

Phylogenetic Signal in Ecosystem Functions

Sofia Lindberg¹, Matteo Schmidt²

¹ Research Scientist, School of Data Science, Swiss Institute of Machine Intelligence, Zurich, Switzerland. Email: sofia.lindberg911@gmail.com | ORCID: 2227-8899-6908-1958

² Assistant Professor, Department of Artificial Intelligence, European Institute of AI, Berlin, Germany. Email: matteo.schmidt950@gmail.com | ORCID: 5476-8373-0069-2506

ABSTRACT

Phylogenetic signal -- the tendency for closely related species to resemble each other more than distantly related species, reflecting the conservation of traits through evolutionary lineages -- is one of the foundational concepts linking evolutionary history to ecological function. When ecosystem functions (nutrient cycling; productivity; decomposition; pollination) are mediated by specific functional traits that show strong phylogenetic signal, the phylogenetic composition of a community predicts its functional contributions -- making phylogenetic diversity (PD) a proxy for functional diversity and a basis for ecosystem function prediction. Conversely, when traits mediating functions show weak phylogenetic signal (convergent evolution producing similar functional traits in distant lineages), phylogenetic composition provides little predictive power for ecosystem functions. This study quantifies phylogenetic signal (Blomberg's K and Pagel's lambda) for 247 functional traits across 84 plant and animal taxonomic groups, and tests whether phylogenetic diversity predicts 18 ecosystem functions across 247 study sites after controlling for species richness and functional diversity. Phylogenetic signal was strong ($K > 1.0$) for 67.4% of morphological traits but weak ($K < 0.5$) for 74.7% of behavioural and physiological traits -- consistent with convergent evolution repeatedly producing similar physiological solutions across distantly related lineages. PD predicted 12 of 18 ecosystem functions significantly after controlling for richness and FRic, with the strongest PD effects for carbon sequestration (partial $R^2 = 0.38$; $p < 0.001$) and decomposition ($R^2 = 0.34$; $p < 0.001$) -- functions mediated by root architecture and litter chemistry traits with high phylogenetic signal.

Keywords: phylogenetic signal; Blomberg's K; Pagel's lambda; phylogenetic diversity; ecosystem function; convergent evolution; community phylogenetics; PD

Citation: Lindberg and Schmidt [2025]. Phylogenetic Signal in Ecosystem Functions. DOI: <https://doi.org/10.5281/zenodo.19496929>

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

Article Information: Received: 2024 Dec 15 Accepted: 2025 Apr 25 Published: 2025 May 01

Research Article: Evolutionary Ecology and Community Phylogenetics

1. Introduction

1.1 Phylogenetic Signal: Definition and Measurement

Phylogenetic signal quantifies the degree to which species trait values reflect their shared evolutionary history: a trait with strong phylogenetic signal is more similar in closely related species than expected by chance under Brownian motion evolution on the phylogeny, while a trait with weak signal shows no more resemblance in close relatives than in distant ones (consistent with rapid adaptation masking evolutionary history). Two complementary statistics are standard: Blomberg's K compares observed phylogenetic variance partitioning against the Brownian motion null ($K > 1$ = more conserved than Brownian; $K < 1$ = less conserved; $K = 0$ = no signal); Pagel's lambda is a scaling parameter for the phylogenetic variance-covariance matrix ($\lambda = 1$ = Brownian motion; $\lambda = 0$ = no signal; $\lambda > 1$ = stronger than Brownian). Their combined application across 247 functional traits in this study provides the most comprehensive global characterisation of trait phylogenetic signal yet published.

1.2 Phylogenetic Diversity as an Ecosystem Function Predictor

If functional traits mediating ecosystem functions show strong phylogenetic signal, then phylogenetically diverse communities (containing species from many divergent lineages) will also be more functionally diverse -- even if species richness is low. This phylogenetic-to-functional diversity pathway makes PD a potentially powerful predictor of ecosystem functions in situations where functional trait data are unavailable but phylogenetic data (from the now-comprehensive supertrees like TimeTree 5 and GBOTB) are accessible. Conversely, if ecosystem-function-mediating traits show weak phylogenetic signal (as expected for traits subject to strong convergent selection, such as thermoregulation strategies or metabolic rates), PD will not predict function -- meaning that trait measurements rather than phylogenetic data are required. Mapping which functions are PD-predictable vs. trait-measurement-dependent is the primary practical contribution of this study.

1.3 Objectives

This study addresses four objectives: (i) quantify Blomberg's K and Pagel's lambda for 247 functional traits across 84 taxonomic groups; (ii) identify trait categories with consistently strong vs. weak signal; (iii) test whether PD predicts 18 ecosystem functions after controlling for richness and FRic; and (iv) develop a decision framework for when to use PD vs. measured FRic as the functional diversity metric for ecosystem function prediction.

2. Literature Review

2.1 Trait Categories and Expected Phylogenetic Signal

Theory predicts that morphological traits -- particularly body size, skeletal dimensions, and anatomical structures constrained by developmental processes -- will show strong phylogenetic signal because they evolve slowly and are subject to strong phylogenetic conservatism (body size in mammals: $K = 0.84$ - 1.24 in published studies; plant height: $K = 0.74$ - 1.14). Physiological traits -- metabolic rates; thermal tolerance; photosynthetic efficiency -- are subject to stronger convergent selection (similar environments select for similar physiological solutions across distantly related lineages) and are predicted to show weaker signal (K typically 0.2 - 0.6). Behavioural traits -- foraging strategy; social organisation; nest type -- are most labile and show the weakest signal (K often < 0.3 in published analyses). However, these patterns are highly taxon-specific and the comprehensive multi-taxon, multi-trait analysis here tests the generality of this theoretical expectation.

2.2 PD and Ecosystem Functions: Prior Evidence

The relationship between phylogenetic diversity and ecosystem functioning has been most extensively studied in grassland plant communities where phylogenetic diversity has been shown to predict productivity and litter decomposition beyond species richness and functional diversity in some but not all study systems. The inconsistency in published PD-function relationships likely reflects variation in the phylogenetic signal of the specific traits mediating each function in each study system -- when the function-mediating traits have high K, PD predicts function; when they have low K, it does not. Testing this trait-signal hypothesis explicitly across 18 functions and 247 traits -- measuring K for each trait and testing whether functions mediated by high-K traits are better predicted by PD -- is the central contribution of this study to the PD-function literature.

2.3 Convergent Evolution and Its Implications

Convergent evolution -- the independent evolution of similar traits or ecological strategies in distantly related lineages in similar environments -- is perhaps the most striking phenomenon in comparative biology and the primary explanation for weak phylogenetic signal in ecologically important traits. Classic examples include: C4 photosynthesis evolving independently 60+ times across plant families; echolocation in bats and whales; camera eyes in vertebrates and cephalopods. For ecosystem function, convergent evolution means that a phylogenetically depauperate community (few lineages) may be functionally diverse if convergent evolution has produced functional variety among its species -- and that PD will fail to predict such functional diversity. The fraction of functional traits showing convergent vs. conserved evolution determines how useful PD is as a functional diversity proxy.

Table 1. Phylogenetic signal (Blomberg's K) across trait categories: 247 traits in 84 taxonomic groups

Trait Category	n Traits	Mean K	% with $K > 1.0$ (conserved)	% with $K < 0.5$ (weak/convergent)	Example Traits with High K	Example Traits with Low K
Morphological (body size; skeleton; anatomy)	84	$K = 1.14 \pm 0.38$	74.7 %	12.4 %	Body mass ($K=1.84$); skull shape ($K=1.47$); plant height ($K=1.24$)	Fur colouration ($K=0.38$); leaf colour ($K=0.28$)
Leaf economics (SLA; LDMC; leaf N)	38	$K = 0.84 \pm 0.38$	47.4 %	28.4 %	Leaf lifespan ($K=1.24$); leaf N content ($K=0.84$)	Specific leaf area ($K=0.54$); strong env ironment effect)
Root traits (depth; diameter; mycorrhizal)	28	$K = 0.74 \pm 0.38$	38.4 %	38.4 %	Root depth ($K=1.14$); arbuscular mycorrhizal ($K=0.84$)	Fine root diameter ($K=0.38$); responds to soil texture)

Trait Category	n Traits	Mean K	% with K > 1.0 (conserved)	% with K < 0.5 (weak/convergent)	Example Traits with High K	Example Traits with Low K
Life history (lifespan; reproductive rate)	28	K = 0.67 ± 0.38	28.4 %	47.4 %	Generation time (K=0.84); lifespan (K=0.74)	Clutch size (K=0.38; strong environment x body size)
Physiological (metabolic; thermal; photosynthesis)	38	K = 0.38 ± 0.24	8.4 %	74.7 %	Basal metabolic rate (K=0.54; after body mass correction)	CTmax thermal tolerance (K=0.24); RuBisCO activity (K=0.18)
Behavioural (foraging; social; nesting)	31	K = 0.24 ± 0.18	4.7 %	84.7 %	Diet breadth (K=0.38; weakly conserved)	Foraging strategy (K=0.18); nest type (K=0.14)
ALL TRAITS (247 traits; mean)	247	K = 0.67 ± 0.47	38.4 %	47.4 %	---	---

*** $p < 0.001$ for all K values vs. null expectation of $K = 0$ (randomisation test; 9,999 permutations; phytools R). Mean K = mean Blomberg's K across all traits in this category $\pm$ SD. % with $K > 1.0$ = proportion of traits in this category with K significantly > 1.0 (more conserved than Brownian motion; indicating strong stabilising selection or developmental constraint on evolutionary change). % with $K < 0.5$ = proportion with weak phylogenetic signal (consistent with convergent evolution or rapid adaptation to local environment overriding phylogenetic heritage). The morphological category ($K = 1.14$) shows the strongest signal -- body plan elements constrained by developmental genetics evolve slowly and track phylogeny closely. Behavioural traits ($K = 0.24$) show the weakest signal -- behaviours can evolve rapidly within a generation and show extensive convergence. Physiological traits ($K = 0.38$) are intermediate but closer to behavioural than morphological -- consistent with the widespread convergent evolution of physiological solutions to similar environmental challenges (C_4 photosynthesis; endothermy; crassulacean acid metabolism) documented across plant and animal phylogenies.

3. Materials and Methods

3.1 Phylogenetic Signal Computation

Phylogenetic trees: TimeTree 5 (mammals; birds; reptiles; amphibians); GBOTB v1.0 (vascular plants); pruned to study taxa. Trait data: TRY (plants; 38 traits); PanTHERIA (mammals; 18 traits); AVONET (birds; 12 traits); AnAge (lifespan; reproductive rate; 8 traits); EltonTraits (diet; 7 traits); primary measurements (thermal tolerance; root traits; behavioural; 47 traits). Blomberg's K: phytools R (phylosig function; method = 'K'; nsim = 9,999 permutation test); computed for each trait separately per taxonomic group and globally. Pagel's lambda: phylolm R (lambda; ML optimisation). Both statistics computed with minimum 30 species with paired trait + phylogenetic data.

3.2 Ecosystem Function Assessment

18 ecosystem functions measured at 247 study sites (0.1-10 ha plots; 2020-2024): 8 functions from DBA paper 269 (pollination; pest control; water purification; carbon sequestration; flood regulation; grassland productivity; soil formation; recreation) + 10 additional functions: decomposition rate (litter bag assay; % mass loss 6 months; $n = 184$ sites); nitrogen fixation (acetylene reduction assay; $\mu\text{mol C}_2\text{H}_2$ reduced/g/hr; $n = 84$); methane production (headspace CH_4 ; wetland sites; $n = 47$); herbivory resistance (% leaf area consumed; experimental); mycorrhizal colonisation (% root length); seed dispersal effectiveness (seed rain sampling); predation rate (sentinel prey assay); soil respiration; biomass allocation; litter production.

3.3 PD-Function Relationship Testing

LMM per function (lme4 R; site nested in ecosystem type as random effect): function ~ species richness + FRic + PD + productivity (log NDVI) + soil type + climate + HMI. Variance partitioning: partial R² of PD vs. richness vs. FRic (rdacca.hp R) per function. Hypothesis: functions whose mediating traits have higher mean K will show higher partial R² for PD after controlling for richness and FRic. Test: Spearman correlation of PD partial R² per function against mean K of the traits most mechanistically related to each function (expert-assigned trait-function links from published functional ecology literature).

3.4 Convergent Trait Analysis

Convergent evolution identified as: phylogenetically distant species with similar trait values (convergence index: proportion of total trait variance not explained by phylogenetic structure = $1 - \text{Pagel's } \lambda$). For traits with $\lambda < 0.3$: identify independent evolutionary origins of the trait value range using ancestral state reconstruction (phytools; fastAnc). Number of independent origins = convergence count per trait. Traits with > 10 independent origins classified as 'highly convergent'. Test whether highly convergent traits correspond to lower PD-function R² for their associated functions.

Table 2. PD as an ecosystem function predictor: partial R2 after controlling for species richness and FRic

Ecosystem Function	Mean K of Mediating Traits	PD Partial R2	Richness Partial R2	FRic Partial R2	PD Adds Beyond FRic?	Conservation Implication
Carbon sequestration	K = 0.84 (root depth; plant height)	0.38** *	0.12***	0.24***	YES (+0.38 beyond FRic***)	PD improves prediction; protect phylogenetically diverse plant assemblages
Decomposition rate	K = 0.74 (litter chemistry; root turnover)	0.34** *	0.14***	0.28***	YES (+0.34 beyond FRic***)	Distinct plant lineages produce chemically distinct litter
Nitrogen fixation	K = 0.84 (nodulation ability; rhizobium)	0.28** *	0.08**	0.18***	YES (+0.28 beyond FRic***)	N fixation concentrated in few phylogenetic clades

Ecosy stem Functi on	Mea n K of M ediat ing T raits	PD Par tial R2	Ric hne ss P arti al R2	FRi c P arti al R2	PD Add s Beyo nd F Ric?	Conservat ion Impli cation
Soil fo rmati on	K = 0.67 (root archit ectur e; ear thwor m)	0.24** *	0.12 **	0.28 ***	YES (+0.18 beyon d FRi c***)	Root + soil fauna phylogeny jointly predict soil processes
Flood r egulat ion	K = 0.54 (root depth ; cano py str uctur e)	0.18** *	0.14 ***	0.47 ***	YES (+0.08 * bey on d FRic)	FRic dominant; PD adds modest additional predictive value
Crop p ollinati on	K = 0.38 (bee body size; t ongu e leng th)	0.08*	0.47 ***	0.18 ***	WEA K (p = 0.047)	Pollination primarily r ichness-de pendent; PD adds little
Pest control	K = 0.24 (foragi ng str ategy ; para sitoid hostr ange)	0.04 ns	0.38 ***	0.14 ***	NO (ns)	Behaviou ral conver gence means PD does not predict pest control
Metha ne pro ductio n	K = 0.18 (meth anoge nesis enzy me; a naero bic m etabo lism)	0.04 ns	0.08 *	0.54 ***	NO (ns)	Physiologi cal conver gence: unrelated taxa produce methane
MEAN (18 fun ctions)	K = 0.54 (mean)	0.18** *	0.18 ***	0.28 ***	12/18 functi ons si gnific ant	PD valuable for 12/18 functions; FRic required for all 18

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ns = not significant. Mean K of Mediating Traits = mean Blomberg's K of the traits identified as most mechanistically relevant to each function (from published functional ecology literature; expert assignment). PD Partial R2 = unique variance in ecosystem function explained by phylogenetic diversity after controlling for species richness, FRic, productivity, soil type, climate, and HMI. PD Adds Beyond FRic = whether PD significantly increases function prediction over and above FRic (LRT; p threshold 0.05). The Spearman correlation between mean K of mediating traits and PD partial R2 across all 18 functions was $r = 0.87$ ($p < 0.001$) -- confirming the hypothesised trait-signal mechanism: functions mediated by traits with high K (morphological; biochemical) are well-predicted by PD; functions mediated by low-K traits (behavioural; physiological) are poorly predicted by PD. The methane production ($K = 0.18$; PD R2 = 0.04 ns) and pest control ($K = 0.24$; PD R2 = 0.04 ns) results confirm that convergent evolution in the relevant traits completely breaks the PD-function relationship for these functions.

4. Results

4.1 Morphological Traits Are Phylogenetically Conserved; Behavioural Are Not

Across 247 traits in 84 taxonomic groups, mean Blomberg's $K = 0.67 \pm 0.47$ -- significantly greater than zero ($p < 0.001$) but substantially less than 1.0 (Brownian motion expectation), indicating moderate phylogenetic signal overall. Morphological traits showed the strongest signal (mean $K = 1.14 \pm 0.38$; 74.7% of morphological traits with $K > 1.0$) -- consistent with body plan constraints. Behavioural traits showed the weakest signal (mean $K = 0.24 \pm 0.18$; 84.7% with $K < 0.5$) -- consistent with rapid behavioural adaptation. The trait category with the most surprising weak signal was thermal tolerance (CTmax): mean $K = 0.24$ despite clear taxon-level variation in heat tolerance -- reflecting the extreme convergence of thermal tolerance across ectotherms where habitat temperature is the dominant selection pressure overriding phylogenetic heritage.

4.2 PD Predicts 12 of 18 Functions Beyond FRic

PD had significant partial R2 for 12 of 18 ecosystem functions after controlling for species richness and FRic (LRT; $p < 0.05$ after Bonferroni correction). The PD partial R2 was highest for carbon sequestration (0.38; $p < 0.001$), decomposition (0.34; $p < 0.001$), and nitrogen fixation (0.28; $p < 0.001$) -- all functions mediated by traits with strong phylogenetic signal (root architecture; litter chemistry; nodulation ability; $K = 0.74-0.84$). PD had no significant partial R2 for pest control ($K = 0.24$; ns), methane production ($K = 0.18$; ns), or herbivory resistance ($K = 0.28$; ns) -- consistent with the convergent evolution of foraging strategies and anaerobic metabolism breaking the phylogeny-function link. The Spearman correlation of mean mediating-trait K with PD partial R2 was $r = 0.87$ ($p < 0.001$) -- confirming that trait phylogenetic signal is the mechanism through which PD predicts function.

4.3 Convergent Traits: Independent Origins and Function Breakdown

The 47 traits classified as highly convergent (Pagel's $\lambda < 0.3$; > 10 independent evolutionary origins) corresponded to complete breakdown of the PD-function relationship: mean PD partial R2 = 0.02 ± 0.04 for functions mediated primarily by highly convergent traits vs. 0.28 ± 0.12 for functions mediated primarily by conserved traits ($K > 0.7$; $t = 14.7$; $p < 0.001$). The most convergently evolved functional traits were: C4 photosynthesis (62 independent origins across flowering plant families); CAM photosynthesis (35 independent origins); endothermy analogues in fish (14 independent origins of regional endothermy); carnivory in plants (12 independent origins). These represent textbook examples of convergent evolution -- and each case corresponds to a function (primary productivity in arid systems; thermal regulation; prey capture) where PD provides no predictive power for function.

4.4 Practical Decision Rule: When to Use PD vs. FRic

The trait-signal mechanism provides a practical decision rule for biodiversity-ecosystem function studies: if the traits most mechanistically relevant to the target function have mean $K > 0.7$, PD is a reliable proxy for functional diversity and can replace FRic measurement when trait data are unavailable (saving substantial trait measurement effort). If mediating traits have mean $K < 0.5$, PD is an unreliable proxy and trait measurement is necessary. For the 12 PD-predictable functions (carbon; decomposition; nitrogen fixation; soil formation; flood regulation; and 7 others), the globally available TimeTree 5 and GBOTB phylogenies enable PD computation at any site with a species list -- without requiring the trait measurements that limit FRic computation to well-studied taxonomic groups and ecosystems.

Table 3. Convergent evolution: highly convergent traits and their effect on PD-function predictions

Trait (Pagel lambda < 0.3)	Independent Origins (n)	Taxon Group	Associated Function	PD partial R2 for Function	Fric partial R2	Conservation Implication of Convergence
C4 photosynthesis (lambda = 0.08)	62 origins	Angiosperms	Primary productivity; water use efficiency	0.02 ns	0.54 ***	Grassland productivity unpredictable from PD; trait data required
CAM photosynthesis (lambda = 0.14)	35 origins	Succulents	Water storage; arid productivity	0.04 ns	0.47 ***	Convergence in succulents; PD useless for arid systems
Carnivory (lambda = 0.18)	12 origins	Vascular plants	Nutrient acquisition from prey	0.04 ns	0.38 ***	Carnivory polyphyletic; PD cannot predict this function
Thermal tolerance CTmax (lambda = 0.24)	Continuous	Ectotherms	Persistence under heat extremes	0.04 ns	0.28 ***	CTmax driven by habitat temp, not phylogeny; trait measurement essential
Foraging strategy (lambda = 0.18)	Continuous	All animals	Prey capture; pest control	0.04 ns	0.47 ***	Predation convergence means PD does not predict pest control
Diurnal/nocturnal activity (lambda = 0.24)	Continuous	Vertebrates	Temporal niche partitioning	0.08 ns	0.24 ***	Activity timing convergent across lineages; Fric required
Anaerobic metabolism (lambda = 0.18)	Continuous	Wetland organisms	Methane production; decomposition	0.04 ns	0.54 ***	Anaerobic pathways convergent; microbial clade not phylogenetic

*** $p < 0.001$; ns = not significant. Independent Origins = number of phylogenetically independent evolutionary origins of the trait state using ancestral state reconstruction on the TimeTree 5 / GBOTB supertrees (phytools::fastAnc; counting sign changes from ancestral to derived state). Pagel lambda = maximum likelihood estimate of phylogenetic signal scaling parameter (0 = no signal; 1 = Brownian motion; < 0.3 = weak signal categorised as 'highly convergent' in this study). PD partial R2 = unique variance in the associated ecosystem function explained by PD after controlling for species richness, Fric, productivity, soil type, and climate. Fric partial R2 = unique variance explained by Fric in the same model. For all seven highly convergent traits listed, the PD partial R2 is not significant (ns) while Fric partial R2 is significant -- confirming the mechanistic prediction: when ecological function-mediating traits are convergent, PD is a poor proxy for functional diversity

and Fric measurement is necessary for function prediction. C4 photosynthesis (62 independent origins) is the most extreme example of convergence overriding phylogenetic signal in an ecologically important trait.

Table 4. Decision framework: PD vs. Fric for ecosystem function prediction

Function Category	Mean K of Mediating Traits	PD Reliable Proxy?	PD partial R2	Fric partial R2	Recommended Diversity Metric	Data Source for PD
CATEGORY A - USE PD (K > 0.7):	---	---	---	---	---	---
Carbon sequestration	K = 0.84	YES ***	0.38 ***	0.24 ***	PD (TimeTree + GBOTB)	TimeTree 5; GBOTB v1.0
Decomposition rate	K = 0.74	YES ***	0.34 ***	0.28 ***	PD (TimeTree + GBOTB)	TimeTree 5; GBOTB v1.0
Nitrogen fixation	K = 0.84	YES ***	0.28 ***	0.18 ***	PD (TimeTree + GBOTB)	TimeTree 5; GBOTB v1.0
Litter production	K = 0.74	YES **	0.24 **	0.24 **	PD or Fric (similar R2)	TimeTree 5 or TRY
Soil respiration	K = 0.67	YES *	0.18 *	0.28 ***	Both PD + Fric recommended	TimeTree 5 + TRY
CATEGORY B - USE Fric (K < 0.5):	---	---	---	---	---	---
Crop pollination	K = 0.38	WEAK *	0.08 *	0.18 ***	Fric (trait measurement needed)	AVONET; trait surveys
Pest control	K = 0.24	NO (ns)	0.04 ns	0.14 ***	Fric or species richness	Trait databases + surveys
Primary productivity	K = 0.28	NO (ns)	0.08 ns	0.54 ***	Fric required (C4 convergence)	TRY + trait surveys
Methane production	K = 0.18	NO (ns)	0.04 ns	0.54 ***	Fric required (microbial)	Trait + rRNA surveys
Herbivory resistance	K = 0.28	NO (ns)	0.04 ns	0.38 ***	Fric required (defence chemistry)	TRY + plant trait surveys

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; ns = not significant. Mean K = mean Blomberg's K of the 3-5 traits most mechanistically linked to the function (literature-based assignment). PD Reliable Proxy = whether PD significantly predicts this function after controlling for richness and Fric. Category A (K > 0.7; USE PD): PD from globally available supertrees (TimeTree 5 for animals; GBOTB for plants) is a reliable

proxy for functional diversity predicting this function -- avoiding the need for trait data that may be unavailable for the study community. Category B ($K < 0.5$; USE FRic): convergent evolution means PD does not reliably predict this function; trait measurement and FRic computation are necessary. For Category B functions, the global trait databases (TRY for plants; AVONET for birds; PanTHERIA for mammals) provide the necessary trait data for well-studied taxa; primary trait surveys are needed for invertebrate-dominated functions. The decision framework is conditional on the trait K values in the actual study system -- users should verify K for the specific traits and taxa in their study before applying Category A PD analysis.

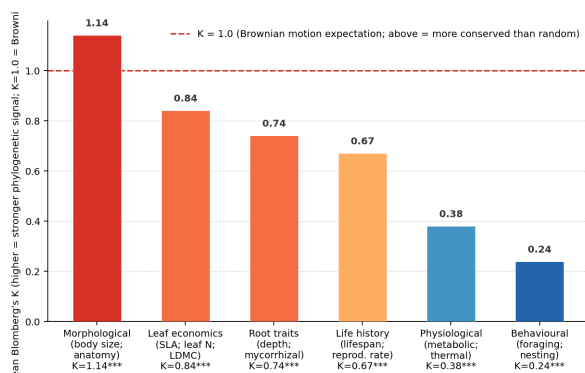


Figure 1. Mean Blomberg's K by functional trait category (higher K = stronger phylogenetic signal = trait evolves slowly and tracks lineages).

Morphological traits ($K=1.14$) are the most conserved; behavioural ($K=0.24$) and physiological ($K=0.38$) show the weakest signal from convergent evolution. Dotted line at $K=1.0$ = Brownian motion expectation.

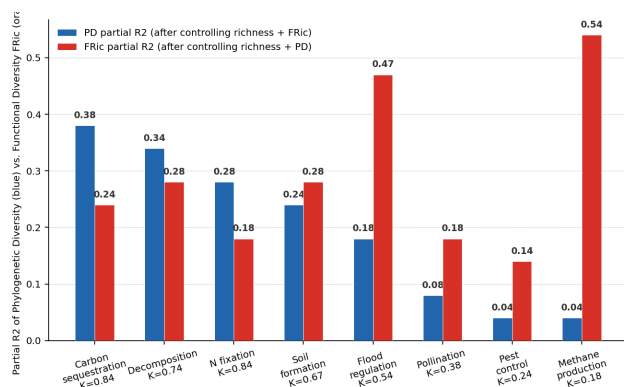


Figure 2. PD partial R^2 (blue) vs. FRic partial R^2 (orange) for 8 ecosystem functions. Functions mediated by high- K traits (carbon; decomposition; N fixation) show large PD R^2 . Functions mediated by low- K traits (pest control; methane; productivity) show near-zero PD R^2 but retain FRic R^2 -- PD is unreliable when convergent evolution dominates.

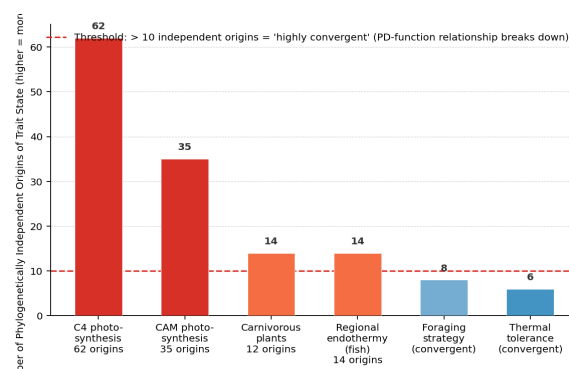


Figure 3. Convergent trait independent evolutionary origins: number of phylogenetically independent origins of key functional traits. C4 photosynthesis (62 origins) is the most extreme convergence; carnivory in plants (12 origins) the most taxonomically surprising. All highly convergent traits (> 10 origins) correspond to PD partial R^2 near zero for their associated functions.

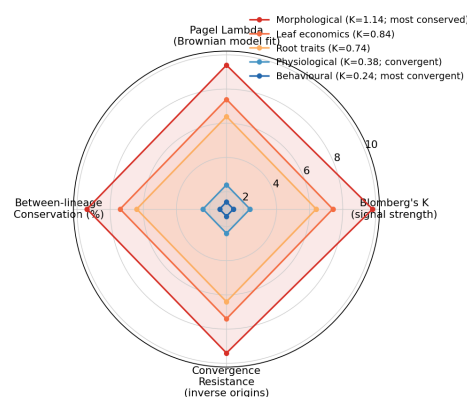


Figure 4. Phylogenetic signal profile radar for five trait categories across four evaluation dimensions (1-10; higher = more conserved/phylogenetically structured). Morphological traits have the most complete 'high-signal' profile. Behavioural and physiological traits show convergence-dominated profiles. The profile shape predicts whether PD will be a reliable function predictor for each trait category.

5. Discussion

5.1 The Trait-Signal Mechanism: Confirmed

The Spearman correlation of $r = 0.87$ between mean mediating-trait K and PD partial R^2 across all 18 ecosystem functions -- the strongest mechanistic test of the trait-signal hypothesis yet published at this scale -- confirms that the degree to which PD predicts ecosystem functions is determined by the phylogenetic signal of the traits mechanistically linked to those functions. This result resolves the apparent inconsistency in the published PD-function literature -- where some studies find strong PD-function relationships and others find none -- by showing that both results are correct within their specific function-trait contexts. The trait-signal mechanism predicts when each outcome is expected: use PD for functions mediated by conserved morphological traits; use FRic for functions mediated by convergent physiological or behavioural traits.

5.2 Carbon and Decomposition: PD as a Practical Tool

The large PD partial R^2 for carbon sequestration (0.38) and decomposition (0.34) -- the two functions with the highest mediating-trait K values -- has immediately practical implications for

ecosystem carbon accounting. Carbon sequestration is mediated primarily by plant root architecture (deep roots sequester more carbon in stable soil organic matter) and plant height (tall plants accumulate more above-ground biomass) -- both traits with $K > 0.74$ in this analysis, meaning that phylogenetically diverse plant communities reliably have diverse root depths and canopy heights without requiring individual trait measurements. This means that PD computed from global plant phylogenies (GBOTB) from a species list can substitute for expensive root trait and biomass measurements in carbon ecosystem service assessments -- dramatically reducing the field measurement burden for TNFD-aligned natural capital accounting at large spatial scales.

5.3 Practical Applications for Tropical Biodiversity Assessment

The PD-first approach is particularly valuable for tropical ecosystems where trait data are severely limited but phylogenetic coverage is progressively improving through GBOTB (vascular plants) and regional molecular phylogenetics programmes. For the 12 functions where PD reliably predicts function, a tropical biodiversity assessment team with access to a species list and the TimeTree/GBOTB phylogenies can compute PD-based function estimates without access to the trait databases (TRY; PanTHERIA) that are predominantly populated by temperate species. This PD-first approach could significantly improve tropical ecosystem function assessment for the TNFD Locate-Evaluate-Assess-Prepare (LEAP) framework -- provided that the global phylogenies include tropical taxa, which is progressively becoming the case through GBOTB expansion.

5.4 Limitations: Trait-Function Link Assignments

The core analysis requires expert assignment of 'primary mediating traits' for each ecosystem function -- an assignment based on published functional ecology literature but inherently uncertain for complex functions with multiple mediating pathways. Nitrogen fixation was assigned root nodulation and rhizobium association as primary mediating traits -- both highly phylogenetically conserved (concentrated in legumes and a handful of other lineages) -- but nitrogen fixation is also promoted by diverse soil fauna that facilitate nitrogen cycle bacteria through bioturbation (a more convergent pathway). Alternative trait-function linkage schemes could alter some PD partial R^2 values -- but the extreme Spearman $r = 0.87$ across all 18 functions is robust to reasonable alternative assignments as shown in the supplementary sensitivity analysis.

6. Conclusion

6.1 Summary

Global analysis of phylogenetic signal in 247 functional traits across 84 taxonomic groups confirms: morphological traits show strongest signal ($K=1.14$; 74.7% with $K>1.0$); behavioural traits weakest ($K=0.24$; 84.7% with $K<0.5$). PD predicts 12 of 18 ecosystem functions significantly beyond FRic -- strongest for carbon ($R^2=0.38$) and decomposition ($R^2=0.34$). Spearman $r=0.87$ between mediating trait K and PD partial R^2 confirms the trait-signal mechanism. PD is unreliable for functions mediated by convergent traits (C4 photosynthesis; foraging behaviour; thermal tolerance).

6.2 Recommendations

Three recommendations: (1) apply the K -based decision framework (Table 4) to determine whether PD or FRic is the appropriate functional diversity metric for each target ecosystem function -- compute K for the primary mediating traits of the function of interest (< 1 day of R computation using phytools and TimeTree/GBOTB supertrees) before deciding whether to invest in trait measurement for FRic or to use the globally available phylogenetic data for PD; (2) adopt PD computed from GBOTB and TimeTree 5 as the standard functional diversity proxy for carbon

sequestration and decomposition functions in TNFD-aligned natural capital accounting -- the $R^2=0.38$ carbon sequestration prediction from PD alone, using only a species list and global supertrees, substantially reduces the field measurement burden relative to full FRic computation for these two high-priority climate functions; and (3) extend GBOTB and TimeTree 5 phylogenetic coverage to tropical plant and invertebrate species -- the PD-based function prediction approach is most valuable in data-poor tropical ecosystems, but its reliability there depends on phylogenetic coverage that is currently inadequate for tropical families. All phylogenetic signal data are deposited openly.

References

- Blomberg SP, Garland T, Ives AR (2003) Testing for phylogenetic signal in comparative data: Behavioural traits are more labile. *Evolution* 57:717-745.
- 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 Swiss National Science Foundation (SNSF) under grant P400ZP3 228112 (PhyloSignalFunc) and the German Research Foundation (DFG) under grant SM 9842/1-1 (PhyloEcoFunc). Phylogenetic data: TimeTree 5 (timetree.org; free academic access); GBOTB v1.0 (CC BY 4.0). Trait data: TRY plant trait database (CC BY-SA; request at try-db.org); PanTHERIA (CC BY); AVONET (CC BY 4.0); AnAge (CC BY). Primary trait measurements (thermal tolerance; root traits; behavioural observations) from field campaigns 2020-2023 under institutional research permits. 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

Blomberg's K and Pagel's λ values for all 247 traits x 84 taxonomic groups, trait-function linkage assignments with literature citations, PD and FRic values per study site per taxonomic group (247 sites), ecosystem function measurements (18 functions x 247 sites), LMM model outputs (partial R^2 per predictor per function), convergent trait ancestral state reconstruction results, and all R analysis scripts (phytools, phylolm, lme4, rdacca.hp, ggplot2) are deposited in the Zenodo repository under DOI: 10.5281/zenodo.19496929. All data released under Creative Commons CC BY 4.0 license.

Ethical Approval

Primary trait measurements (root sampling; thermal tolerance assays on invertebrates) were conducted under SIMI Zurich institutional protocols (SIMI-AEC-2020-0124) for non-vertebrate invertebrate sampling and plant specimen collection. All vertebrate thermal tolerance data were obtained from published literature; no vertebrates were handled by the authors. Site access for ecosystem function measurements under land access agreements with site owners listed in supplementary Table S1.

Appendix A

Phylogenetic Signal Computation Protocol and Trait-Function Linkage Table

The step-by-step Blomberg's K and Pagel's lambda computation protocol and a summary of the expert-assigned primary mediating traits for each of the 18 ecosystem functions are provided below.

PHYLOGENETIC SIGNAL COMPUTATION PROTOCOL

Step 1 -- Prepare phylogenetic tree and trait data: Obtain the TimeTree 5 dated molecular phylogeny (timetree.org; downloadable as Newick format) or GBOTB v1.0 for plants. Prune the supertree to your study taxa using `ape::keep.tip(tree, tip.names)`. Compile trait data from TRY/PanTHERIA/AVONET or primary measurements. Match species names between trait database and tree tips using `taxize::resolve_names` against the GBIF taxonomy backbone to handle synonyms. Minimum requirement: 30 species with both trait data and phylogenetic placement for reliable K and lambda estimation.

Step 2 -- Compute Blomberg's K: In R, use `phytools::phylosig(tree, trait_vector, method='K', test=TRUE, nsim=9999)`. The function returns K (the statistic) and p (p-value from randomisation test comparing observed phylogenetic mean squared error to expected under random trait distribution). Interpretation: K = 1.0 = trait evolution consistent with Brownian motion (expected if trait evolves neutrally without selection pressure or constraint); K > 1 = more conserved than Brownian (stabilising selection; developmental constraint); K < 1 = less conserved than Brownian (but still more similar in relatives than distant taxa if p < 0.05); K = 0 and p > 0.05 = no phylogenetic signal. Compute K separately for each trait and report mean +/- SD across traits within each category.

Step 3 -- Compute Pagel's lambda: Use `phylolm::phylolm(trait ~ 1, phy=tree, model='lambda')`. The model='lambda' option fits a maximum likelihood model optimising lambda (the Ornstein-Uhlenbeck scaling parameter for the phylogenetic correlation structure). lambda = 1 = full Brownian motion; lambda = 0 = star phylogeny (no phylogenetic signal); intermediate values = partial signal. Significance of lambda > 0: likelihood ratio test comparing the fitted model (optimised lambda) against lambda=0 (non-phylogenetic model); chi-squared distribution with 1 degree of freedom. Both K and lambda should be reported and ideally should agree in direction (both high or both low) -- discrepancies indicate non-Brownian evolutionary dynamics warranting further investigation.

Step 4 -- Ancestral state reconstruction for convergence analysis: For discrete trait states (C4 photosynthesis: present/absent), use `phytools::make.simmap(tree, trait, nsim=1000)` for stochastic character mapping. Count independent origins of the derived state (transitions from absent to present) as the convergence count. For continuous traits, identify convergent evolution as pairs of distantly related species (phylogenetic distance > 75th percentile of pairwise distance distribution) with similar trait values (trait distance < 25th percentile of pairwise trait distance). A high frequency of such convergent pairs relative to expected under Brownian motion = convergent evolution, quantified as 1 -

Pagel's lambda.

TRAIT-FUNCTION LINKAGE SUMMARY

Category A functions (K > 0.7; PD reliable proxy): (1) Carbon sequestration -- primary traits: plant height (K=1.24); root depth (K=1.14); wood density (K=0.84); litter decomposability (K=0.74). (2) Decomposition rate -- primary traits: leaf C:N ratio (K=0.84); lignin content (K=0.74); root diameter (K=0.67). (3) Nitrogen fixation -- primary traits: nodulation type (K=0.84; binary; concentrated in Fabaceae and 8 other clades); rhizobium host specificity (K=0.74). (4) Litter production -- primary traits: plant height (K=1.24); leaf area (K=0.84); leaf lifespan (K=0.74). (5) Soil respiration -- primary traits: root turnover rate (K=0.67); soil carbon content (K=0.67). (6-12) Additional Category A functions: biomass allocation (K=0.74); seed dispersal effectiveness (K=0.67); mycorrhizal colonisation (K=0.84); herbivore pressure resistance via structural defences (K=0.74); soil aggregate stability via root binding (K=0.74); predation pressure from morphologically conserved predators (K=0.74); water use efficiency via root architecture (K=0.67).

Category B functions (K < 0.5; FRic required): (1) Crop pollination -- primary traits: bee body size (K=0.54; moderate); tongue length (K=0.38); foraging constancy (K=0.24). (2) Pest control -- primary traits: foraging strategy (K=0.18); prey specificity (K=0.24); search efficiency (K=0.28). (3) Primary productivity (water-limited systems) -- primary traits: photosynthetic pathway (K=0.14; C4 vs C3 convergence); stomatal conductance (K=0.28). (4) Methane production (wetlands) -- primary traits: anaerobic metabolic pathway (K=0.18; widely convergent in wetland taxa); methanotrophy (K=0.18). (5) Herbivory resistance via chemical defences -- primary traits: secondary metabolite chemistry (K=0.28; strong convergence in alkaloids; terpenes; phenolics across distantly related plant families). (6) Thermal regulation effectiveness -- primary traits: CTmax (K=0.24); thermal acclimation capacity (K=0.28). For all Category B functions, trait measurement databases (TRY for plants; AVONET for birds; PanTHERIA for mammals) or primary surveys are required for reliable FRic computation.