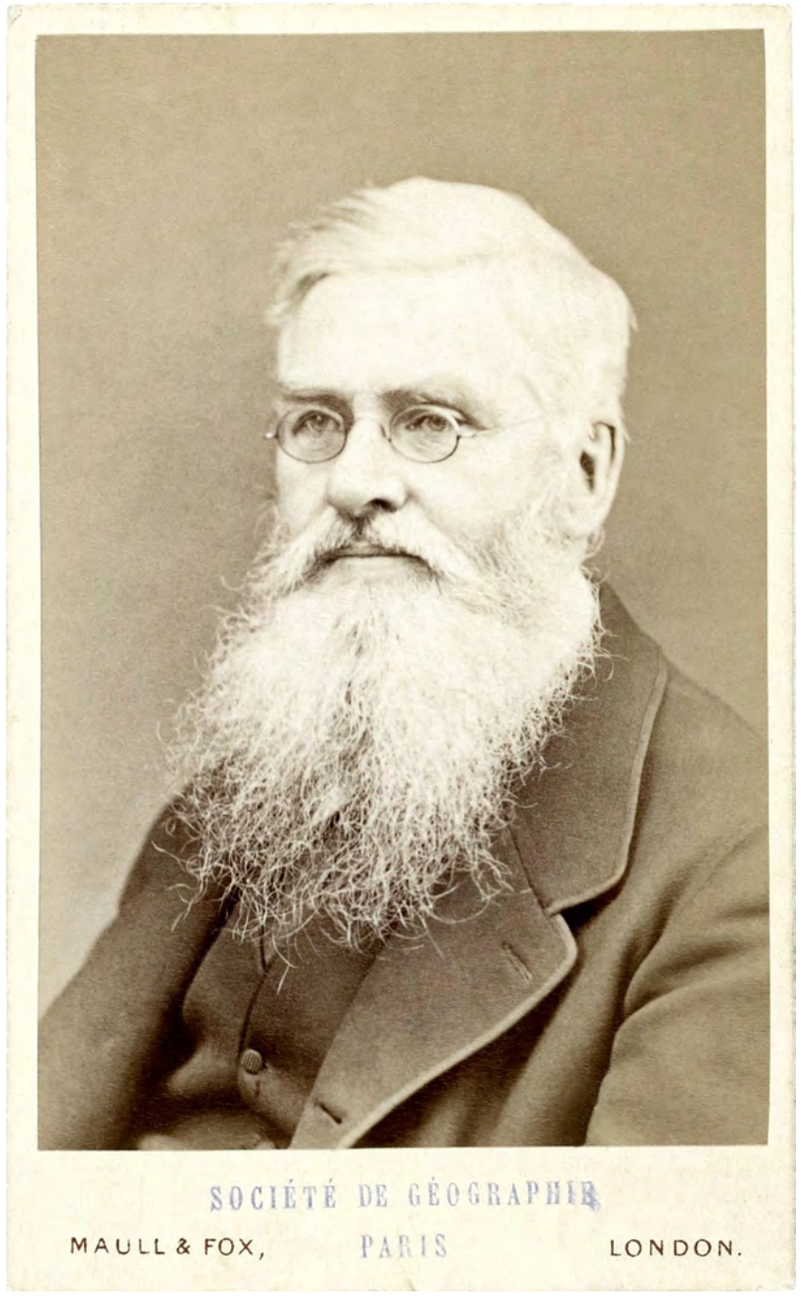
# Can we measure the importance of distance in biogeography?

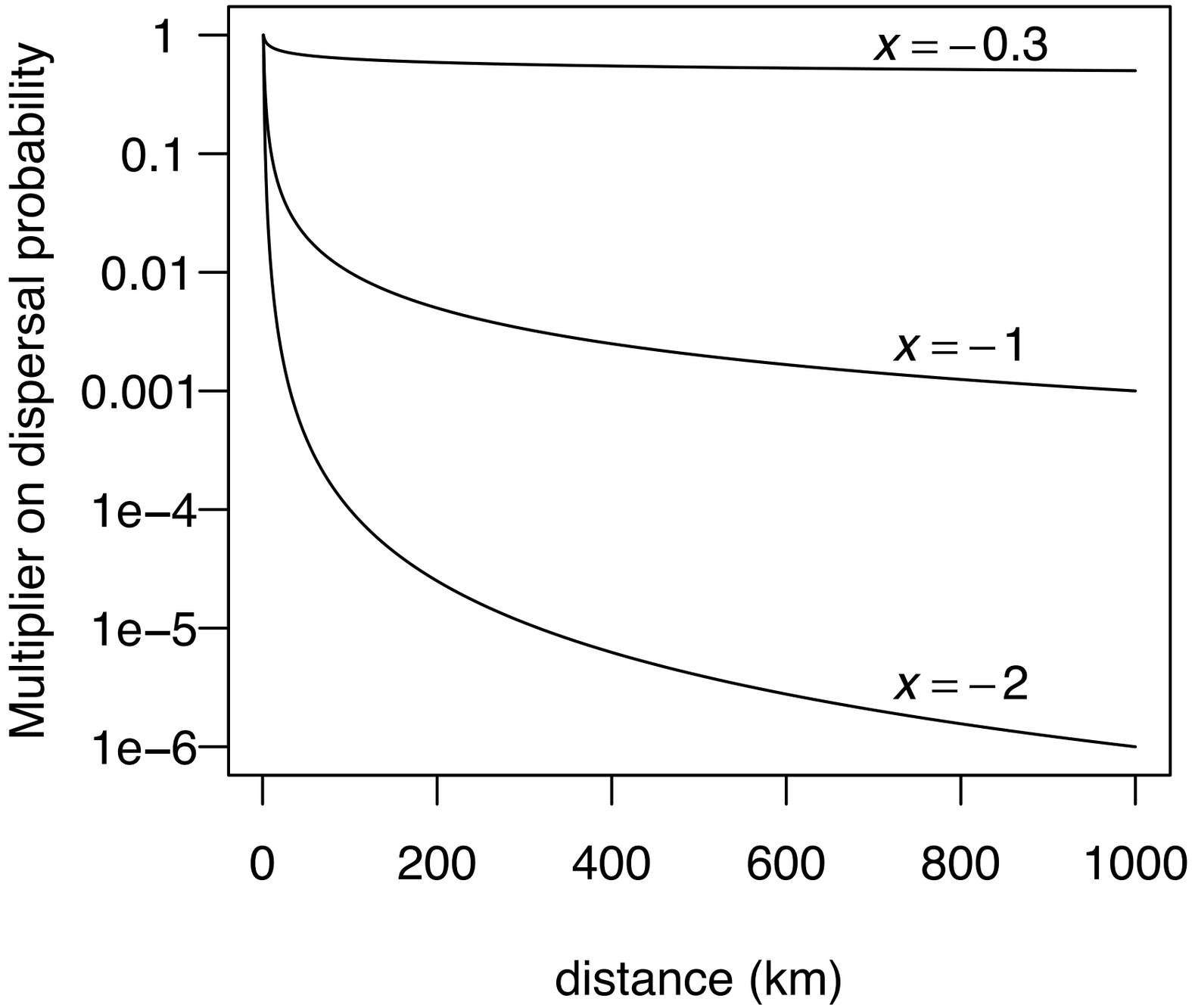
# Historical biogeography: Wallace and Darwin




**Wallace (1855): “Every species has come into existence coincident both in space and time with a pre-existing closely allied species.”**

**Darwin (1859): “the more nearly any two forms are related in blood, the nearer they will generally stand to each other in time and space”.**

# The “+x” model: modify dispersal probability by distance^*x*



*Journal of Biogeography* (J. *Biogeogr.*) (2017) 44, 887–899

ORIGINAL ARTICLE

# Bayesian estimation of the global biogeographical history of the Solanaceae

Julia Dupin[1]*, Nicholas J. Matzke[2], Tiina Sarkinen[3], Sandra Knapp[4], Richard G. Olmstead[5], Lynn Bohs[6] and Stacey D. Smith[1]

Journal of Biogeography (J. Biogeogr.) (2017) **44**, 887–899

ORIGINAL ARTICLE

# Bayesian estimation of the global biogeographical history of the Solanaceae

Julia Dupin[1]*, Nicholas J. Matzke[2], Tiina Särkinen[3], Sandra Knapp[4], Richard G. Olmstead[5], Lynn Bohs[6] and Stacey D. Smith[1]

http://solcap.msu.edu/img/germplasm2.jpg

Journal of Biogeography (J. Biogeogr.) (2017) **44**, 887–899

ORIGINAL ARTICLE

# Bayesian estimation of the global biogeographical history of the Solanaceae

Julia Dupin[1]*, Nicholas J. Matzke[2], Tiina Särkinen[3], Sandra Knapp[4], Richard G. Olmstead[5], Lynn Bohs[6] and Stacey D. Smith[1]

**50 to 24 Ma**

| | SAm | CAm | Car | NAm | AF | EU | OZ |
|---|---|---|---|---|---|---|---|
| SAm | 1 | 0.4 | 0.4 | 0.2 | 0.1 | 0.1 | 0.1 |
| CAm | 0.4 | 1 | 0.7 | 0.4 | 0.1 | 0.1 | 0.1 |
| Car | 0.4 | 0.7 | 1 | 0.4 | 0.1 | 0.1 | 0.1 |
| NAm | 0.2 | 0.4 | 0.4 | 1 | 0.7 | 1 | 0.3 |
| AF | 0.1 | 0.1 | 0.1 | 0.7 | 1 | 0.7 | 0.3 |
| EU | 0.1 | 0.1 | 0.1 | 0.8 | 0.7 | 1 | 0.3 |
| OZ | 0.1 | 0.1 | 0.1 | 0.3 | 0.3 | 0.3 | 1 |

**24 to 10 Ma**

| | SAm | CAm | Car | NAm | AF | EU | OZ |
|---|---|---|---|---|---|---|---|
| SAm | 1 | 0.6 | 0.5 | 0.5 | 0.1 | 0.2 | 0.1 |
| CAm | 0.6 | 1 | 0.7 | 0.6 | 0.1 | 0.1 | 0.1 |
| Car | 0.5 | 0.7 | 1 | 0.5 | 0.1 | 0.1 | 0.1 |
| NAm | 0.5 | 0.6 | 0.5 | 1 | 0.4 | 0.5 | 0.1 |
| AF | 0.1 | 0.1 | 0.1 | 0.4 | 1 | 1 | 0.3 |
| EU | 0.2 | 0.1 | 0.1 | 0.5 | 1 | 1 | 0.5 |
| OZ | 0.1 | 0.1 | 0.1 | 0.1 | 0.3 | 0.5 | 1 |

**10 Ma to Present**

| | SAm | CAm | Car | NAm | AF | EU | OZ |
|---|---|---|---|---|---|---|---|
| SAm | 1 | 1 | 0.75 | 1 | 0.1 | 0.2 | 0.1 |
| CAm | 1 | 1 | 0.7 | 1 | 0.1 | 0.1 | 0.1 |
| Car | 0.75 | 0.7 | 1 | 0.8 | 0.1 | 0.1 | 0.1 |
| NAm | 1 | 1 | 0.8 | 1 | 0.1 | 0.2 | 0.1 |
| AF | 0.1 | 0.1 | 0.1 | 0.1 | 1 | 1 | 0.7 |
| EU | 0.2 | 0.1 | 0.1 | 0.2 | 1 | 1 | 0.7 |
| OZ | 0.1 | 0.1 | 0.1 | 0.1 | 0.7 | 0.7 | 1 |

Time

# Global Solanaceae:
# distance, Biogeographical Stochastic Mapping

| | Models | Dispersal multipliers | Free parameters | | | | | Log-Likelihood | AIC | Δ AIC | AIC weights |
|---|---|---|---|---|---|---|---|---|---|---|---|
| # | Basic Models | | Number | $d^{(1)}$ | $e^{(2)}$ | $j^{(3)}$ | $w^{(4)}$ | | | | |
| 1 | DEC_NonTS | Non-TS | 2 | 0.011 | 0.000 | 0.000 | 1.000 | -1341.413 | 2686.826 | 317.604 | 0.000 |
| 2 | DIVALIKE_NonTS | Non-TS | 2 | 0.013 | 0.000 | 0.000 | 1.000 | -1401.936 | 2807.871 | 438.648 | 0.000 |
| 3 | BayAreaLIKE_NonTS | Non-TS | 2 | 0.009 | 0.049 | 0.000 | 1.000 | -1528.828 | 3061.656 | 692.433 | 0.000 |
| | **Time-Stratified Models** | | | | | | | | | | |
| 4 | DEC_TS | TS | 2 | 0.025 | 0.000 | 0.000 | 1.000 | -1210.832 | 2425.663 | 56.441 | 0.000 |
| 5 | DIVALIKE_TS | TS | 2 | 0.029 | 0.000 | 0.000 | 1.000 | -1252.831 | 2509.661 | 140.438 | 0.000 |
| 6 | BayAreaLIKE_TS | TS | 2 | 0.023 | 0.044 | 0.000 | 1.000 | -1430.614 | 2865.227 | 496.004 | 0.000 |
| | **+*j* Models** | | | | | | | | | | |
| 7 | DEC_ NonTS _*j* | Non-TS | 3 | 0.010 | 0.000 | 0.003 | 1.000 | -1324.525 | 2655.050 | 285.828 | 0.000 |
| 8 | DIVALIKE_NonTS_*j* | Non-TS | 3 | 0.011 | 0.000 | 0.003 | 1.000 | -1387.768 | 2781.536 | 412.313 | 0.000 |
| 9 | BayAreaLIKE_NonTS_*j* | Non-TS | 3 | 0.007 | 0.004 | 0.009 | 1.000 | -1335.728 | 2677.457 | 308.234 | 0.000 |
| 10 | DEC_TS_*j* | TS | 3 | 0.022 | 0.000 | 0.008 | 1.000 | -1191.907 | 2389.813 | 20.590 | 0.000 |
| 11 | DIVALIKE_ TS_*j* | TS | 3 | 0.025 | 0.000 | 0.007 | 1.000 | -1239.123 | 2484.246 | 115.024 | 0.000 |
| 12 | BayAreaLIKE_ TS_*j* | TS | 3 | 0.015 | 0.001 | 0.020 | 1.000 | -1213.992 | 2433.984 | 64.762 | 0.000 |
| | **+*w* Models** | | | | | | | | | | |
| 13 | DEC_TS_*w* | TS | 3 | 0.026 | 0.000 | 0.000 | 1.114 | -1204.980 | 2415.960 | 46.738 | 0.000 |
| 14 | DIVALIKE_ TS_*w* | TS | 3 | 0.032 | 0.000 | 0.000 | 1.201 | -1250.615 | 2507.231 | 138.008 | 0.000 |
| 15 | BayAreaLIKE_ TS_*w* | TS | 3 | 0.026 | 0.045 | 0.000 | 1.307 | -1425.518 | 2857.036 | 487.813 | 0.000 |
| | **+*j* +*w* Models** | | | | | | | | | | |
| ***16*** | ***DEC_TS_ j_w*** | ***TS*** | ***4*** | ***0.029*** | ***0.000*** | ***0.009*** | ***1.887*** | ***-1180.611*** | ***2369.223*** | ***0.000*** | ***0.999*** |
| 17 | DIVALIKE_TS_*j_w* | TS | 4 | 0.027 | 0.000 | 0.007 | 1.344 | -1221.864 | 2451.729 | 82.506 | 0.000 |
| 18 | BayAreaLIKE_TS_*j_w* | TS | 4 | 0.016 | 0.001 | 0.021 | 1.117 | -1210.247 | 2428.494 | 59.271 | 0.000 |

# Global Solanaceae:
# distance, Biogeographical Stochas3c Mapping

| | Models | Dispersal multipliers | Free parameters | | | | | Log-Likelihood | AIC | Δ AIC | AIC weights |
|---|---|---|---|---|---|---|---|---|---|---|---|
| # | Basic Models | | Number | $d^{(1)}$ | $e^{(2)}$ | $j^{(3)}$ | $w^{(4)}$ | | | | |
| 1 | DEC_NonTS | Non-TS | 2 | 0.011 | 0.000 | 0.000 | 1.000 | -1341.413 | 2686.826 | 317.604 | 0.000 |
| 2 | DIVALIKE_NonTS | Non-TS | 2 | 0.013 | 0.000 | 0.000 | 1.000 | -1401.936 | 2807.871 | 438.648 | 0.000 |
| 3 | BayAreaLIKE_NonTS | Non-TS | 2 | 0.009 | 0.049 | 0.000 | 1.000 | -1528.828 | 3061.656 | 692.433 | 0.000 |
| | **Time-Stratified Models** | | | | | | | | | | |
| 4 | DEC_TS | TS | 2 | 0.025 | 0.000 | 0.000 | 1.000 | -1210.832 | 2425.663 | 56.441 | 0.000 |
| 5 | DIVALIKE_TS | TS | 2 | 0.029 | 0.000 | 0.000 | 1.000 | -1252.831 | 2509.661 | 140.438 | 0.000 |
| 6 | BayAreaLIKE_TS | TS | 2 | 0.023 | 0.044 | 0.000 | 1.000 | -1430.614 | 2865.227 | 496.004 | 0.000 |
| | **+*j* Models** | | | | | | | | | | |
| 7 | DEC_ NonTS _*j* | Non-TS | 3 | 0.010 | 0.000 | 0.003 | 1.000 | -1324.525 | 2655.050 | 285.828 | 0.000 |
| 8 | DIVALIKE_NonTS_*j* | Non-TS | 3 | 0.011 | 0.000 | 0.003 | 1.000 | -1387.768 | 2781.536 | 412.313 | 0.000 |
| 9 | BayAreaLIKE_NonTS_*j* | Non-TS | 3 | 0.007 | 0.004 | 0.009 | 1.000 | -1335.728 | 2677.457 | 308.234 | 0.000 |
| 10 | DEC_TS_*j* | TS | 3 | 0.022 | 0.000 | 0.008 | 1.000 | -1191.907 | 2389.813 | 20.590 | 0.000 |
| 11 | DIVALIKE_ TS_*j* | TS | 3 | 0.025 | 0.000 | 0.007 | 1.000 | -1239.123 | 2484.246 | 115.024 | 0.000 |
| 12 | BayAreaLIKE_ TS_*j* | TS | 3 | 0.015 | 0.001 | 0.020 | 1.000 | -1213.992 | 2433.984 | 64.762 | 0.000 |
| | **+*w* Models** | | | | | | | | | | |
| 13 | DEC_TS_*w* | TS | 3 | 0.026 | 0.000 | 0.000 | 1.114 | -1204.980 | 2415.960 | 46.738 | 0.000 |
| 14 | DIVALIKE_ TS_*w* | TS | 3 | 0.032 | 0.000 | 0.000 | 1.201 | -1250.615 | 2507.231 | 138.008 | 0.000 |
| 15 | BayAreaLIKE_ TS_*w* | TS | 3 | 0.026 | 0.045 | 0.000 | 1.307 | -1425.518 | 2857.036 | 487.813 | 0.000 |
| | **+*j* +*w* Models** | | | | | | | | | | |
| ***16*** | ***DEC_TS_ j_w*** | ***TS*** | ***4*** | ***0.029*** | ***0.000*** | ***0.009*** | ***1.887*** | ***-1180.611*** | ***2369.223*** | ***0.000*** | ***0.999*** |
| 17 | DIVALIKE_TS_*j_w* | TS | 4 | 0.027 | 0.000 | 0.007 | 1.344 | -1221.864 | 2451.729 | 82.506 | 0.000 |
| 18 | BayAreaLIKE_TS_*j_w* | TS | 4 | 0.016 | 0.001 | 0.021 | 1.117 | -1210.247 | 2428.494 | 59.271 | 0.000 |

# Akaike Information Criterion

**"True model"**

Better fit

lower AIC = lower distance

▼ - distance from truth

AIC = estimator of relative distance from true model

higher AIC = higher distance

Worse fit

↑ - higher data log-likelihood (LnL)

↑ - fewer free parameters

**Akaike Information Criterion (AIC) =**
-2(LnL - # of parameters)

↓ - lower data log-likelihood (LnL)

↓ - more free parameters

# Global Solanaceae:
# distance, Biogeographical Stochastic Mapping

| | Models | Dispersal multipliers | Free parameters | | | | | Log-Likelihood | AIC | Δ AIC | AIC weights |
|---|---|---|---|---|---|---|---|---|---|---|---|
| # | Basic Models | | Number | $d^{(1)}$ | $e^{(2)}$ | $j^{(3)}$ | $w^{(4)}$ | | | | |
| 1 | DEC_NonTS | Non-TS | 2 | 0.011 | 0.000 | 0.000 | 1.000 | -1341.413 | 2686.826 | 317.604 | 0.000 |
| 2 | DIVALIKE_NonTS | Non-TS | 2 | 0.013 | 0.000 | 0.000 | 1.000 | -1401.936 | 2807.871 | 438.648 | 0.000 |
| 3 | BayAreaLIKE_NonTS | Non-TS | 2 | 0.009 | 0.049 | 0.000 | 1.000 | -1528.828 | 3061.656 | 692.433 | 0.000 |
| | **Time-Stratified Models** | | | | | | | | | | |
| 4 | DEC_TS | TS | 2 | 0.025 | 0.000 | 0.000 | 1.000 | -1210.832 | 2425.663 | 56.441 | 0.000 |
| 5 | DIVALIKE_TS | TS | 2 | 0.029 | 0.000 | 0.000 | 1.000 | -1252.831 | 2509.661 | 140.438 | 0.000 |
| 6 | BayAreaLIKE_TS | TS | 2 | 0.023 | 0.044 | 0.000 | 1.000 | -1430.614 | 2865.227 | 496.004 | 0.000 |
| | **+*j* Models** | | | | | | | | | | |
| 7 | DEC_ NonTS _*j* | Non-TS | 3 | 0.010 | 0.000 | 0.003 | 1.000 | -1324.525 | 2655.050 | 285.828 | 0.000 |
| 8 | DIVALIKE_NonTS_*j* | Non-TS | 3 | 0.011 | 0.000 | 0.003 | 1.000 | -1387.768 | 2781.536 | 412.313 | 0.000 |
| 9 | BayAreaLIKE_NonTS_*j* | Non-TS | 3 | 0.007 | 0.004 | 0.009 | 1.000 | -1335.728 | 2677.457 | 308.234 | 0.000 |
| 10 | DEC_TS_*j* | TS | 3 | 0.022 | 0.000 | 0.008 | 1.000 | -1191.907 | 2389.813 | 20.590 | 0.000 |
| 11 | DIVALIKE_ TS_*j* | TS | 3 | 0.025 | 0.000 | 0.007 | 1.000 | -1239.123 | 2484.246 | 115.024 | 0.000 |
| 12 | BayAreaLIKE_ TS_*j* | TS | 3 | 0.015 | 0.001 | 0.020 | 1.000 | -1213.992 | 2433.984 | 64.762 | 0.000 |
| | **+*w* Models** | | | | | | | | | | |
| 13 | DEC_TS_*w* | TS | 3 | 0.026 | 0.000 | 0.000 | 1.114 | -1204.980 | 2415.960 | 46.738 | 0.000 |
| 14 | DIVALIKE_ TS_*w* | TS | 3 | 0.032 | 0.000 | 0.000 | 1.201 | -1250.615 | 2507.231 | 138.008 | 0.000 |
| 15 | BayAreaLIKE_ TS_*w* | TS | 3 | 0.026 | 0.045 | 0.000 | 1.307 | -1425.518 | 2857.036 | 487.813 | 0.000 |
| | **+*j* +*w* Models** | | | | | | | | | | |
| ***16*** | ***DEC_TS_ j_w*** | ***TS*** | ***4*** | ***0.029*** | ***0.000*** | ***0.009*** | ***1.887*** | ***-1180.611*** | ***2369.223*** | ***0.000*** | ***0.999*** |
| 17 | DIVALIKE_TS_*j_w* | TS | 4 | 0.027 | 0.000 | 0.007 | 1.344 | -1221.864 | 2451.729 | 82.506 | 0.000 |
| 18 | BayAreaLIKE_TS_*j_w* | TS | 4 | 0.016 | 0.001 | 0.021 | 1.117 | -1210.247 | 2428.494 | 59.271 | 0.000 |

# Global Solanaceae:
# distance, Biogeographical Stochastic Mapping

| # | Models | Dispersal multipliers | Free parameters | | | | | Log-Likelihood | AIC | Δ AIC | AIC weights |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Number | $d^{(1)}$ | $e^{(2)}$ | $j^{(3)}$ | $w^{(4)}$ | | | | |
| | **Basic Models** | | | | | | | | | | |
| 1 | DEC_M | | | | | | | | | 317.604 | 0.000 |
| 2 | DIVALIKE_M | | | | | | | | | 438.648 | 0.000 |
| 3 | BayAreaLIKE_M | | | | | | | | | 692.433 | 0.000 |
| | **Time-S** | | | | | | | | | | |
| 4 | DEC_TS | | | | | | | | | 56.441 | 0.000 |
| 5 | DIVALIKE_TS | | | | | | | | | 140.438 | 0.000 |
| 6 | BayAreaLIKE_TS | | | | | | | | | 496.004 | 0.000 |
| | ***+j* Models** | | | | | | | | | | |
| 7 | DEC_TS_*j* | | | | | | | | | 285.828 | 0.000 |
| 8 | DIVALIKE_TS_*j* | | | | | | | | | 412.313 | 0.000 |
| 9 | BayAreaLIKE_TS_*j* | | | | | | | | | 308.234 | 0.000 |
| 10 | DEC_TS_*w* | | | | | | | | | 20.590 | 0.000 |
| 11 | DIVALIKE_TS_*w* | | | | | | | | | 115.024 | 0.000 |
| 12 | BayAreaLIKE_TS_*w* | | | | | | | | | 64.762 | 0.000 |
| | ***+w* Models** | | | | | | | | | | |
| 13 | DEC_TS_*w* | | | | | | | | | 46.738 | 0.000 |
| 14 | DIVALIKE_TS_*w* | | | | | | | | | 138.008 | 0.000 |
| 15 | BayAreaLIKE_ TS_*w* | TS | 3 | 0.026 | 0.045 | 0.000 | 1.307 | -1425.518 | 2857.036 | 487.813 | 0.000 |
| | ***+j +w* Models** | | | | | | | | | | |
| ***16*** | ***DEC_TS_ j_w*** | ***TS*** | ***4*** | ***0.029*** | ***0.000*** | ***0.009*** | ***1.887*** | ***-1180.611*** | ***2369.223*** | ***0.000*** | ***0.999*** |
| 17 | DIVALIKE_TS_*j_w* | TS | 4 | 0.027 | 0.000 | 0.007 | 1.344 | -1221.864 | 2451.729 | 82.506 | 0.000 |
| 18 | BayAreaLIKE_TS_*j_w* | TS | 4 | 0.016 | 0.001 | 0.021 | 1.117 | -1210.247 | 2428.494 | 59.271 | 0.000 |

**Best model: 99.9% of AIC weight**

**DEC**

**+ J (jump dispersal)**

**+ changing geography (distances)**

**+ parameter weigh3 ng distance**