

Using network analysis to trace the evolution of biogeography through geologic time: A case study

Steffen Kiel*

Swedish Museum of Natural History, Department of Palaeobiology, Box 50007, 10405 Stockholm, Sweden

ABSTRACT

The biogeographic distribution of organisms has continuously changed through Earth's history as plate tectonics changed the configurations of land masses, ocean basins, and climate zones. Yet, methods to investigate this dynamic through geologic time are limited. Here, network analysis is used to explore and to visualize the biogeographic history of brachiopods through the entire Triassic period. Many previously recognized biogeographic provinces are found, and in addition, the stratigraphic ranges of these provinces were identified. Provinces in the Tethys Ocean show the lowest degree of connectedness, which can be linked to higher evolutionary rates in this tropical ocean basin and possibly also to higher habitat heterogeneity. Stratigraphically, the Tethyan provinces are separated largely along the boundaries of the Early, Middle, and Late Triassic. This suggests that the events resulting in faunal changes among the index fossils used to define these sub-periods also affected the brachiopods. However, through the ~50 m.y. of the Triassic period, geographic proximity played a greater role in producing faunal similarity than proximity in geologic age. Thus network analysis is a viable tool to better understand the dynamic evolution of biogeography through geologic time.

INTRODUCTION

Many systems of interacting units in the biological, social, and technological world can be seen as complex networks. Obvious examples include food webs, neural and social networks, the Internet, power grids, and many more (Albert et al., 1999; Dunne et al., 2002; Newman and Park, 2002; Greicius et al., 2003; Pagan and Aiello, 2013). Understanding the nature of these networks is crucial for a broad range of applications, like assessing the vulnerability of critical infrastructures to failure, or predicting and preventing the spread of epidemics (Rinaldi et al., 2001; Pastor-Satorras et al., 2015). Of particular interest are methods to detect large-scale structures such as communities and modules in networks (Newman, 2012), and these approaches have recently begun to be used to address biogeographic questions (Moalic et al., 2012; Kougioumoutzis et al., 2014; Kiel, 2016).

Present-day biogeographic distributions have a historical basis, and the fossil record is a key to understand both this historical basis and the long-term evolution of biogeography (Jablonski et al., 2006; Renema et al., 2008). Biogeographic analyses through geologic time often use presence-absence matrices and are typically done in a step-wise fashion, whereby each geologic time interval is analyzed separately and this succession is then described in a narrative way (Brayard et al., 2007; Neubauer et al., 2015; Dunhill et al., 2016). But this approach does not fully capture the dynamics in biogeographic evolution. Here, occurrence

data of the phylum Brachiopoda from the entire Triassic period (~50 m.y.) are analyzed as a single complex network. The aims are to (1) delimit biogeographic provinces and their range through time, (2) evaluate the relative importance of time and geography in delimiting those provinces, and (3) assess regional differences in connectivity.

DATA

The basis is a recently published data set of late Permian to Late Triassic brachiopod occurrences (Ke et al., 2016). This data set was compiled from published sources, covers all brachiopod clades, and includes only taxa that had been systematically described and illustrated; the taxonomic unit used herein is the genus. The geographic unit used in the original study was called a "paleogeographic analysis unit" and represents a set of collections from a certain region delimited by geographic or tectonic borders (Ke et al., 2016); information on the sampled facies were not given. These units span 1.9°–4.6° of longitude on average, and values >12° (up to 41°) were only reached in very high latitudes (i.e., northern Siberia); in latitude, they span 1°–2.5° on average, with a maximum extension of 10° (the Qiangtang block during the Carnian). These units are also used herein and are referred to as localities or nodes. Excluded from the original data set were the Permian records and the records of Dienerian age (late Induan) because of the very low number of records, and nodes with fewer than three genera.

For a subsequent assessment of the impact of evolutionary rates on the longevity of the

biogeographic provinces, I determined for each node how many of its genera have their first appearance (FA) in the geologic stage to which the respective node belongs. Permian occurrences were added back into the data set when counting the number of first occurrences to avoid inflated origination in the oldest Triassic stage. To analyze the relationship between faunal similarity between two nodes on the one hand, and their geographic and temporal distance on the other, the great-circle distance between the paleo-coordinates of any two nodes was calculated using the haversine formula (the paleo-coordinates were originally calculated using PointTrack version 7.0 [Ke et al., 2016]). The time difference between any two nodes was calculated for the midpoints of their respective geologic stages, using the International Commission on Stratigraphy International Chronostratigraphic Chart version 2016/12 (Cohen et al., 2013).

NETWORK ANALYSIS

The present approach uses weighted networks, in which the localities represent the nodes, and the links between them indicate not only the presence of a connection, but are assigned a value (the weight) indicating the strength of the connection. These weights were calculated using the Bray-Curtis distance measure. They range from zero when the two nodes are identical in faunal composition, to 1 when they share no taxa, resulting in the counterintuitive situation that a link with a high weight indicates low similarity between the nodes, and vice versa.

A thresholding approach was used to illustrate and quantitatively analyze relationships between nodes (Moalic et al., 2012; Kivelä et al., 2015; Kiel, 2016). The threshold is the maximum weight for the links used to construct the network; links beyond that threshold (indicating lower similarity) are not considered (Kivelä et al., 2015). The result is that at low thresholds, the network is broken up into components, which have strong links within each component but only weak links to the remaining network. This pattern is characteristic for communities within networks (Newman, 2012), and the components were therefore used as the foundation to define paleobiogeographic provinces with both a spatial and a temporal dimension. With increasing thresholds, i.e., with the addition of

*E-mail: steffen.kiel@nrm.se

successively weaker links, the components will “grow” by the addition of further nodes and will merge with other components. The thresholding approach thus allows looking at how patterns change along a sliding scale of threshold values, and networks at different thresholds can provide different insights into biogeographic relationships (i.e., Fig. 1). When all components merge into one giant component, the so-called percolation point is reached (Kivelä et al., 2015), and networks at thresholds beyond this point are typically of limited value for delimiting paleobiogeographic provinces.

Construction and most analyses of the networks were performed using the software package EDENetworks version 2.18 (Kivelä et al.,

2015); additional analyses were performed using PAST version 2.17c (Hammer et al., 2001).

RESULTS AND DISCUSSION

Paleobiogeographic Provinces

The biogeographic provinces derived from the thresholding approach (Fig. 1) are listed in Table 1, and many of them correspond to previously recognized provinces (Ke et al., 2016). In addition, the network analysis retrieved their durations in geologic time as described in the original study (Ke et al., 2016). This is regarded here as “proof of concept” that network analysis is a viable tool to investigate biogeography through geologic time. Figure 2 shows the

network plotted onto a series of paleogeographic maps. It illustrates the distinctiveness and longevity of the Austrazea province compared to all other provinces, and shows that the main breaks in time between the Tethyan provinces are along the boundaries of the Early, Middle, and Late Triassic. These boundaries were originally defined based on ammonites and microfossils from low latitudes (Visscher, 1992; Kozur, 2003). Finding this zonation also among Tethyan brachiopod provinces suggests that the mechanisms resulting in faunal changes among ammonites and the microfossils also affected the brachiopods.

Impact of Geography and Geologic Age on Faunal Similarity

Adding stratigraphic duration to the properties of biogeographic provinces raises the question of whether geography or geologic age is the main force shaping paleobiogeographic pattern. To address this question, the impacts of each factor were first analyzed individually, and then relative to each other. Regarding age, there is a strong positive correlation between the weights of the links and their geologic ranges (ordinary least-square test, $p < 0.0001$; Fig. 3, left panel). In other words, the closer the localities are in age, the more similar they are in faunal composition. Regarding geography, there is a strong positive correlation between the weights of the links and the geographic distance of their nodes ($p < 0.0001$; Fig. 3, right panel). Thus the closer two nodes are geographically, the greater their faunal similarity. Because geographic distance was calculated as great-circle distance and thus without taking geographic barriers such as land bridges into account, a second comparison was made using ocean basins to define geographic proximity. Mean and median weights of links between nodes from the same ocean basin are in almost all cases lower than those of links between nodes from different ocean basins (Table 2), confirming the result from the comparison to geographic distance.

In summary, both geography and age influenced faunal similarity among Triassic brachiopods, so which factor had a greater impact? I compared the fractions of links “within ocean basins” to those of links “within geologic stages”. Because different numbers of ocean basins and geologic stages are compared (five ocean basins versus seven geologic stages), these fractions need to be compared to some expected values. The expected values are derived from the fully connected network (a hypothetical version of the network in which all nodes are connected to all others) and represent the fractions of links “within ocean basins” and “within geologic stages” in this fully connected network. In the actual networks, there are consistently more links “within ocean basins” compared to the expected value than there are links “within geologic stages” (Table DR1 in the GSA Data

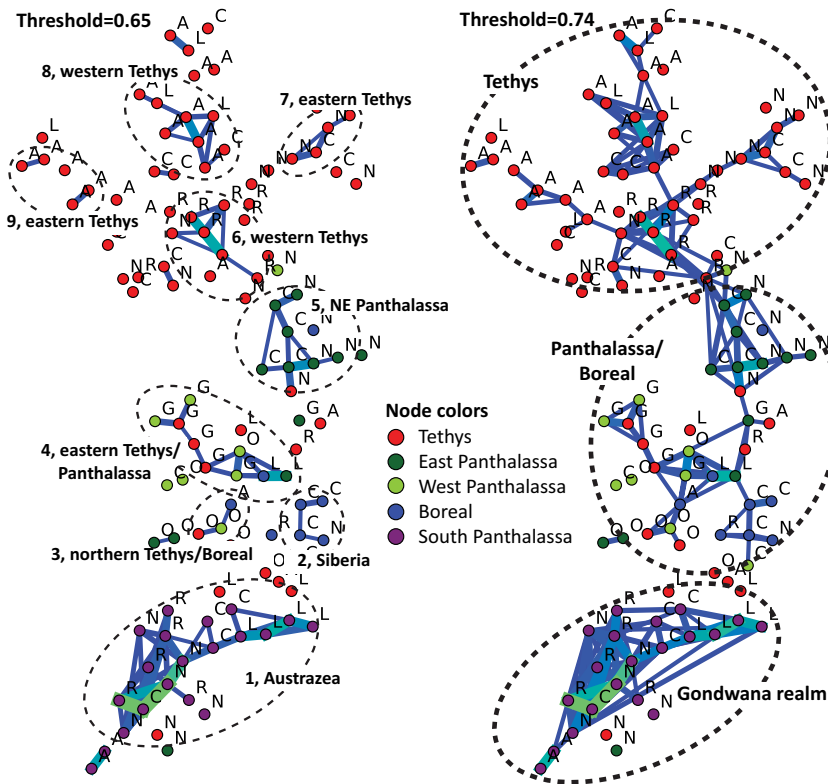


Figure 1. Triassic brachiopod paleobiogeographic network at two different thresholds. Color coding of nodes is by ocean basin; thickness and color coding of links is by weight (thicker and lighter = higher faunal similarity); numbers correspond to those in Table 1. Node letters indicate their geologic age: G—Griesbachian; O—Olenekian; A—Anisian; L—Ladinian; C—Carnian; N—Norian; R—Rhaetian.

TABLE 1. PALEOBIOGEOGRAPHIC PROVINCES FOR TRIASSIC BRACHIOPODS

Province	Age	Province in Ke et al. (2016)
1. Austrazea	Ladinian to Rhaetian	Austrazean
2. Siberia	Carnian to Norian	Siberian-Greenland
3. Northern Tethys (Olen.), Siberia (Anisian)	Olenekian to Anisian	Sino-Mongolian-Japanese?
4. Eastern Tethys, Panthalassa, Siberia	Early to Middle Triassic	—
5. Northeastern Panthalassa	Carnian to Norian	Grandian
6. Western Tethys	Norian to Rhaetian	Western Tethyan
7. Eastern Tethys	Carnian to Norian	Cathaysian
8. Western Tethys	Anisian to Carnian	Western Tethyan
9. Eastern Tethys	Anisian	Cathaysian

Note: Numbers correspond to those in Figure 1. Olen.—Olenekian.

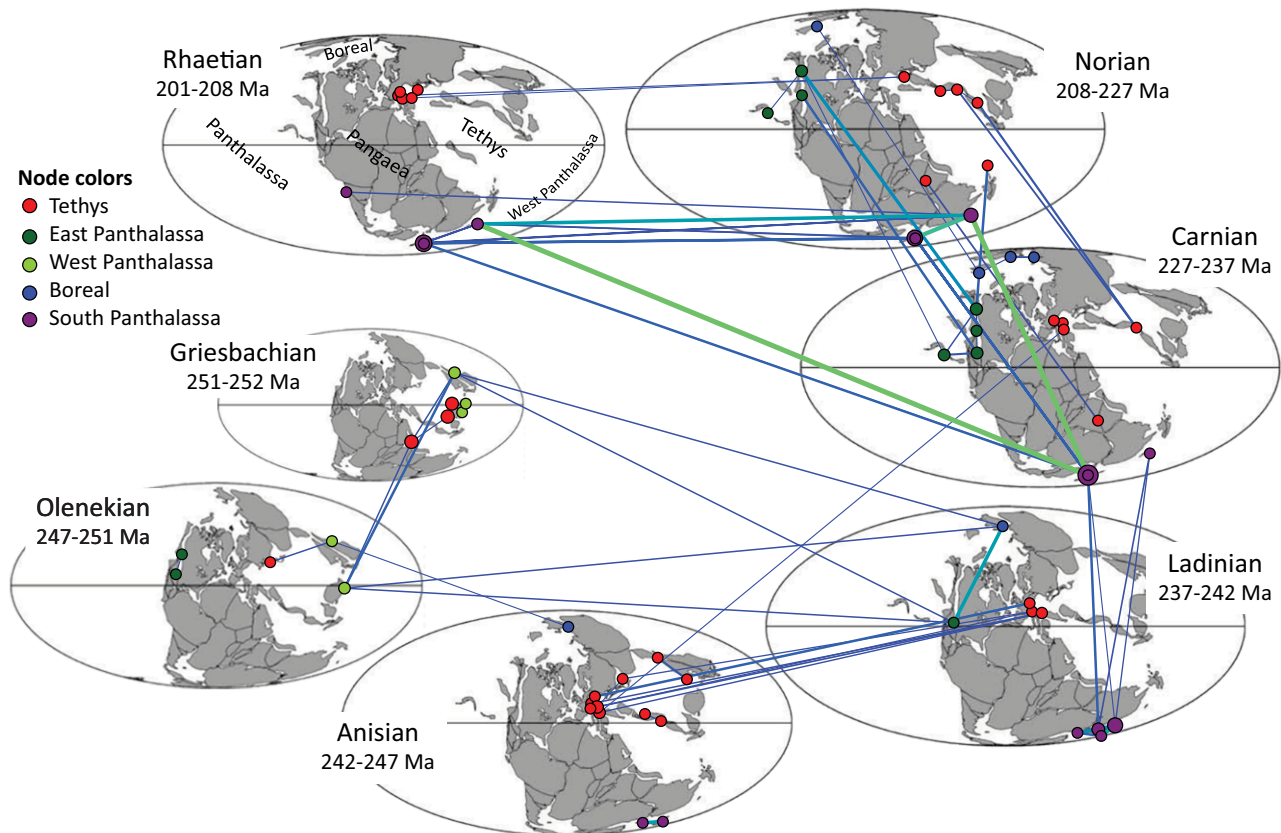


Figure 2. Triassic brachiopod network thresholded at 0.65 plotted on a series of paleogeographic maps (adapted from Scotese, 2004); node and line colors as in Figure 1.

Repository¹). This indicates that geographic proximity has a stronger impact on faunal similarity than proximity in geologic age. Network topology supports this conclusion. At thresholds around the percolation point, three geographic realms are evident: (1) Gondwana realm,

(2) northern Panthalassa and the Boreal Ocean, and (3) the tropical Tethyan realm (Fig. 1). If geologic age had a stronger impact on faunal similarity, then the major divisions at these high thresholds should have been along stratigraphic boundaries.

of connectedness, whereas those in the Tethys Ocean have the lowest (Table 2). This indicates that localities in the Tethys Ocean are the most heterogeneous in faunal composition. This pattern could have various causes, including sampling biases, varying evolutionary rates, habitat heterogeneity, or other physical or biological heterogeneities. Higher evolutionary rates in the

¹GSA Data Repository item 2017229, Table DR1 (summaries of the links by geologic stage and ocean basin) and Table DR2 (correlations between sampling intensity and latitude), is available online at <http://www.geosociety.org/datarepository/2017/> or on request from editing@geosociety.org.

Regional Differences in Connectedness

At a threshold of 0.99, the average weights of links within and across ocean basins show that localities in the southern Panthalassic Ocean have by far the highest average degree

TABLE 2. AVERAGE WEIGHTS OF LINKS IN THE TRIASSIC BRACHIOPOD PALEOBIOGEOGRAPHIC NETWORK WITHIN AND BETWEEN OCEAN BASINS

Ocean basin(s)	Mean / median
SP	0.74 / 0.78
Boreal	0.82 / 0.83
EP	0.82 / 0.86
WP	0.86 / 0.88
EP-WP	0.87 / 0.89
Boreal-EP	0.89 / 0.90
SP-WP	0.90 / 0.91
Boreal-WP	0.87 / 0.91
Tethys	0.89 / 0.91
SP-EP	0.90 / 0.91
Tethys-EP	0.91 / 0.93
Boreal-SP	0.93 / 0.93
Tethys-WP	0.91 / 0.94
Tethys-Boreal	0.93 / 0.94
Tethys-SP	0.94 / 0.95

Note: Values for the network at a threshold of 0.99. SP= southern Panthalassa; EP—eastern Panthalassa; WP—western Panthalassa.

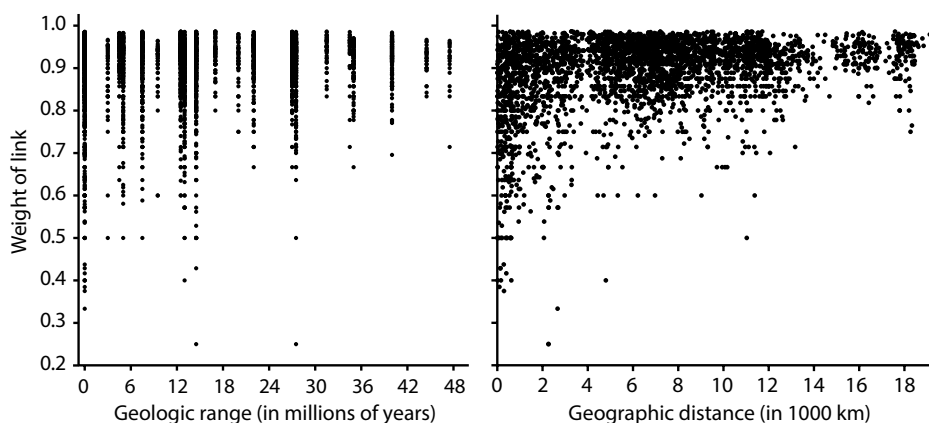


Figure 3. Relationships between weights of links in Triassic brachiopod network and their geologic ranges (left panel) and geographic distance they span (right panel), at threshold of 0.99.

tropics (Jablonski et al., 2006; Kiessling et al., 2010) could result in lower biogeographic connectedness simply because newly evolved taxa are unlikely to have a wide geographic distribution. The average numbers of FAs per ocean basin are in the opposite order as the average weights of the links within those basins: Tethyan sites have the highest average number of FAs, whereas the southern Panthalassic sites have the lowest (Tethys = 6.03; western Panthalassa = 5.33; eastern Panthalassa = 2.85; Boreal = 1.75; southern Panthalassa = 1.6). This indicates that higher evolutionary rates indeed result in lower biogeographic connectedness.

The tropical Tethys Ocean could potentially have had a higher habitat diversity than the other ocean basins: reefs and carbonate platforms were common in the Tethys Ocean, but were largely absent from the other regions (Martindale et al., 2015). Rigorously testing this hypothesis might be possible by compiling data on habitat types at the investigated localities, but this was beyond the scope of the present paper. Uneven sampling can potentially distort the results of any quantitative approach to biogeography, and network analysis is no exception to this. I tested the data set for latitudinal sampling disparities that could be inflating origination rates in the Tethys Ocean relative to other regions. The nodes showed no correlation between their latitude and the number of individual sites per node (Table DR2). Thus sampling bias is unlikely to have affected the regional differences in biogeographic connectedness.

CONCLUSIONS

Network analysis was used here to investigate the evolution of Triassic brachiopod biogeographic provinces through geologic time. The method retrieved many previously recognized paleobiogeographic provinces; in addition, it (1) recognized the stratigraphic ranges of these provinces, which had only been outlined in a narrative way in the original study (Ke et al., 2016), and (2) allowed an intuitive visualization of the results. Furthermore, the approach allowed the detection and quantification of regional differences in duration and connectedness of the paleobiogeographic provinces, showing that provinces in the tropical Tethys Ocean were short lived and had low connectedness, whereas the highest connectedness and longevity were found in the southern Panthalassic ocean. This pattern is linked to evolutionary rates: FAs of genera were highest in the Tethys and tropical western Panthalassa, and lowest in high latitudes.

Perhaps the most remarkable property of the Triassic brachiopod paleobiogeographic network is that through ~50 m.y. of time, geographic proximity had a stronger impact on faunal similarity than geologic age. It would be interesting to investigate (1) how much geologic time it takes until geologic age plays a greater role (if at all), (2) which role mass extinction

events might play in this context, and (3) which effect the taxonomic level (species, genus, family) might have on these patterns.

The strong impact of geographic proximity on faunal similarity emphasizes that present-day biogeographic distributions also have a strong historical basis, and that investigating this historical basis will lead to a better understanding of the biogeography of the world today (Sanmartín and Ronquist, 2004; Krug et al., 2010). Other future research directions in network paleobiogeography include the exploration of directed networks (Leicht and Newman, 2008), modeling of ideal scenarios and comparing them to real world examples, assessing the impact of sampling biases, and many more. I hope that the approach outlined here will be explored further and will result in a better understanding of the dynamic evolution of biogeography through geologic time.

ACKNOWLEDGMENTS

I thank Michael Hautmann (Zürich) and two anonymous reviewers for their insightful comments that greatly improved this manuscript.

REFERENCES CITED

- Albert, R., Jeong, H., and Barabási, A.-L., 1999, Internet: Diameter of the World-Wide Web: *Nature*, v. 401, p. 130–131, doi:10.1038/43601.
- Brayard, A., Escarguel, G., and Bucher, H., 2007, The biogeography of Early Triassic ammonoid faunas: Clusters, gradients, and networks: *Geobios*, v. 40, p. 749–765, doi:10.1016/j.geobios.2007.06.002.
- Cohen, K.M., Finney, S.C., Gibbard, P.L., and Fan, J.-X., 2013, The ICS International Chronostratigraphic Chart: *Episodes*, v. 36, p. 199–204.
- Dunhill, A.M., Bestwick, J., Narey, H., and Sciberras, J., 2016, Dinosaur biogeographical structure and Mesozoic continental fragmentation: A network-based approach: *Journal of Biogeography*, v. 43, p. 1691–1704, doi:10.1111/jbi.12766.
- Dunne, J.A., Williams, R.J., and Martinez, N.D., 2002, Food-web structure and network theory: The role of connectance and size: *Proceedings of the National Academy of Sciences of the United States of America*, v. 99, p. 12,917–12,922, doi:10.1073/pnas.192407699.
- Greicius, M.D., Krasnow, B., Reiss, A.L., and Menon, V., 2003, Functional connectivity in the resting brain: A network analysis of the default mode hypothesis: *Proceedings of the National Academy of Sciences of the United States of America*, v. 100, p. 253–258, doi:10.1073/pnas.0135058100.
- Hammer, Ø., Harper, D.A.T., and Ryan, P.D., 2001, PAST: Palaeontological statistics software package for education and data analysis: *Palaeontologia Electronica*, v. 4, no. 1, 9 p.
- Jablonski, D., Roy, K., and Valentine, D.L., 2006, Out of the tropics: Evolutionary dynamics of the latitudinal diversity gradient: *Science*, v. 314, p. 102–106, doi:10.1126/science.1130880.
- Ke, Y., Shen, S., Shi, G.R., Fan, J.-x., Zhang, H., Qiao, L., and Zeng, Y., 2016, Global brachiopod paleobiogeographical evolution from Changhsingian (Late Permian) to Rhaethian (Late Triassic): *Palaeogeography, Palaeoclimatology, Palaeoecology*, v. 448, p. 4–25, doi:10.1016/j.palaeo.2015.09.049.
- Kiel, S., 2016, A biogeographic network reveals evolutionary links between deep-sea hydrothermal vent and methane seep faunas: *Proceedings of the Royal Society of London B*, v. 283, 20162337, doi:10.1098/rspb.2016.2337.
- Kiessling, W., Simpson, C., and Foote, M., 2010, Reefs as cradles of evolution and sources of

- biodiversity in the Phanerozoic: *Science*, v. 327, p. 196–198, doi:10.1126/science.1182241.
- Kivelä, M., Arnaud-Haond, S., and Saramäki, J., 2015, EDENetworks: A user-friendly software to build and analyse networks in biogeography, ecology and population genetics: *Molecular Ecology Resources*, v. 15, p. 117–122, doi:10.1111/1755-0998.12290.
- Kougioumoutzis, K., Simaiakis, S.M., and Tiniakou, A., 2014, Network biogeographical analysis of the central Aegean archipelago: *Journal of Biogeography*, v. 41, p. 1848–1858, doi:10.1111/jbi.12342.
- Kozur, H.W., 2003, Integrated ammonoid, conodont and radiolarian zonation of the Triassic: *Hallesches Jahrbuch für Geowissenschaften*, v. B25, p. 49–79.
- Krug, A.Z., Jablonski, D., Roy, K., and Beu, A.G., 2010, Differential extinction and the contrasting structure of polar marine faunas: *PLoS One*, v. 5, e15362, doi:10.1371/journal.pone.0015362.
- Leicht, E.A., and Newman, M.E.J., 2008, Community structure in directed networks: *Physical Review Letters*, v. 100, 118703, doi:10.1103/PhysRevLett.100.118703.
- Martindale, R.C., Corsetti, F.A., James, N.P., and Botter, D.J., 2015, Paleogeographic trends in Late Triassic reef ecology from northeastern Panthalassa: *Earth-Science Reviews*, v. 142, p. 18–37, doi:10.1016/j.earscirev.2014.12.004.
- Moalic, Y., Desbruyères, D., Duarte, C.M., Rozenfeld, A.F., Bachraty, C., and Arnaud-Haond, S., 2012, Biogeography revisited with network theory: Retracing the history of hydrothermal vent communities: *Systematic Biology*, v. 61, p. 127–137, doi:10.1093/sysbio/syr088.
- Neubauer, T.A., Harzhauser, M., Kroh, A., Georgopoulou, E., and Mandic, O., 2015, A gastropod-based biogeographic scheme for the European Neogene freshwater systems: *Earth-Science Reviews*, v. 143, p. 98–116, doi:10.1016/j.earscirev.2015.01.010.
- Newman, M.E.J., 2012, Communities, modules and large-scale structure in networks: *Nature Physics*, v. 8, p. 25–31, doi:10.1038/nphys2162.
- Newman, M.E.J., and Park, J., 2002, Why social networks are different from other types of networks: *Physical Review E*, v. 68, 036122, doi:10.1103/PhysRevE.68.036122.
- Pagani, G.A., and Aiello, M., 2013, The Power Grid as a complex network: A survey: *Physica A*, v. 392, p. 2688–2700, doi:10.1016/j.physa.2013.01.023.
- Pastor-Satorras, R., Castellano, C., Miegheem, P.V., and Vespignani, A., 2015, Epidemic processes in complex networks: *Reviews of Modern Physics*, v. 87, p. 925, doi:10.1103/RevModPhys.87.925.
- Renema, W., et al., 2008, Hopping hotspots: Global shifts in marine biodiversity: *Science*, v. 321, p. 654–657, doi:10.1126/science.1155674.
- Rinaldi, S.M., Peerenboom, J.P., and Kelly, T.K., 2001, Identifying, understanding, and analyzing critical infrastructure interdependencies: *IEEE Control Systems*, v. 21, p. 11–25, doi:10.1109/37.969131.
- Sanmartín, I., and Ronquist, F., 2004, Southern Hemisphere biogeography inferred by event-based models: Plant versus animal patterns: *Systematic Biology*, v. 53, p. 216–243, doi:10.1080/10635150490423430.
- Scotese, C.R., 2004, A continental drift flipbook: *The Journal of Geology*, v. 112, p. 729–741, doi:10.1086/424867.
- Visscher, H., 1992, The new STS Triassic stage nomenclature: *Albertiana*, v. 10, p. 1–2.

Manuscript received 8 December 2016
 Revised manuscript received 13 April 2017
 Manuscript accepted 17 April 2017

Printed in USA