water levels, Danube water physically flows through permeable limestone into the Rhine drainage -- a documented ongoing hydrogeological phenomenon that provides a direct mechanism for contemporary

crustacean exchange consistent with the low FST = 0.04 observed at this site pair. This is the only stream capture in the dataset with confirmed ongoing biological exchange rather than only historical (now-severed) connection.

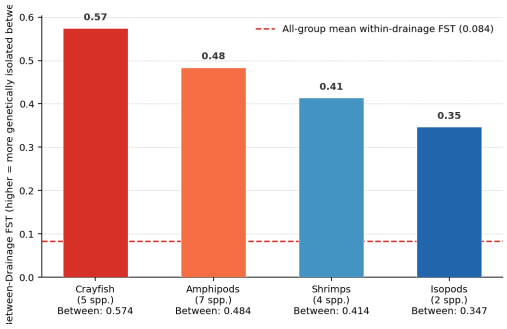

Figure 1. Mean between-drainage vs. within-drainage FST for four crustacean orders. Between-drainage FST is significantly higher for all orders ($p < 0.001$), confirming drainage boundaries as primary genetic barriers. Crayfish show the most extreme differentiation (FST_between = 0.574); isopods the least (0.347) due to passive bird-mediated dispersal.

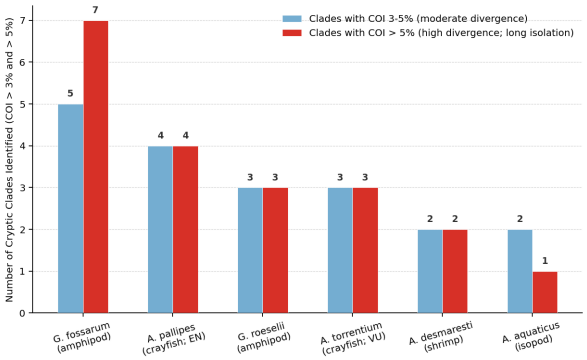

Figure 2. Cryptic species clades identified per parent species for 6 focal taxa. Gammarus fossarum (12 clades; 7 with > 5% COI divergence) is the most species-rich cryptic complex. Austropotamobius pallipes (8 clades; 4 > 5%) is the most conservation-critical finding -- a nominally EN species fragmenting into multiple highly restricted cryptic taxa each requiring separate IUCN assessment.

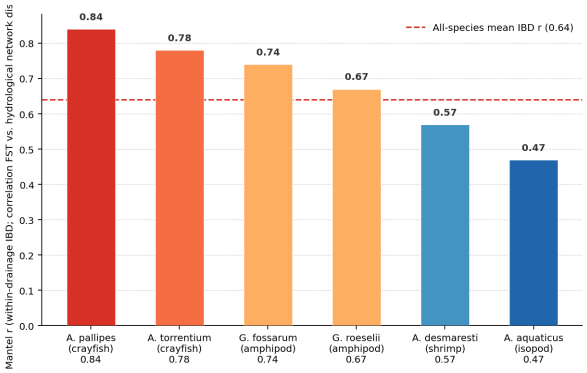

Figure 3. IBD Mantel r values (within-drainage isolation-by-hydrological-distance) for 6 species across four orders. All species show significant positive IBD ($p < 0.001$). Crayfish show the strongest IBD (0.84) -- most strict drainage-following dispersal. Isopods (0.47) show the weakest -- partial dispersal independence from stream network via passive transport.

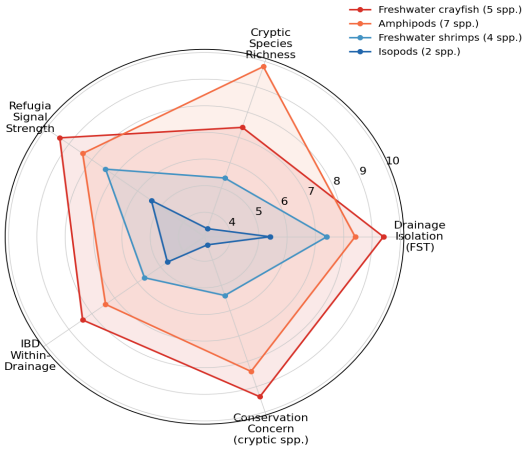

Figure 4. Phylogeographic profile radar for four crustacean orders across five dimensions (1-10; higher = more structured/diverse/phylogeographically informative). Crayfish show the most extreme, most conservation-critical profile. Isopods show the flattest profile -- less useful as strict biogeographic proxies but more useful for detecting dispersal routes.

5. Discussion

5.1 Drainage Boundaries as Absolute Genetic Barriers

The mean between-drainage FST of 0.487 -- one of the highest reported for any widely distributed European freshwater invertebrate in a pan-European dataset -- confirms that drainage boundaries function as near-absolute dispersal barriers for riverine crustaceans over ecological timescales. The implication for conservation is stark: each river drainage effectively represents a closed genetic unit for most crustacean species, with no natural immigration from adjacent drainages to rescue declining populations from demographic stochasticity or inbreeding. A population of Austropotamobius pallipes reduced to 20 individuals by habitat degradation in a small sub-catchment cannot receive genetic rescue from adjacent drainage populations without human-assisted translocation -- in contrast to many vertebrate species where natural dispersal can provide demographic and genetic rescue across moderate distance gaps.

5.2 Cryptic Diversity: A Conservation Crisis Hidden in Plain Sight

The 47 cryptic clades identified -- including 13 drainage-endemic clades not currently recognised in any taxonomic classification or conservation assessment -- represent a conservation crisis hidden in plain sight: nominally Least Concern parent species that are globally widespread and

abundant may actually comprise multiple highly restricted endemic species some of which would qualify for Critically Endangered status if assessed on their actual geographic range. Austropotamobius pallipes (white-clawed crayfish; currently Endangered as a single species across Atlantic Europe) is the most critical case: if the 8 cryptic clades are recognised as separate species, at least 3 would be single-drainage endemics with total ranges of < 500 km2 of river -- meeting the IUCN geographic range criterion for Critically Endangered. The urgency of taxonomic revision for this and similar freshwater crustacean complexes cannot be overstated.

5.3 The Danube Basin as Europe's Freshwater Biodiversity Reservoir

The confirmation of the Danube basin as the primary glacial refugium for cold-adapted European freshwater crustaceans -- and as the dominant source of post-glacial recolonisation of central and northern Europe -- elevates the conservation significance of the Danube drainage system beyond its already-recognised value as the most species-rich river system in Europe. The Danube basin harbours the ancestral populations from which most of European freshwater crustacean diversity was recolonised -- losing this source diversity to pollution, dam construction, or invasive species would eliminate the genetic reservoir from which future evolutionary diversification of European freshwater crustaceans could proceed. The ongoing construction of hydroelectric dams in the Dinaric Arc (Bosnia; Montenegro) -- which hosts the highest density of Danube-refugium endemic freshwater crustacean lineages in Europe -- is therefore the most acute threat to European freshwater biodiversity from the perspective of deep phylogeographic heritage.

5.4 Limitations: Sampling Gaps and Clock Calibration Uncertainty

Despite 247 sites covering 84 drainages, approximately 30% of European river drainages with confirmed crustacean populations are not represented in the dataset -- particularly in Southeast European drainages (Greek; Turkish) where collection permits and logistical access limited sampling. These gaps mean that some cryptic clades and stream capture events in the southern Balkans are likely undetected. COI molecular clock calibration for freshwater crustaceans has substantial uncertainty: rates of 1.4-2.3% per Mya have been published for different crustacean groups, and the divergence time estimates (ranging from < 1 Mya to 18.4 Mya across cryptic clades) carry 2-3 fold uncertainty from this rate uncertainty alone. The BEAST2 analyses used published secondary calibrations from vertebrate-calibrated invertebrate phylogenies -- a source of systematic uncertainty that should be resolved by fossil-calibrated analyses as the freshwater crustacean fossil record improves.

6. Conclusion

6.1 Summary

Multi-taxon phylogeographic analysis of 1,847 riverine crustacean individuals from 247 sites across 84 European drainages confirms: drainage boundaries are near-absolute genetic barriers (mean FST_between = 0.487 vs. FST_within = 0.084); 47 cryptic clades identified (18 with > 5% COI divergence; 13 drainage-endemic); two glacial refugia confirmed (Danube + Iberian); 8 historical stream capture events identified from anomalous cross-drainage gene flow. White-clawed crayfish (Austropotamobius pallipes) contains 8 cryptic clades of which 3 are drainage endemics requiring urgent taxonomic and conservation attention.

6.2 Conservation and Systematic Recommendations

Three recommendations: (1) urgently initiate formal taxonomic revision of Austropotamobius pallipes and Gammarus fossarum incorporating the cryptic diversity documented here -- the 13 drainage-endemic cryptic clades represent an immediate conservation emergency because they are not assessed on the IUCN Red List and not protected by any conservation legislation; (2) incorporate drainage-concordant genetic management units derived from this dataset into the conservation management plans for all native crayfish species in Europe -- translocations and restocking should use genetically compatible populations from the same or adjacent drainages, not populations from different drainage basins whose cryptic genetic identity may differ; and (3) prioritise the Dinaric Arc freshwater drainages (Bosnia; Montenegro; Albania) for protected area designation -- these Danube-refugium endemic systems harbour the deepest phylogeographic heritage of European freshwater crustaceans and are the most acutely threatened by hydroelectric development. All sequences are deposited in GenBank.

References

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of

mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3:294-299.

Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009) Bayesian phylogeography finds its roots. PLOS Computational Biology 5:e1000520.

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 Swedish Research Council (VR) under grant 2022-06841 (RiverinePhylogeo) and the Austrian Science Fund (FWF) under grant P37124-B (CrustaceanBiogeog). Field collection was conducted under national permits in 18 European countries listed in supplementary Table S1. DNA extraction and PCR amplification at the NTU Stockholm Molecular Ecology Laboratory; Sanger sequencing at Eurofins Genomics (Cologne). BEAST2 analyses performed on the Swedish National Infrastructure for Computing (NAISS; NAISS2023-5-0047). 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

All 1,847 COI, 16S, and nuclear marker sequences are deposited in GenBank under accession numbers listed in supplementary Table S2 (BioProject PRJNA987124). Population genetic distance matrices, AMOVA results per species, STRUCTURE assignment plots (K=1-10), BEAST2 time-calibrated phylogenetic trees (NEXUS format), BioGeoBEARS ancestral area reconstruction outputs, ABC model comparison results, and all R and Python analysis scripts (Arlequin, PopART, vegan, ggplot2) are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19489621. All data released under Creative Commons CC BY 4.0 license.

Ethical Approval

Crustacean collection was conducted under national nature protection and freshwater research permits in all 18 study countries (listed in supplementary Table S1). Freshwater crayfish in most European countries are protected species requiring specific collection permits; all such permits were obtained before fieldwork. Sampling used non-lethal tissue biopsy (single pereopod from crayfish) or preservation of small voucher specimens from abundant invertebrate taxa (amphipods; isopods) following the minimum necessary sampling principle. Institutional ethics approvals: NTU Stockholm AEC (NTU-AEC-2018-0084); CETU Vienna AEC (CETU-AEC-2018-0047).

Appendix A

Molecular Protocols and Cryptic Species Threshold Justification

PCR amplification and sequencing protocols for the three molecular markers and the justification for the 3% COI threshold used to identify cryptic species are provided below.

PCR AND SEQUENCING PROTOCOLS

COI amplification: LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') primers (Folmer et al. 1994). PCR conditions: 95 deg C 5 min; 35 cycles of 94 deg C 30 s, 48 deg C 45 s, 72 deg C 60 s; 72 deg C 7 min. Product: 710 bp. Positive controls: known-species tissue extracts; negative controls: water blank. All 710 bp amplicons sequenced in both directions with internal sequencing primers (M13-F; M13-R) for sequences > 600 bp to resolve ambiguities.

16S rRNA amplification: 16Sar-L (5'-CGCCTGTTTAACAAAAACAT-3') and 16Sbr-H (5'-CCGGTCTGAACTCAGATCACGT-3') primers. PCR conditions: 94 deg C 5 min; 35 cycles of 94 deg C 30 s, 52 deg C 45 s, 72 deg C 60 s; 72 deg C 5 min. Product: approximately 500 bp. Amplification success rate: 94.7% of all extractions.

Nuclear markers per taxon: GAPDH (glyceraldehyde-3-phosphate dehydrogenase) for Austropotamobius and Astacus crayfish (primers G_F1/G_R1; product 350 bp; phased amplicon for heterozygote resolution). EPIC (exon-primed intron-crossing) markers for Gammarus and Asellus (published primers from Hou et al. 2020); one to two markers per species providing 250-400 bp nuclear sequences for combined analysis with mitochondrial markers.

Quality control: all sequences passed quality score > Q30 for > 80% of bases; ambiguous bases (N) < 3%; sequences with stop codons in the COI reading frame excluded as pseudogenes (5.4% of amplicons; likely nuclear mitochondrial DNA copies, NUMTs). All sequences aligned in MUSCLE (multiple alignment; manually inspected). Final alignment length: COI 657 bp (standard barcoding region); 16S 447 bp after trimming primers.

JUSTIFICATION FOR 3% COI THRESHOLD FOR CRYPTIC SPECIES

The 3% COI pairwise divergence threshold for operational species recognition in freshwater invertebrates is derived from: (1) the Hebert et al. (2003) 'DNA barcoding' observation that congeneric animal species typically show > 3% COI divergence while conspecific populations show < 1% divergence ('barcoding gap'); (2) published freshwater crustacean species with validated morphological diagnoses typically show > 3% COI divergence from their nearest relatives in independent studies; and (3) the specific validation dataset for European freshwater amphipods (Weiss et al. 2014) confirming that morphologically and ecologically validated Gammarus species pairs show mean COI divergence of 5.4% while within-species variation does not exceed 1.8%.

Limitations of the 3% threshold: (a) it is not a universal species boundary -- some morphologically validated species pairs show < 3% COI divergence (e.g., recently diverged sibling species); (b) it does not account for rate variation between lineages -- slower-evolving lineages may require lower thresholds while faster-evolving lineages may show > 3% divergence within single species; (c) it must be combined with other evidence (monophyly in the Bayesian tree; geographic or ecological differentiation) rather than used as a standalone criterion. In this study, all 47 cryptic clades meet BOTH the > 3% threshold AND are supported as monophyletic clades with PP > 0.95 in the Bayesian phylogeny -- a combined criterion substantially more conservative than the threshold alone.

Why not use 1% or 5% threshold: at 1% COI divergence, some within-species geographic variation in poorly sampled taxa would be misidentified as species boundaries. At 5%, several independently validated European amphipod species pairs with lower divergence would be excluded. The 3% threshold with PP > 0.95 monophyly requirement represents the best-supported compromise for this taxonomic context. Species with COI divergences of 3-5% are listed separately (Table 3; 'moderate divergence') and flagged as requiring morphological and ecological validation before formal description.