Inheritance and inequality under communal land tenure: convergence on cultural–evolutionary regimes

Nicolas Montalva 0000-0001-8612-663X[1*],
Francisca Vásquez-Estay 0009-0006-8013-1693[1,2],
Margarita Reyes-Madrid 00000-0003-1521-0994[1,3],
Pablo Varas-Enríquez 0000-0003-1521-0994[4,5,6]

[1*]Society and Health Research Center, Universidad Mayor, José Toribio Medina 29, Santiago, 8340589, Región Metropolitana, Chile.
[2]Escuela de Salud Pública, Facultad de Medicina, Universidad de Chile, Av. Independencia 1027, Santiago, 8380453, Región Metropolitana, Chile.
[3]Núcleo Interdisciplinario de Biología y Genética, Instituto de Ciencias Biomédicas (ICBM), Facultad de Medicina, Universidad de Chile, Av. Independencia 1027, Santiago,8380453, Región Metropolitana, Chile.
[4]Department of Social and Behavioral Sciences, Toulouse School of Economics, Université Toulouse Capitole, 1, Esplanade de l'Université, Toulouse, 31000, France.
[5]BirthRites Lise Meitner Research Group, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig, 04103, Germany.
[6]Department of Human Behavior, Ecology, and Culture, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, Leipzig, 04103, Germany.

Contributing authors: nicolas.montalva@umayor.cl;
fran.vasquez.e@gmail.com; ma.reyesmadrid@gmail.com;
pablo.varas-enriquez@tse-fr.eu;

1 Extended results

1.1 Trees correspondence

Topological comparisons indicated that T_s and T_g shared moderate similarity (Baker's $\gamma = 0.542$, $p = 0.013$; entanglement = 0.242), whereas both showed little correspondence with the consensus tree T_c (Baker's $\gamma < 0.30$, non–significant). These results suggest that while surname and genetic trees overlap to some extent, they are not interchangeable in fine–scale topology.

Table 1: **Descriptive statistics of inferred trees.** T_a is the surname tree for all communities, T_s the surname tree for the genotyped subset, T_g the STR-based genetic tree, and T_c the consensus of T_s and T_g.BL = branch length;Colless and Sackin are normalised imbalance indices.

	T_a	T_s	T_g	T_c
Tips	170	15	15	15
Ultrametric	Yes	Yes	Yes	No
Height	0.4921	0.4486	0.4528	0.6942
Mean BL	0.1466	0.1744	0.1445	0.1793
Colless	1.1491	—	—	0.7402
Sackin	1.4820	—	—	1.4287

— indicates values not computed.

Table 2: **Pairwise Mantel correlations among surname (D_s), genetic (D_g), and geographic (D_{geo}) distance matrices for the sampled subset ($n = 15$ communities).** Entries show Mantel r (two–sided p).

	D_s	D_g	D_{geo}
D_s		$0.52(< 0.01)^*$	$0.60(< 0.01)^*$
D_g			0.38(0.05)
D_{geo}			

* $p < 0.01$.

1.2 Other estimators of phylogenetic signal

For completeness, we report phylogenetic signal estimates using both Blomberg's K and Pagel's λ, and for both the full surname tree (T_a) and the consensus tree (T_c). These results are supplementary to the main text, which focuses on Blomberg's K on T_a. All traits were analysed on the logit–transformed scale (with continuity correction), and significance was evaluated by permutation (9,999 permutations). Estimates based on T_c should be interpreted cautiously given the small sample size ($n = 15$) and the construction of the consensus tree. For ease of comparison, Table 3 repeats the $T_a - K$ results reported in the main text.

Table 3: **Supplementary phylogenetic signal estimates across trees and estimators.**Pagel's λ and Blomberg's K for surname diversity (S), agnatic bias (A), inequality (G), and multigeniture (M) on the full surname tree (T_a) and the consensus tree (T_c).Traits were analysed on the logit–transformed scale (continuity correction);p values are from permutation tests (9,999 permutations).

Tree	Trait	λ	p_λ	K	p_K
T_a	S	0.000	1.000	0.611	0.296
	A	0.000	1.000	0.653	0.245
	G	0.291	0.094	0.687*	0.008*
	M	0.060	0.863	0.654*	0.037*
T_c	S	0.161	0.543	0.318	0.120
	A	0.000	1.000	0.201	0.610
	G	0.991*	0.014*	0.707*	0.038*
	M	1.025*	¡0.001*	1.201*	¡0.001*

* $p < 0.05$.

To validate that the detected phylogenetic signal was not an artefact of chance alignment between trait values and tree labels, we conducted a randomisation test for G and M on both T_a and T_c. For each tree, we generated 1,000 randomised replicates by permuting tip labels (equivalently, permuting trait values across tips) while keeping tree topology and branch lengths fixed. For each replicate, we re–estimated Blomberg's K (and Pagel's λ where reported). The observed statistics were then compared against the null distributions from the randomised replicates. For K in T_a, where signal was reported in the main analyses, the observed values lay in the extreme tail of the corresponding null distribution (Figure 1).

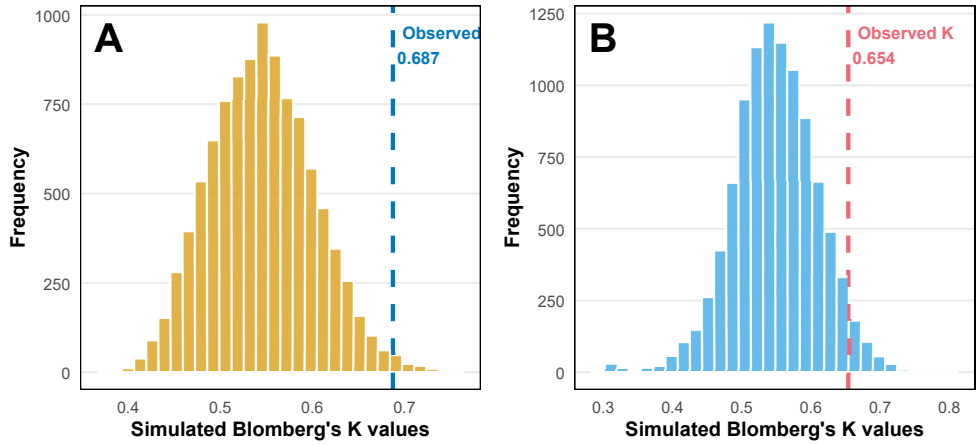

Fig. 1: **Randomisation test for phylogenetic signal on the full surname tree (T_a).** Null distributions of Blomberg's K for inequality (G; panel A) and multigeniture (M; panel B) obtained from 1,000 tip-label permutations (tree topology and branch lengths held fixed). Dashed vertical lines indicate the observed K values. Traits were analysed on the logit–transformed scale (with continuity correction).

1.3 Results of phylogenetic regression

Full model specification, priors, and MCMC settings for the Bayesian zero–inflated beta phylogenetic regression are reported in Online Resource 2. Here we provide the full posterior summary including distributional and phylogenetic variance parameters (Table 4).

As a sensitivity analysis consistent with the Gaussian/logit framework used by SURFACE model fitting, we fitted a frequentist PGLS model to logit–transformed variables (Table 5), using the same continuity correction described in the Methods. Phylogenetic dependence was modelled by estimating Pagel's λ by maximum likelihood.

1.4 Evolutionary model fitting and SURFACE outputs

For the best–fitting SURFACE models, we report the number of inferred shifts (k), the number of distinct regimes after convergence (k'), the proportion of convergent shifts (c/k), and OU parameters (α, σ^2). For interpretability, optima (θ) are shown both on the fitted (logit) scale and back–transformed to the original $[0, 1]$ scale (Table 6).

In some fits, SURFACE estimated an additional optimum with no communities assigned after the backward phase; such unoccupied regimes are reported for completeness but are not interpreted.

Table 4: **Full posterior summary for the Bayesian phylogenetic regression (surname tree T_a, $n = 170$).** The response (G) was modelled with a zero–inflated beta likelihood; predictors (M, A, S) were standardised. Entries are posterior means (b) and 95% credible intervals (LCrI, UCrI). $\phi_{\text{Intercept}}$ is the (log) precision intercept, $zi_{\text{Intercept}}$ is the (logit) zero–inflation intercept, and sd(phylo) is the phylogenetic random–effect standard deviation.

Term	b	LCrI	UCrI
Intercept	-1.85	-2.10	-1.63
M (std.)	0.65	0.53	0.77
A (std.)	-0.09	-0.25	0.07
S (std.)	-0.05	-0.19	0.09
$\phi_{\text{Intercept}}$ (log)	3.06	2.64	3.81
$zi_{\text{Intercept}}$ (logit)	-0.71	-1.03	-0.39
sd(phylo)	0.24	0.01	0.58
Bayesian R^2	0.45	0.32	0.58

Table 5: **Phylogenetic GLS sensitivity analysis (surname tree T_a, $n = 170$).** PGLS fitted to logit–transformed variables with Gaussian errors, estimating Pagel's λ by maximum likelihood. Entries show coefficients (b), 95% confidence intervals (LCI, UCI), and p values.

Term	b	LCI	UCI	p
Intercept	-0.68	-1.15	-0.21	¡0.01
M	0.70	0.63	0.78	¡0.01
A	-0.10	-0.38	0.18	0.47
S	-0.14	-0.33	0.06	0.18
Fit: logLik $= -258.68$, $\lambda = 0.57$, $R^2 = 0.73$				

Table 6: **SURFACE summary parameters for multi–optima OU models on** T_a**.**α and σ^2 are estimated on the fitted (logit) scale. k is the number of inferred regime shifts, k' the number of distinct regimes after convergence, and c/k the proportion of convergent shifts.

Trait	α	σ^2	k	k'	c/k
G	22.36	155.27	9	3	0.889
M	8.50	65.23	14	5	0.857