Supplementary materials for

"Are nonsynonymous transversions generally more deleterious than nonsynonymous transitions?"

Z. Zou and J. Zhang (jianzhi@umich.edu)

The supplementary materials include:
Legends of supplementary figures
Figures S1-S8
Table S1

Legends of supplementary figures

Fig. S1. Correlations between the estimated η and estimated (**a**) transition bias at the mutational level (κ), (**b**) fixation probability of nonsynonymous transversions relative to that of synonymous transversions (ω), and (**c**) genetic distance (d) in the 90 clades sampled across the tree of life. Spearman's correlation coefficient ρ and associated P-value are shown in each panel.

Fig. S2. Estimated η's show minor deviations from the true values in simulations under Eq. (1) with among-site ω variation, under different values of (**a**) genetic distance (d), (**b**) transition bias at the mutational level (κ), and (**c**) the fixation probability of nonsynonymous transversions relative to that of synonymous transversions (mean ω), or (**d**) with among-site variations in both ω and equilibrium frequencies $\boldsymbol{\pi}$. In (a)-(c), only the gamma shape parameter α of the among-site ω variation varied, as shown on the x-axis. In (d), only the mean ω varied. In all panels, each dot is one η estimate plotted against the α or mean ω used in the simulation. The true value of an η estimate is indicated by its color, and the dotted lines correspond to the true η values for easy comparison. In each plot, except for the parameter varied, the other parameters used in the simulation are $d = 1$, $\kappa = 2$, mean $\omega = 0.06$, and $\alpha = 1$.

Fig. S3. Even in the presence of among-site ω variation, variations in κ, ω_0, $\boldsymbol{\pi}$, d, or their combination are insufficient to explain the observed large among-clade variation in the estimated η. Here, η's estimated from sequence alignments simulated under Eq. (2) are plotted against the true values of (**a**) κ, (**b**) ω_0, (**c**) $\boldsymbol{\pi}$, (**d**) d, or (**e**) combination of κ, ω_0, $\boldsymbol{\pi}$, and d used in the simulation. In (a), (b), and (d), only the parameter on the x-axis varied in the simulations. In (c) and (e), the observed codon frequencies and combination of κ, ω_0, $\boldsymbol{\pi}$, and d of the 90 clades are respectively used in the simulations, and the insets plot the η estimated from the real sequence alignment of each clade against that estimated from the simulated alignment. For each parameter value, the η estimates from 10 replicate simulations are shown as dots (in panels a, b, and d) or boxplots (in panels c and e). In each boxplot, the lower and upper edges of a box represent the first (qu_1) and third (qu_3) quartiles, respectively, the horizontal line inside the box indicates the median (md), and the whiskers extend to the most extreme values inside inner fences, $\mathrm{md}\pm1.5(\mathrm{qu}_3-\mathrm{qu}_1)$. In the insets of panels (c) and (e), the mean estimate from the 10 replicate simulations is shown on the y-axis. In each plot, except for the parameter(s) varied, the other parameters used in the simulation are $d = 1$, $\kappa = 2$, $\omega_0 = 0.06$, and $\alpha = 1$.

Fig. S4. In the presence of among-site ω variation, variation of relative amino acid exchangeabilities (REs) among clades can still explain the observed η variation. (**a**) Simulations with RE values derived from the Grantham matrix (see Materials and Methods). For each of the 10 new RE sets at a given level of deviation from or shuffled from the original values, five replicate sequence evolution simulations are conducted and the corresponding η estimates are plotted. Different RE sets at each deviation level and from each independent shuffle are distinguished by different (randomly assigned) colors. (**b**) The η's estimated from the 90 clades simulated using the corresponding RE values of the real clades are plotted against the η's estimated from the real clades. Dots are colored by the corresponding taxonomic group of the clades, as shown in the legend. The dashed red line indicates $y = x$. The y-axis value of each dot is the mean estimate from 10 replicate simulations.

Fig. S5. The observed η variation among clades cannot be caused primarily by mutational biases unconsidered in η inference. Shown are comparisons between η's estimated from real data and those estimated from data simulated using mutational biases of the real data, when the simulation is performed in the (**a**) absence or (**b**) presence of among-site ω variation. The gamma shape parameter for ω variation in (b) is 1. Mutational bias is estimated from four-fold degenerate sites of the real data for each clade using a General-Reverse-Time model. The dashed red line indicates $y = x$. The y-axis value of each dot is the mean estimate from 10 replicate simulations.

Fig. S6. Synonymous (κ) and nonsynonymous ($\kappa\eta$) transition biases for each of the 90 clades, estimated under Eq. (1) (modified Goldman-Yang model; blue dots) or Eq. (4) (modified FMutSel model; orange dots).

Fig. S7. The η values of the 90 clades estimated under Eq. (4) (modified FMutSel model) correlate well with those estimated under Eq. (1) (modified Goldman-Yang model). The red line indicates $y = x$.

Fig. S8. The RE estimator of Zou and Zhang (2019) is robust to the violation of the assumption of no selection on synonymous mutations. We simulated 90 clades according to a modified version of the FMutSel model with the RE terms incorporated into the transition rates of corresponding codon changes, using REs estimated by Zou and Zhang (2019) and ***F*** estimated in the present study from the real data of the 90 clades. We then estimated REs using the method of Zou and Zhang (2019). Shown are the frequency distributions of (A) Pearson and (B) Spearman correlations between the true REs used in the simulation and the estimated REs from the simulated data for the 90 clades simulated.

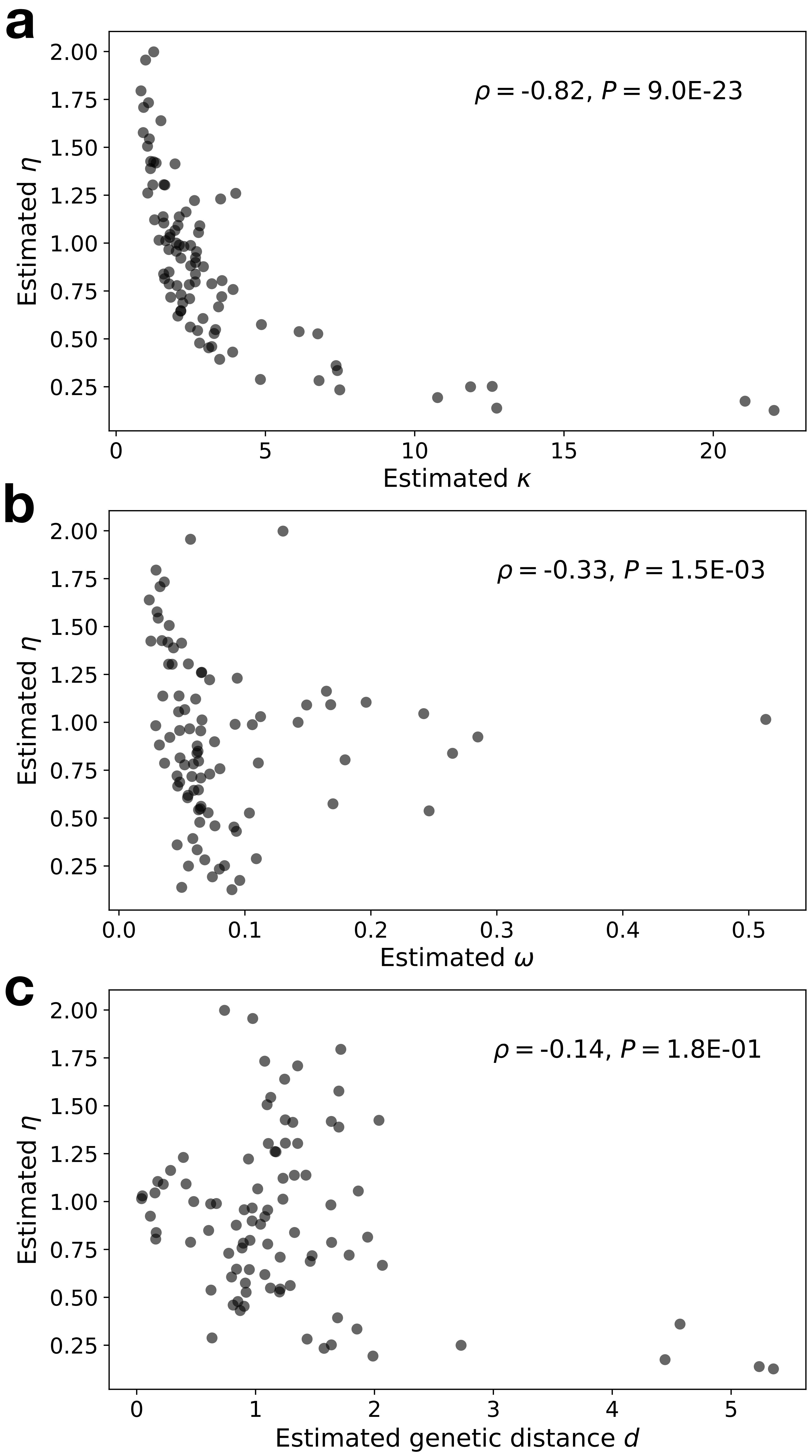
a
Estimated η
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
$\rho = $-0.82, $P = $9.0E-23
0
5
10
15
20
Estimated κ
b
Estimated η
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
$\rho = $-0.33, $P = $1.5E-03
0.0
0.1
0.2
0.3
0.4
0.5
Estimated ω
c
Estimated η
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
$\rho = $-0.14, $P = $1.8E-01
0
1
2
3
4
5
Estimated genetic distance d

Figure S1

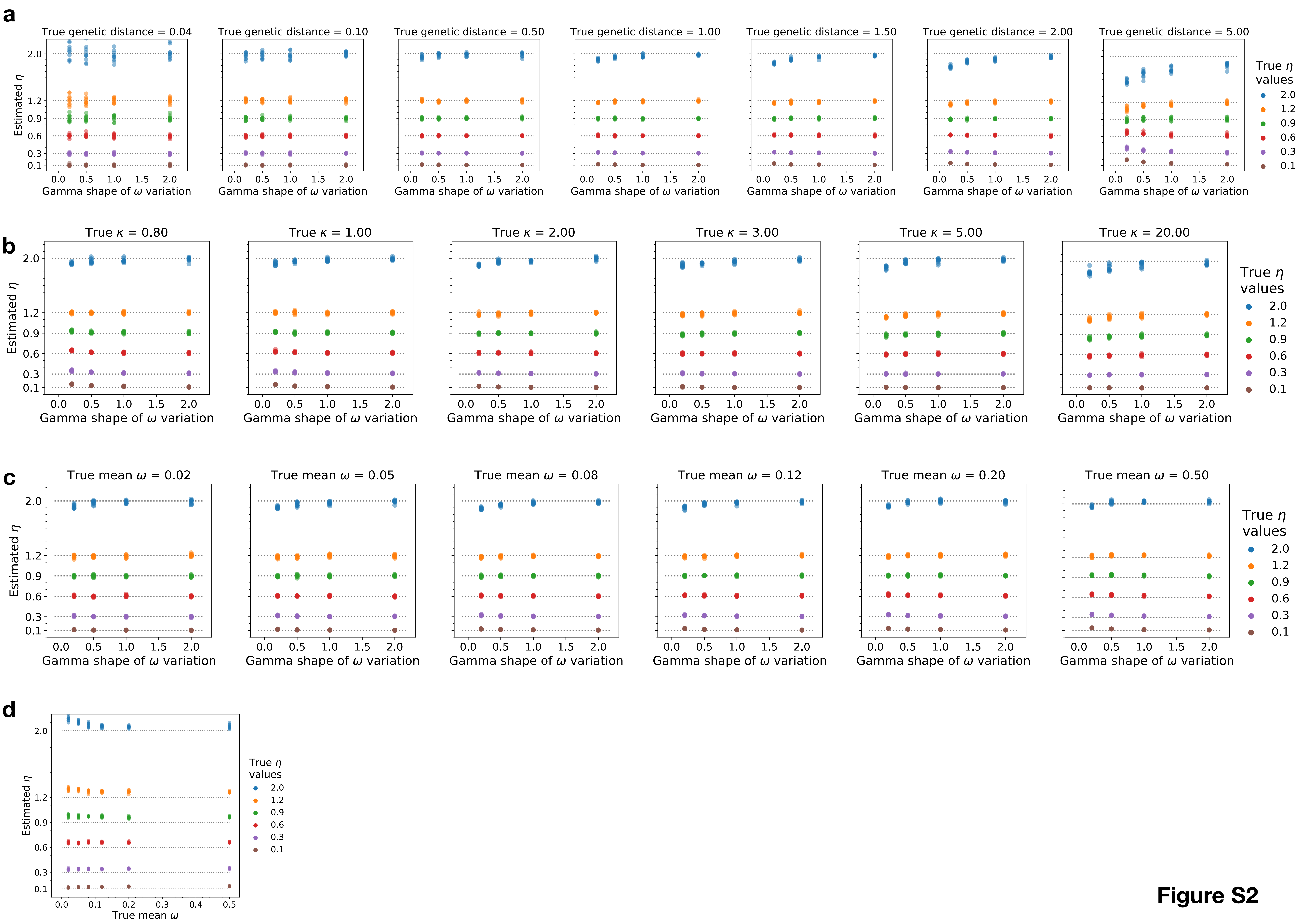
a
True genetic distance = 0.04
True genetic distance = 0.10
True genetic distance = 0.50
True genetic distance = 1.00
True genetic distance = 1.50
True genetic distance = 2.00
True genetic distance = 5.00
Estimated η
0.1
0.3
0.6
0.9
1.2
2.0
0.0
0.5
1.0
1.5
2.0
Gamma shape of ω variation
True η values
2.0
1.2
0.9
0.6
0.3
0.1
b
True κ = 0.80
True κ = 1.00
True κ = 2.00
True κ = 3.00
True κ = 5.00
True κ = 20.00
c
True mean ω = 0.02
True mean ω = 0.05
True mean ω = 0.08
True mean ω = 0.12
True mean ω = 0.20
True mean ω = 0.50
d
0.0
0.1
0.2
0.3
0.4
0.5
True mean ω

Figure S2

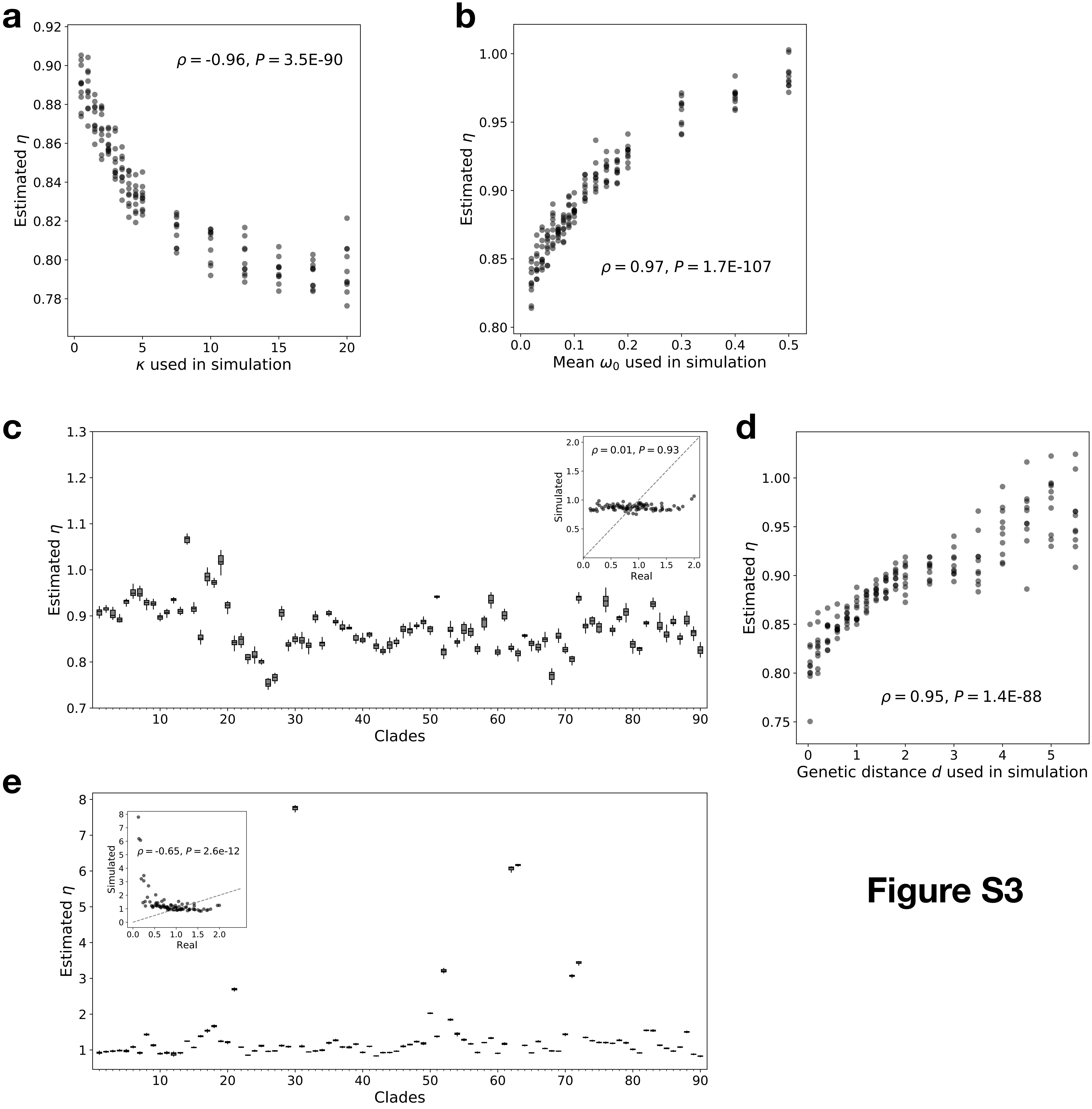
a
0.92
0.90
0.88
0.86
0.84
0.82
0.80
0.78
Estimated η
ρ = -0.96, P = 3.5E-90
0
5
10
15
20
κ used in simulation
b
1.00
0.95
0.90
0.85
0.80
Estimated η
ρ = 0.97, P = 1.7E-107
0.0
0.1
0.2
0.3
0.4
0.5
Mean ω0 used in simulation
c
1.3
1.2
1.1
1.0
0.9
0.8
0.7
Estimated η
10
20
30
40
50
60
70
80
90
Clades
ρ = 0.01, P = 0.93
Simulated
2.0
1.5
1.0
0.5
0.0
Real
d
1.00
0.95
0.90
0.85
0.80
0.75
Estimated η
ρ = 0.95, P = 1.4E-88
0
1
2
3
4
5
Genetic distance d used in simulation
e
8
7
6
5
4
3
2
1
Estimated η
10
20
30
40
50
60
70
80
90
Clades
ρ = -0.65, P = 2.6e-12
Simulated
Real

Figure S3

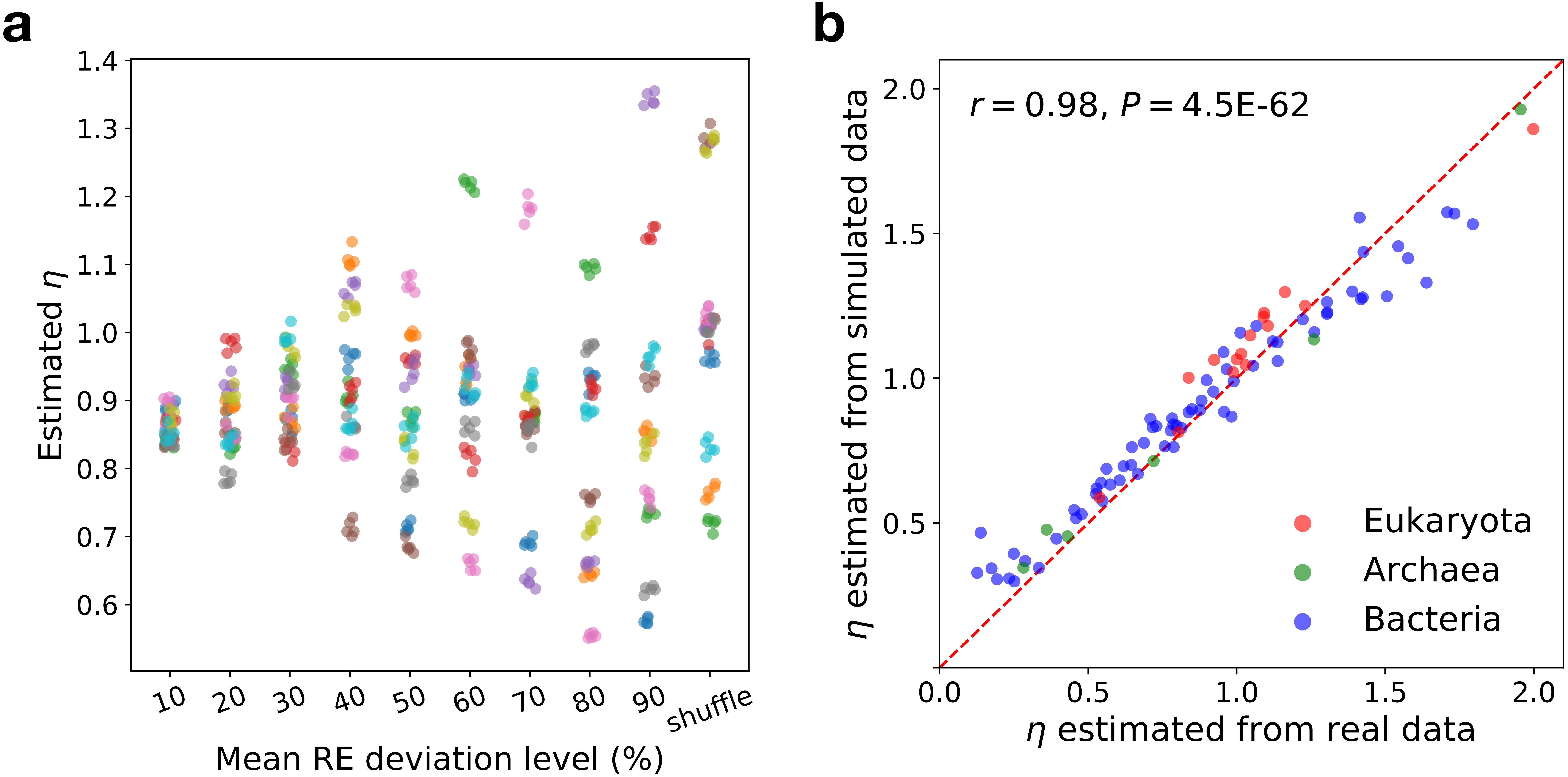
a
Estimated η
1.4
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
10
20
30
40
50
60
70
80
90
shuffle
Mean RE deviation level (%)
b
$r = 0.98$, $P = 4.5$E-62
η estimated from simulated data
2.0
1.5
1.0
0.5
0.0
0.5
1.0
1.5
2.0
η estimated from real data
Eukaryota
Archaea
Bacteria

Figure S4

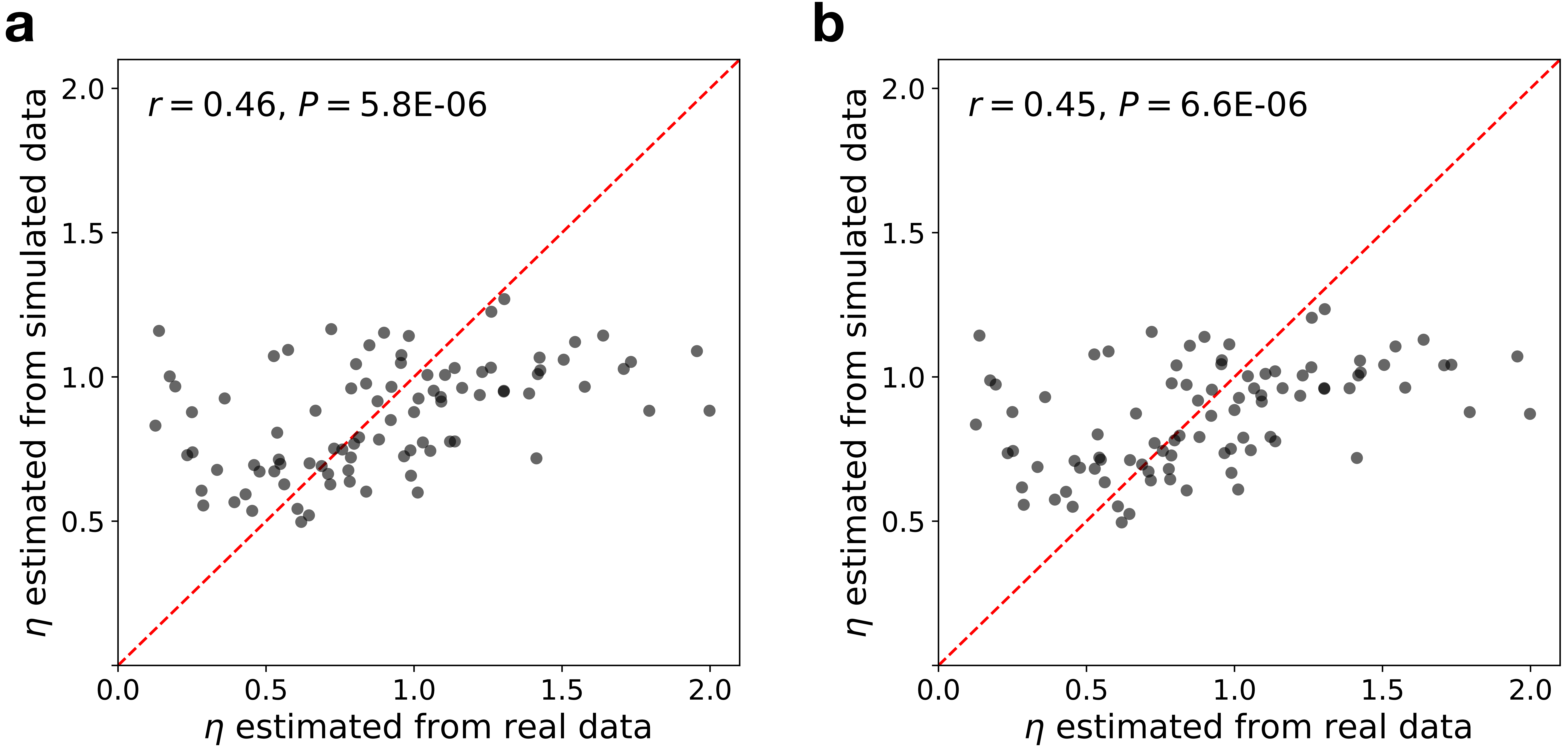
a
b
$r = 0.46$, $P = 5.8$E-06
$r = 0.45$, $P = 6.6$E-06
η estimated from simulated data
η estimated from real data
0.0
0.5
1.0
1.5
2.0

Figure S5

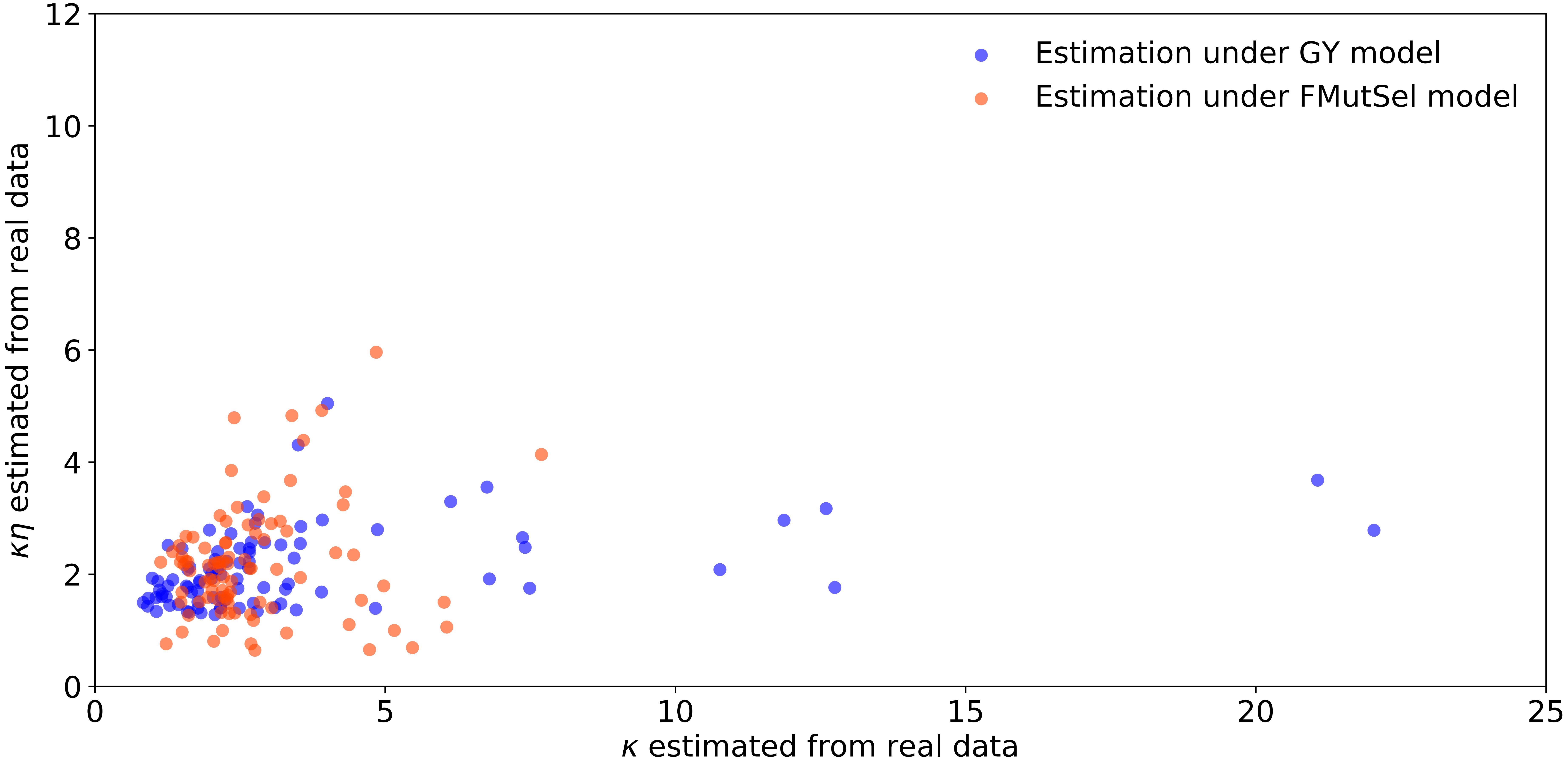
Estimation under GY model
Estimation under FMutSel model
$\kappa\eta$ estimated from real data
κ estimated from real data
0
2
4
6
8
10
12
0
5
10
15
20
25

Figure S6

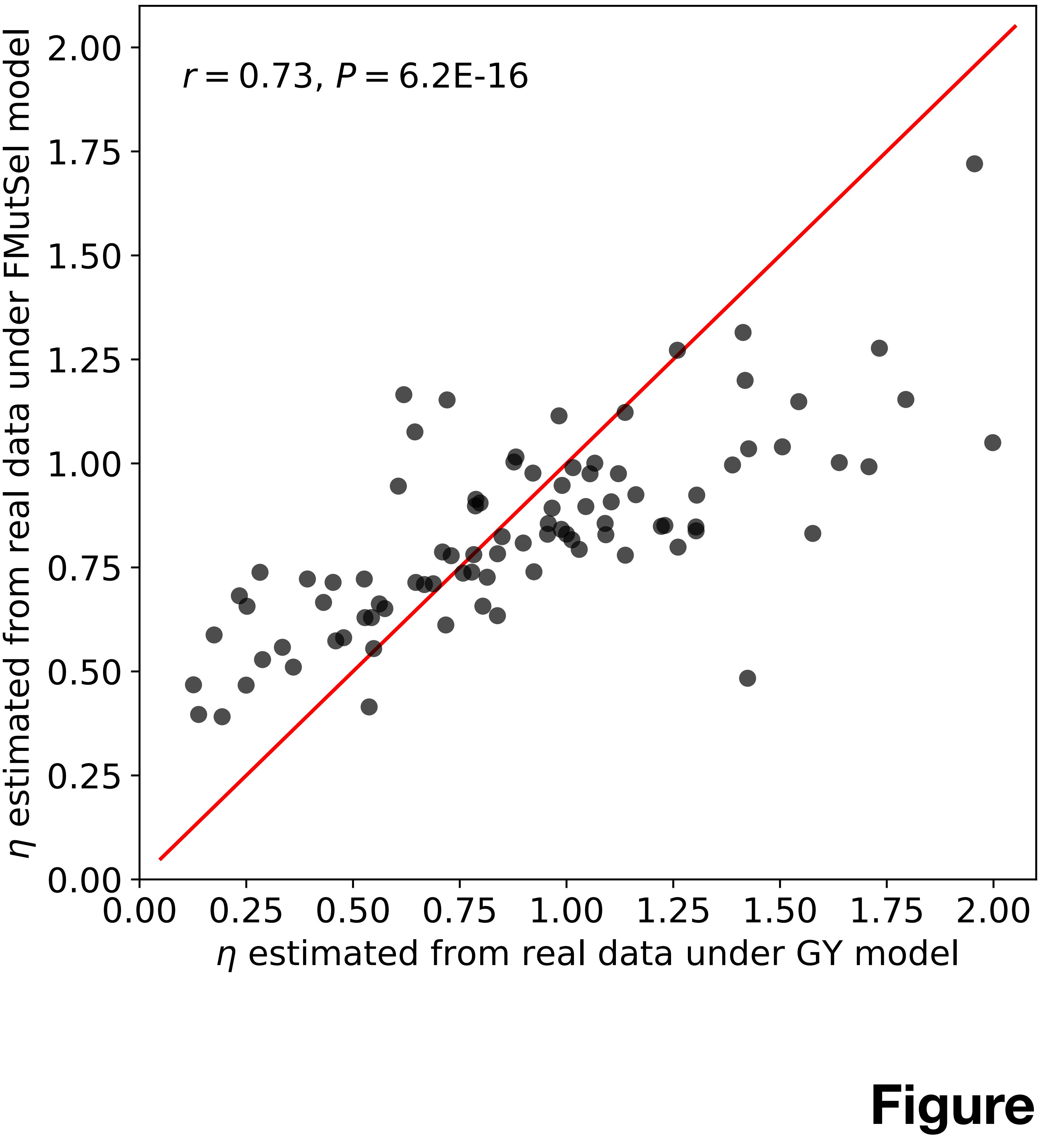
$r = 0.73$, $P = 6.2\text{E-}16$
η estimated from real data under FMutSel model
η estimated from real data under GY model
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00

Figure S7

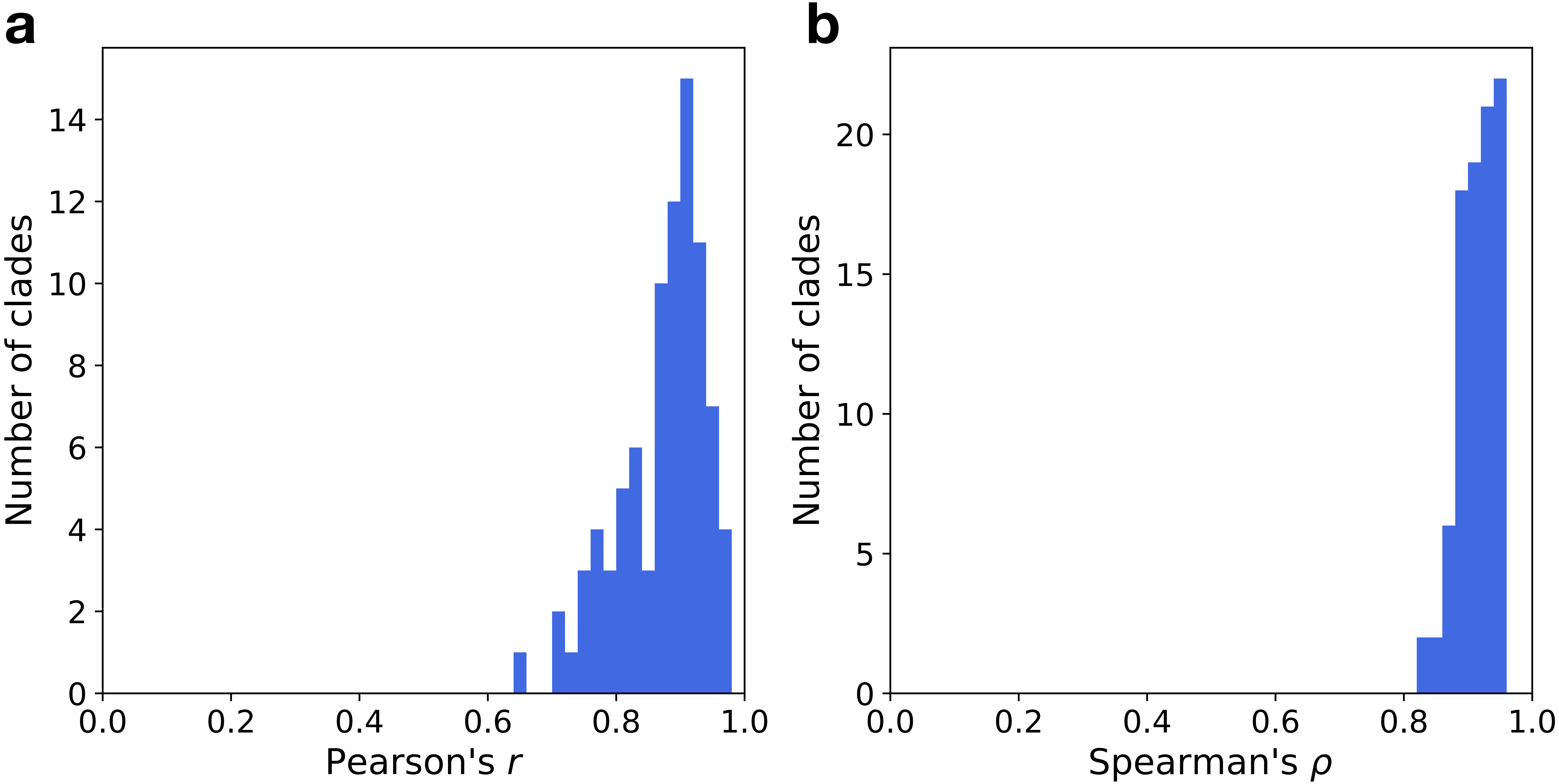
a
b
Number of clades
0
2
4
6
8
10
12
14
0.0
0.2
0.4
0.6
0.8
1.0
Pearson's r
Number of clades
0
5
10
15
20
0.0
0.2
0.4
0.6
0.8
1.0
Spearman's ρ

Figure S8

Table S1. Taxon compositions, data sources and estimated η values of all 90 clades for which η's are estimated. The order and indexing are consistent with those in figures.

	Clade name	Domain	η estimation	Taxon 1 (designated in shuffling test)	Taxon 2	Data source
1	03_HsaMmu	Eukaryota	0.92	*Homo sapiens*	*Macaca mulatta*	OrthoMaM v9[1]
2	02_MusRat	Eukaryota	1.09	*Mus musculus*	*Rattus norvegicus*	OrthoMaM v9
3	01_FcaCaf	Eukaryota	1.16	*Canis lupus familiaris*	*Felis catus*	OrthoMaM v9
4	09_SarMon	Eukaryota	1.09	*Monodelphis domestica*	*Sarcophilus harrisii*	OrthoMaM v9
5	04_Gallus	Eukaryota	0.84	*Gallus gallus*	*Meleagris gallopavo*	Ensembl 84[2]
6	12_fugu	Eukaryota	1.00	*Takifugu rubripes*	*Tetraodon nigroviridis*	Ensembl 84
7	05_Droso	Eukaryota	1.03	*Drosophila sechellia*	*Drosophila simulans*	Flybase[3]
8	10_ants	Eukaryota	0.54	*Atta cephalotes*	*Solenopsis invicta*	Ensembl Metazoa[4]
9	07_yeast	Eukaryota	1.23	*Saccharomyces cerevisiae*	*Saccharomyces paradoxus*	http://www.saccharomycessensustricto.org/
10	14_Fusarium	Eukaryota	0.80	*Fusarium graminearum*	*Fusarium pseudograminearum*	Ensembl Fungi[5]
11	08_Arab	Eukaryota	1.10	*Arabidopsis thaliana*	*Arabidopsis lyrata*	Ensembl Plants[6]
12	15_Oryza	Eukaryota	1.02	*Oryza sativa Japonica*	*Oryza glaberrima*	Ensembl Plants
13	16_Solanum	Eukaryota	1.05	*Solanum tuberosum*	*Solanum lycopersicum*	Ensembl Plants
14	11_malaria	Eukaryota	2.00	*Plasmodium vivax*	*Plasmodium knowlesi*	Ensembl Protists[7]
15	13_Phytophthora	Eukaryota	0.99	*Phytophthora infestans*	*Phytophthora parasitica*	Ensembl Protists
16	ATGC095	Archaea	1.26	Vulcanisaeta_moutnovskia_768_28.GCF_000190315.1	Vulcanisaeta_distributa_DSM_14429.GCF_000148385.1	ATGC[8]
17	ATGC130	Archaea	0.28	Haloferax_mediterranei_ATCC_33500.GCF_000306765.2	Haloferax_gibbonsii.GCF_001190965.1	ATGC
18	ATGC152	Archaea	0.72	Thermococcus_guaymasensis_DSM_11113.GCF_000816105.1	Thermococcus_AM4.GCF_000151205.2	ATGC
19	ATGC202	Archaea	1.96	Methanothermobacter_thermautotrophicus_Delta_H.GCF_000008645.1	Methanothermobacter_marburgensis_Marburg.GCF_000145295.1	ATGC
20	ATGC277	Archaea	0.43	Halorhabdus_utahensis_DSM_12940.GCF_000023945.1	Halorhabdus_tiamatea_SARL4B.GCF_000470655.1	ATGC
21	ATGC279	Archaea	0.36	Thermococcus_sibiricus_MM_739.GCF_000022545.1	Thermococcus_2319x1.GCF_001484685.1	ATGC
22	ATGC001	Bacteria	0.78	Shigella_flexneri_4c.GCF_001579965.1	Salmonella_enterica_enterica_serovar_Gallinarum_str_287_91.GCF_000009525.1	ATGC
23	ATGC013	Bacteria	1.26	Streptococcus_parasanguinis_FW213.GCF_000262145.1	Streptococcus_A12.GCF_001587175.1	ATGC
24	ATGC015	Bacteria	1.39	Bacillus_velezensis_FZB42.GCF_000015785.1	Bacillus_subtilis.GCF_001604995.1	ATGC
25	ATGC016	Bacteria	0.76	Geobacillus_thermodenitrificans_NG80_2.GCF_000015745.1	Geobacillus_GHH01.GCF_000336445.1	ATGC
26	ATGC021	Bacteria	0.96	Chlamydia_trachomatis_E_SW3.GCF_000304495.1	Chlamydia_muridarum.GCF_000772145.1	ATGC
27	ATGC022	Bacteria	0.90	Chlamydia_psittaci_GR9.GCF_000298415.1	Chlamydia_felis_Fe_C_56.GCF_000009945.1	ATGC
28	ATGC028	Bacteria	0.64	Mycobacterium_vanbaalenii_PYR_1.GCF_000015305.1	Mycobacterium_gilvum_PYR_GCK.GCF_000016365.1	ATGC
29	ATGC044	Bacteria	0.79	Rickettsia_rickettsii_Morgan.GCF_000831545.1	Rickettsia_prowazekii_Rp22.GCF_000022785.1	ATGC
30	ATGC045	Bacteria	0.13	Ehrlichia_ruminantium_Welgevonden.GCF_000026005.1	Ehrlichia_HF.GCF_000632845.1	ATGC
31	ATGC050	Bacteria	1.22	Helicobacter_pylori_26695_1CH.GCF_000829135.1	Helicobacter_cetorum_MIT_99_5656.GCF_000259275.1	ATGC
32	ATGC056	Bacteria	1.42	Lactobacillus_rhamnosus_LOCK908.GCF_000418495.1	Lactobacillus_paracasei_N1115.GCF_000582665.1	ATGC
33	ATGC058	Bacteria	1.64	Lactobacillus_amylovorus_GRL1118.GCF_000194115.1	Lactobacillus_acidophilus_La_14.GCF_000389675.2	ATGC
34	ATGC069	Bacteria	1.12	Corynebacterium_deserti_GIMN1_010.GCF_001277995.1	Brevibacterium_flavum.GCF_000987865.1	ATGC
35	ATGC071	Bacteria	0.48	Pseudomonas_parafulva.GCF_000800255.1	Pseudomonas_CCOS_191.GCF_001007005.1	ATGC
36	ATGC073	Bacteria	0.56	Pseudomonas_stutzeri_DSM_4166.GCF_000195105.1	Pseudomonas_balearica_DSM_6083.GCF_000818015.1	ATGC
37	ATGC075	Bacteria	0.97	Pseudomonas_syringae_actinidiae_ICMP_18884.GCF_000648735.2	Pseudomonas_cichorii_JBC1.GCF_000517305.1	ATGC
38	ATGC078	Bacteria	0.65	Pseudomonas_chlororaphis.GCF_001602135.1	Pseudomonas_URMO17WK12_I11.GCF_001511755.1	ATGC
39	ATGC099	Bacteria	0.72	Pantoea_vagans_C9_1.GCF_000148935.1	Pantoea_rwandensis.GCF_000759475.1	ATGC
40	ATGC110	Bacteria	1.43	Vibrio_harveyi.GCF_001558435.1	Vibrio_EJY3.GCF_000241385.1	ATGC
41	ATGC111	Bacteria	0.73	Aeromonas_veronii_B565.GCF_000204115.1	Aeromonas_hydrophila_4AK4.GCF_000512185.1	ATGC
42	ATGC112	Bacteria	1.54	Vibrio_fischeri_ES114.GCF_000011805.1	Aliivibrio_wodanis.GCF_000953695.1	ATGC
43	ATGC117	Bacteria	1.30	Aggregatibacter_aphrophilus.GCF_001262035.1	Aggregatibacter_actinomycetemcomitans_ANH9381.GCF_000241025.1	ATGC
44	ATGC120	Bacteria	1.30	Shewanella_putrefaciens_CN_32.GCF_000016585.1	Shewanella_MR_4.GCF_000014685.1	ATGC
45	ATGC127	Bacteria	1.14	Yersinia_ruckeri.GCF_000964565.1	Yersinia_pestis_biovar_Medievalis_Harbin_35.GCF_000186725.1	ATGC
46	ATGC134	Bacteria	0.46	Xanthomonas_oryzae_oryzicola.GCF_001042795.1	Xanthomonas_campestris_campestris_str_8004.GCF_000012105.1	ATGC
47	ATGC136	Bacteria	1.07	Ochrobactrum_anthropi_ATCC_49188.GCF_000017405.1	Brucella_pinnipedialis_B2_94.GCF_000221005.1	ATGC
48	ATGC138	Bacteria	0.98	Francisella_tularensis_holarctica_LVS.GCF_000009245.1	Francisella_noatunensis_orientalis_FNO24.GCF_001042545.1	ATGC
49	ATGC143	Bacteria	0.92	Campylobacter_jejuni.GCF_001506225.1	Campylobacter_coli_RM5611.GCF_000583795.1	ATGC
50	ATGC145	Bacteria	0.53	Borrelia_turicatae_91E135.GCF_000012085.1	Borrelia_duttonii_Ly.GCF_000019685.1	ATGC
51	ATGC146	Bacteria	1.41	Thermotoga_neapolitana_DSM_4359.GCF_000018945.1	Thermotoga_maritima_MSB8.GCF_000978555.1	ATGC
52	ATGC155	Bacteria	0.19	Buchnera_aphidicola_G002_Myzus_persicae.GCF_000521565.1	Buchnera_aphidicola_Ak_Acyrthosiphon_kondoi.GCF_000225445.1	ATGC
53	ATGC159	Bacteria	0.33	Propionibacterium_avidum_44067.GCF_000367205.1	Propionibacterium_acnes_ATCC_11828.GCF_000231215.1	ATGC
54	ATGC169	Bacteria	0.23	Thermus_scotoductus_SA_01.GCF_000187005.1	Thermus_CCB_US3_UF1.GCF_000236585.1	ATGC
55	ATGC171	Bacteria	0.88	Thermoanaerobacter_kivui.GCF_000763575.1	Thermoanaerobacter_italicus_Ab9.GCF_000025645.1	ATGC

56	ATGC173	Bacteria	0.45	Rhodopseudomonas_palustris_TIE_1.GCF_000020445.1	Rhodopseudomonas_palustris_BisB5.GCF_000013685.1	ATGC
57	ATGC178	Bacteria	0.85	Prochlorococcus_marinus_pastoris_str_CCMP1986.GCF_000011465.1	Prochlorococcus_marinus_MIT_9515.GCF_000015665.1	ATGC
58	ATGC181	Bacteria	1.42	Caldicellulosiruptor_saccharolyticus_DSM_8903.GCF_000016545.1	Caldicellulosiruptor_bescii_DSM_6725.GCF_000022325.1	ATGC
59	ATGC182	Bacteria	1.14	Thermoanaerobacterium_xylanolyticum_LX_11.GCF_000189775.2	Thermoanaerobacterium_thermosaccharolyticum_M0795.GCF_000328545.1	ATGC
60	ATGC186	Bacteria	1.51	Sulfurospirillum_deleyianum_DSM_6946.GCF_000024885.1	Sulfurospirillum_barnesii_SES_3.GCF_000265295.1	ATGC
61	ATGC188	Bacteria	0.61	Ralstonia_solanacearum.GCF_001299555.1	Ralstonia_mannitolilytica.GCF_000954135.1	ATGC
62	ATGC201	Bacteria	0.17	Candidatus_Bartonella_ancashi.GCF_001281405.1	Bartonella_clarridgeiae_73.GCF_000253015.1	ATGC
63	ATGC223	Bacteria	0.14	Blattabacterium_Nauphoeta_cinerea.GCF_000471965.1	Blattabacterium_Mastotermes_darwiniensis_str_MADAR.GCF_000233435.1	ATGC
64	ATGC232	Bacteria	0.80	Hyphomicrobium_denitrificans_ATCC_51888.GCF_000143145.1	Hyphomicrobium_denitrificans_1NES1.GCF_000230975.2	ATGC
65	ATGC235	Bacteria	1.73	Pseudoalteromonas_haloplanktis_TAC125.GCF_000026085.1	Pseudoalteromonas_Bsw20308.GCF_000310105.2	ATGC
66	ATGC239	Bacteria	0.55	Edwardsiella_tarda.GCF_001011055.1	Edwardsiella_ictaluri_93_146.GCF_000022885.2	ATGC
67	ATGC242	Bacteria	1.01	Janthinobacterium_Marseille.GCF_000013625.1	Herminiimonas_arsenicoxydans.GCF_000026125.1	ATGC
68	ATGC243	Bacteria	0.81	Exiguobacterium_sibiricum_255_15.GCF_000019905.1	Exiguobacterium_MH3.GCF_000496635.1	ATGC
69	ATGC244	Bacteria	1.30	Psychrobacter_arcticus_273_4.GCF_000012305.1	Psychrobacter_alimentarius.GCF_001606025.1	ATGC
70	ATGC246	Bacteria	0.57	Candidatus_Liberibacter_asiaticus_psy62.GCF_000023765.2	Candidatus_Liberibacter_africanus_PTSAPSY.GCF_001021085.1	ATGC
71	ATGC258	Bacteria	0.25	Dictyoglomus_turgidum_DSM_6724.GCF_000021645.1	Dictyoglomus_thermophilum_H_6_12.GCF_000020965.1	ATGC
72	ATGC260	Bacteria	0.25	Candidatus_Arthromitus_SFB_rat_Yit.GCF_000283555.1	Candidatus_Arthromitus_SFB_mouse_NL.GCF_000709435.1	ATGC
73	ATGC301	Bacteria	0.53	Pseudomonas_stutzeri_DSM_10701.GCF_000279165.1	Pseudomonas_20_BN.GCF_000756775.1	ATGC
74	ATGC309	Bacteria	0.88	Campylobacter_subantarcticus_LMG_24377.GCF_000816305.1	Campylobacter_peloridis_LMG_23910.GCF_000816785.1	ATGC
75	ATGC330	Bacteria	0.54	Corynebacterium_singulare.GCF_000833575.1	Corynebacterium_aurimucosum_ATCC_700975.GCF_000022905.1	ATGC
76	ATGC332	Bacteria	0.62	Corynebacterium_marinum_DSM_44953.GCF_000835165.1	Corynebacterium_humireducens_NBRC_106098_DSM_45392.GCF_000819445.1	ATGC
77	ATGC339	Bacteria	0.69	Bordetella_holmesii_H558.GCF_000341485.2	Bordetella_avium_197N.GCF_000070465.1	ATGC
78	ATGC341	Bacteria	0.71	Bifidobacterium_dentium_JCM_1195_DSM_20436.GCF_001042595.1	Bifidobacterium_adolescentis.GCF_000817995.1	ATGC
79	ATGC352	Bacteria	0.29	Xanthomonas_translucens_undulosa.GCF_001021935.1	Xanthomonas_sacchari.GCF_000815185.1	ATGC
80	ATGC361	Bacteria	0.99	Paenibacillus_sabinae_T27.GCF_000612505.1	Paenibacillus_durus.GCF_000756615.1	ATGC
81	ATGC369	Bacteria	0.96	Carnobacterium_inhibens_gilichinskyi.GCF_000493735.1	Carnobacterium_17_4.GCF_000195575.1	ATGC
82	ATGC371	Bacteria	0.67	Fervidobacterium_pennivorans_DSM_9078.GCF_000235405.2	Fervidobacterium_islandicum.GCF_000767275.2	ATGC
83	ATGC372	Bacteria	1.06	Thermotoga_caldifontis_AZM44c09.GCF_000828655.1	Pseudothermotoga_hypogea_DSM_11164_NBRC_106472.GCF_000504105.1	ATGC
84	ATGC374	Bacteria	0.79	Paenarthrobacter_aurescens_TC1.GCF_000014925.1	Arthrobacter_ATCC_21022.GCF_001542565.1	ATGC
85	ATGC376	Bacteria	0.84	Glutamicibacter_arilaitensis_Re117.GCF_000197735.1	Glutamicibacter_arilaitensis.GCF_001302565.1	ATGC
86	ATGC389	Bacteria	1.79	Dokdonia_MED134.GCF_000152925.2	Dokdonia_4H_3_7_5.GCF_000212355.1	ATGC
87	ATGC393	Bacteria	0.78	Collimonas_fungivorans.GCF_001584145.1	Collimonas_arenae.GCF_000786695.1	ATGC
88	ATGC420	Bacteria	0.39	Thioalkalivibrio_versutus.GCF_001020955.1	Thioalkalivibrio_K90mix.GCF_000025545.1	ATGC
89	ATGC428	Bacteria	1.71	Cyclobacterium_marinum_DSM_745.GCF_000222485.1	Cyclobacterium_amurskyense.GCF_001050135.1	ATGC
90	ATGC431	Bacteria	1.58	Thioalkalimicrobium_cyclicum_ALM1.GCF_000214825.1	Thioalkalimicrobium_aerophilum_AL3.GCF_000227665.2	ATGC

---

[1] http://www.orthomam.univ-montp2.fr/orthomam_v9/html/

[2] https://useast.ensembl.org/index.html

[3] https://flybase.org/

[4] https://metazoa.ensembl.org/index.html

[5] http://fungi.ensembl.org/index.html

[6] http://plants.ensembl.org/index.html

[7] http://protists.ensembl.org/index.html

[8] http://dmk-brain.ecn.uiowa.edu/ATGC/