



# Bivalve network reveals latitudinal selectivity gradient at the end-Cretaceous mass extinction

SUBJECT AREAS:

MACROECOLOGY

ASTEROIDS, COMETS AND  
KUIPER BELT

PALAEONTOLOGY

ECOLOGICAL NETWORKS

Daril A. Vilhena<sup>1</sup>, Elisha B. Harris<sup>1,2</sup>, Carl T. Bergstrom<sup>1,3</sup>, Max E. Maliska<sup>1</sup>, Peter D. Ward<sup>1,2,4</sup>,  
Christian A. Sidor<sup>1,2</sup>, Caroline A. E. Strömberg<sup>1,2</sup> & Gregory P. Wilson<sup>1,2</sup>

<sup>1</sup>Department of Biology, University of Washington, Seattle, WA 98195-1800, <sup>2</sup>Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195-3010, <sup>3</sup>Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501, <sup>4</sup>Department of Earth and Space Science, University of Washington, Seattle, WA 98195-1800.

Received  
6 February 2013

Accepted  
8 April 2013

Published  
7 May 2013

Correspondence and  
requests for materials  
should be addressed to  
D.A.V. (daril@uw.edu)  
or G.P.W. (gpwilson@  
u.washington.edu)

**Biogeographic patterns of survival help constrain the causal factors responsible for mass extinction. To test whether biogeography influenced end-Cretaceous (K-Pg) extinction patterns, we used a network approach to delimit biogeographic units (BUs) above the species level in a global Maastrichtian database of 329 bivalve genera. Geographic range is thought to buffer taxa from extinction, but the number of BUs a taxon occurred in superseded geographic range as an extinction predictor. Geographically, we found a latitudinal selectivity gradient for geographic range in the K-Pg, such that higher latitude BUs had lower extinction than expected given the geographic ranges of the genera, implying that (i) high latitude BUs were more resistant to extinction, (ii) the intensity of the K-Pg kill mechanism declined with distance from the tropics, or (iii) both. Our results highlight the importance of macroecological structure in constraining causal mechanisms of extinction and estimating extinction risk of taxa.**

Mass extinctions have disproportionately shaped the evolutionary history of life<sup>1</sup>. During these rare, geologically rapid events, the rules of selectivity that prevail in background extinctions do not always clearly apply<sup>2</sup>. To delimit what may be independent selection processes and constrain possible causal factors, paleontologists have sought biologically meaningful patterns of survivorship in mass extinctions<sup>3,4</sup>. In the marine realm, there are some taxon-specific cases where survivorship is linked to ecological traits – for example, in the K-Pg mass extinction event, reliance on photosymbiosis among scleractinian corals severely reduced survivorship<sup>5</sup> and sea urchin feeding strategy correlates positively with survivorship<sup>6</sup>. However, more often, survivorship in mass extinctions appears to be linked to increased geographic range size above the species level<sup>7</sup>, suggesting that biogeographic history of a taxon plays a vital role in sorting survivors from victims.

The correlation of geographic range at the lineage level with lineage survival, but apparent lack of physiological mechanisms in determining survivorship, suggests that lineages with similar biogeographic histories should have similar chances to survive mass extinction. Histories of ocean surface currents, plate tectonics, and environments shape biogeographic patterns above the species level by allowing lineages (genera) dispersal opportunities. Here, we explore whether these emergent biogeographic regions impact our understanding of extinction processes.

Here we employ a network approach to test whether biogeographic structure above the species level<sup>8–11</sup> correlates with bivalve survivorship in the K-Pg. Network methods have been useful in a broad range of applications, for example, to model the transmission of disease in social networks<sup>12</sup>, to describe the structure of scholarly communication<sup>13</sup>, and to model the stability of ecosystems in response to extinction<sup>14</sup>. A network approach can also reveal spatial patterns of taxa from geographic range data<sup>15</sup>. We generated a Maastrichtian network from the bivalve dataset, and used stratigraphic ranges from the Paleobiology Database to determine which genera survived the K-Pg mass extinction event<sup>16</sup>. We adapt a network-based clustering approach<sup>13</sup> to reveal biogeographic units (BUs) from patterns of geographic ranges (see methods).

The K-Pg event (ca. 66 Ma) is an ideal case to test whether biogeography influences survival in mass extinctions. It was geologically abrupt and was associated with significant changes to marine productivity and ocean chemistry, dramatic restructuring of marine and terrestrial communities, long-term effects on evolutionary rates and biogeography<sup>17,18</sup>, and the extinction of up to 76% of all species<sup>19</sup>. As the most recent of the “big five” mass extinctions, the quantity, quality, and spatial resolution of the geological and paleontological data for the K-Pg interval are also better than those available for more ancient mass extinctions<sup>22</sup>. Bivalves have emerged as a model



system for examining macroevolutionary phenomena due to their excellent preservational record<sup>22</sup>, deep evolutionary history, spatial ubiquity, and the significant effort dedicated to standardizing their systematics<sup>7</sup>.

The bivalve dataset that was used in this study was downloaded from the Paleobiology Database (PBDB<sup>16</sup>) and consists of 3,445 occurrences of 329 bivalve genera from 105 Maastrichtian faunal assemblages<sup>20</sup>. This taxonomically standardized and globally representative dataset is the same one that was previously used to show a geographically uniform pattern of extinction across the K-Pg boundary<sup>20</sup>.

## Results

**Biogeography above the species level.** The ten biogeographic units (BUs) we identified have boundaries that are naturally delimited by the major patterns of geographic ranges of genera, and reflect sudden biogeographic transitions in the biota (Fig. 1). The sampling intensity (number of localities per biogeographic region) do not affect BU delineations unless poorly sampled regions contain few or no endemic taxa, preventing our identification of a biogeographic transition. Australia is perhaps the best example of this – our analysis groups the single Australian locality with the mostly European non-rudist BU ( $M_2$ ). However, the Australian locality contains 17 taxa, 15 of which are found on average in five other BUs<sup>21</sup>. This suggests that, at least given the genus-level data available, Australia had cosmopolitan bivalve fauna that extended down from European shorelines. In turn, this indicates that either i) biogeographic structure above the species level has little to do with Maastrichtian Stage oceanic surface currents, and likely reflects complex ancient dispersal patterns and plate tectonics, or that ii) additional sampling in Australia, for example, and in under-sampled localities, may refine and improve the biogeographic structure we identified.

Overall, the genus-level BUs we identified are not necessarily limited to localized continental shorelines (Fig. 1). BUs  $M_1$  and  $M_2$  are the best sampled and are distributed mainly in the North American Gulf and southern Europe, respectively. BUs  $M_3$  and  $M_8$  are characterized by tropical rudist bivalves (Order Hippuritoida) and are found in the Eastern and Western Hemispheres, respectively. BU  $M_4$  is composed of several high-latitude clam families, and is distributed across South Africa, South America, New Zealand, and the Asian-Alaska land bridge. BU  $M_5$  is located in both central South America and West Africa. BU  $M_6$ , distributed along the east Asian coastline, is a provincial BU with several Asian endemics. BUs  $M_7$  and  $M_9$  are



**Figure 1 | The biogeographic structure of bivalves reveals spatial organization above the species level.** Points correspond to fossil localities, and are colored by BU. For visualization, ten-by-ten degree cells were colored by BU only if the cell contained fossil localities from a single BU, and cells without fossil localities were colored if they were less than 15 degrees from a locality. Although some cells may have been uninhabitable by marine bivalves, they were colored if they met the above criteria. Grids are only for visual aid and were not included in any way in subsequent analysis.

provincial and found along North American coastlines, but both contain several genera with large geographic ranges. BU  $M_{10}$  is small and tied to the European shoreline; it comprises a single locality with many endemics, and has the lowest extinction percentage. However,  $M_{10}$  also has an excellent preservational setting, suggesting that its difference from  $M_2$  may be partially due to sampling effects. The full distribution of genera across BUs is available in the online supplement<sup>21</sup>.

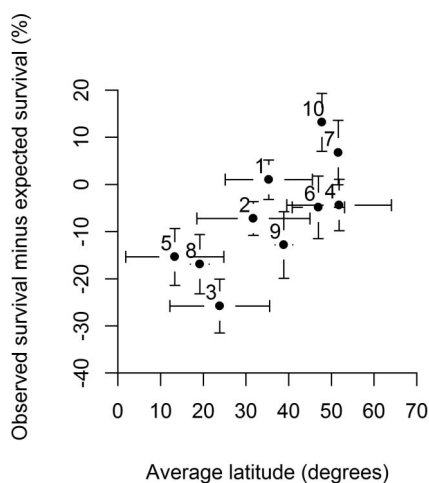
**BU range as an extinction predictor.** The demarcation of BU boundaries has predictive power for per-taxon survivorship. Previous work has indicated that geographic range is the best predictor of genus-level survivorship in both background and mass extinction<sup>7,23</sup>, and presumably buffers taxa from extinction, perhaps because it is correlated with environmental breadth<sup>24</sup>. However, we found that BU range is a better predictor than geographic range, or number of BUs a taxon is distributed in. A logistic regression with BU range alone was the best model to predict K-Pg survivorship (AIC = 419.74,  $P = 10^{-8}$ ), compared with a regression with both (AIC = 421.67,  $P = 0.78$  for geographic range,  $P = 0.015$  for BU range), and geographic range alone (AIC = 425.8,  $P = 10^{-7}$ ). This finding suggests that, at least for bivalves in the end-Cretaceous event, the number of BUs a lineage was distributed in has better predictive power for survivorship than the number of provincial shorelines a lineage was distributed in. These results are robust to the inclusion or removal (any combination) of inoceramid and rudist bivalves, the first whose extinction preceded the K-Pg boundary, and the second whose extinction might be tied to a physiological factor [Ref. 25, but see Ref. 7].

**Geographic-range latitudinal selectivity gradient.** Comparing extinction percentage across BUs requires an adjustment for the correlation of geographic range with extinction probability, because differences in distributions of geographic ranges will bias observed per-BU extinction percentage. To correct this, we estimated expected per-BU extinction percentage given the per-BU distribution of geographic ranges<sup>23</sup>, and analyzed the difference between expected per-BU extinction percentage given geographic range and observed per-BU extinction percentage. The entire assemblage of bivalves that occurred in each BU was included in the calculation of per-BU extinction unless otherwise noted.

After adjusting the per-BU extinction percentage, we examined how extinction percentage varied by BU geography. Whereas Raup and Jablonski<sup>20</sup> used nine geographic regions to infer that K-Pg extinction intensity was globally homogeneous, we used geographic regions determined by the modular structure of the data to show that BUs had different adjusted extinction percentages (Fig. 2). This adjusted extinction percentage did not vary along a paleolongitudinal gradient or with distance to either the bolide impact at Chixculub, Mexico or the Deccan flood basalt volcanism of peninsular India ( $P > 0.05$ ). However, we detected a paleolatitudinal gradient of adjusted extinction percentage, such that BUs with higher average paleolatitudes had higher survival than expected given the geographic ranges of the genera that occurred in those BUs (Fig. 2,  $P = 0.02$ ,  $R^2 = 0.495$ ). This result is robust to the inclusion or exclusion of both rudist and inoceramid bivalves, but the figure shown has inoceramids excluded. The absolute value of the paleolatitude was used to control for differences in sampling between hemispheres.

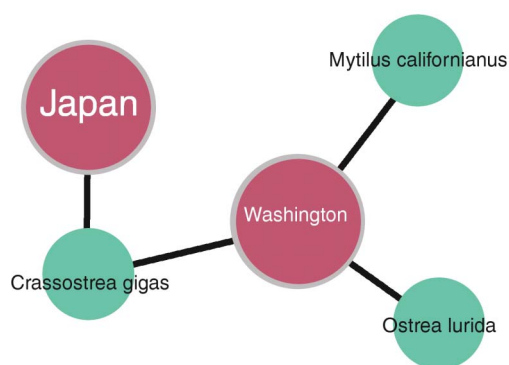
## Discussion

Our study refines the debate regarding the major causal hypotheses for the K-Pg mass extinction. Some researchers have suggested that the bolide impact alone triggered a host of secondary effects, for example, a brief but intense global thermal pulse<sup>26</sup> and a dust cloud that inhibited photosynthesis<sup>27</sup>, which together would have led to catastrophic extinction cascades. Others have contended that



**Figure 2 | Observed survivorship minus expected survivorship, under a model that predicts survivorship solely based on geographic range, correlates with latitude (Regression with unequal variances,  $P = 0.02$ , horizontal and vertical standard error shown).** These points are not independent because BUs contain some of the same genera, but a Mantel test between the Jaccard distance matrix and the extinction differences between those BUs indicates no correlation between the BU assemblages and their adjusted extinction percentages ( $P > 0.05$ ). Above zero indicates survivorship greater than expected given geographic range. BU numbers are shown, and rudist BUs are included. Average latitude of each BU was determined from the absolute latitudes of the localities in each BU. This result is robust to changes of the model parameter, global extinction risk, use of median instead of average latitude, and choice of regression analysis (equal or unequal variance). Inoceramids are excluded here, but the result is not sensitive to their exclusion. If rudists are excluded the  $P$ -value increases slightly ( $P = 0.03$ ).

additional events, such as flood basalt volcanism in India that released massive amounts of sulfur and carbon dioxide and resulted in severe environmental perturbations, combined with the bolide impact to cause the K-Pg mass extinction<sup>28,29</sup>. Our study does not reject these hypotheses, but further constrains the search space of the proximal causal agents. Our results imply that either the severity of the K-Pg kill mechanism declined with distance from the tropics, that



**Figure 3 | A bipartite occurrence network.** *Ostrea lurida*, *Mytilus californianus*, and *Crassostrea gigas* each have a second order relationship with each other (co-occurrence). Japan and Washington have a single second order relationship (shared *Crassostrea gigas*). Both *Ostrea lurida* and *Mytilus californianus* have a single third order relationship with Japan. This does not imply that *Ostrea lurida*, for example, could occur in Japan, but more third order relationships than we would expect due to chance with Japan is evidence for occurrence potential, or depauperate fossilization.

higher latitude BUs were more resistant to extinction, or both. In turn, proposed causal agents and scenarios must either match the decline of latitudinal kill mechanism severity with decreased geographic range selectivity at the BU level, or provide paleoecological evidence for decreased ecosystem extinction risk with latitude. Although neocological evidence suggests that ecosystems at higher latitudes have lower extinction risk<sup>30</sup>, presumably because of wider abiotic tolerances, we cannot be certain that this relationship applies through geologic time given differences in latitudinal diversity gradients and configurations of continents. Analyses of BU-level extinction risk in background intervals are needed to calibrate the effect of latitude on ecosystem extinction risk throughout the Phanerozoic.

Though extinction percentage in the K-Pg extinction event was geographically uniform<sup>20</sup>, our results reveal a selectivity gradient for geographic range, such that extinction vulnerability differed depending on latitude. Our analysis suggests that taxa must be studied in the context of the higher-level biogeographic structure in which they reside. In other words, the extinction vulnerability of organisms during the K-Pg was biogeographically coupled. Biogeographic selectivity, or a difference between BU vulnerability, in the K-Pg mass extinction has theoretical ramifications as well. If BUs have differential extinction, then invasives from BUs with lower extinction might displace BUs with higher extinction as ecospace opens<sup>31</sup>, effectively acting as colonizers for the BU. Given the spatial complexity of the K-Pg recovery<sup>7</sup>, the ability of BUs to displace one another through biological invasions is a possibility. Our study underscores that the macroecological context of taxa should not be ignored because extinction vulnerability is inextricably coupled to biogeographic history.

## Methods

**Paleobiology database download.** We downloaded the Maastrichtian Raup and Jablonski dataset from the Paleobiology Database on March 21, 2012. This dataset contains 3,445 occurrences of 329 bivalve genera in 105 assemblages<sup>16</sup>. These assemblages are basin level resolution from the Maastrichtian Stage. We chose this dataset because it is spatially well sampled and it is taxonomically standardized. We could not use all Maastrichtian marine invertebrates (or even bivalves) from the Paleobiology Database because the majority of the taxa are from a USGS data dump. These data were not used because they are spatially uneven (Gulf of Mexico bias) and taxonomically inconsistent with the rest of the database (causing duplicates of many genera). Additionally, we opted to use this dataset to make our results more comparable to those of Raup and Jablonski<sup>20</sup>.

**Determining survivors.** Survivors were determined from the Paleobiology Database standard stratigraphic range intervals<sup>16</sup>. We could not use the Sepkoski compendium<sup>32</sup> because it lacked range interval data for 68 genera.

**Data availability and software.** Data not immediately accessible in the Paleobiology Database are available in the public repository Dryad (<http://www.datadryad.org>). This includes the bivalve network, BU assignments, and the geographic ranges of the taxa. Shown in the supplement are the distributions of taxa across BUs, geographic ranges, and BU ranges. To infer geographic ranges, we used the province-counting approach outlined in Jablonski and Raup<sup>33</sup>. Their approach used the biogeographic provinces in the Atlas of Palaeobiogeography<sup>34</sup>. We created shape files for these provinces and used the R package *sp* to infer geographic range size.

**A network approach for biogeography.** Here, we introduce bipartite occurrence networks, which contain both localities and taxa as nodes (Fig. 3). The links in this network (connections between nodes) are occurrences. This network has convenient higher-order biological properties. The set of localities a taxon links to is its geographic range, while the number of taxa a locality links to is its richness. A pair of nodes that are two links away from each other have a “second order relationship.” For pairs of taxa, the number of second order relationships is the number of co-occurrences, while for localities, the number of second order relationships is the number of shared taxa (note that a second order relationship cannot exist between a locality and a taxon). Third order relationships in this network are between taxon and locality, and appear when a taxon is connected to a locality through an intermediary taxon and locality (Fig. 3).

Classical approaches for biogeographic analysis of occurrence data, such as ordination or agglomerative clustering of a locality-locality distance matrix, use only second order information<sup>35</sup>. Specifically, one cannot reconstruct the geographic range of a taxon from the distance matrix used for analysis. A network approach allows one to integrate geographic ranges, co-occurrence (taxon-taxon), shared taxa (locality-locality), and higher order relationships. The higher order relationships can help



biogeographic analysis recover from competitive exclusion or taphonomy – for example, aragonitic and calcitic shells may not be preserved together in entire biogeographic regions<sup>36–38</sup>. For smaller spatial scales, such as in a Pacific Northwest intertidal ecosystem, goose barnacles may not occur in the same plot as California mussels, yet we would like them grouped within the same biogeographic structure (intertidal strip). A network approach can therefore be applied to assemblage data collected at any spatial or taxonomic resolution, and has the added advantage that no dissimilarity measure is required for cluster analysis of the network<sup>35</sup>.

The taxon-locality matrix  $M$ , or occurrence matrix, is a representation of the occurrence distributions of taxa (any level) across localities (any spatial scale, from plots to counties to provinces). Each entry in the matrix is either one (taxon present in locality) or zero (absent)

$$M_{ij} = \begin{cases} 1 & \text{if taxon } j \text{ is in locality } i \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

This matrix is the basic representation of a bipartite network, where there are two types of nodes: taxa and localities (Fig. 3). The links (occurrence relationships) are exclusively between localities and taxa, taxa cannot be linked to taxa, and localities cannot be linked to localities.

Taxa will be connected to one another in complex patterns through the localities they occupy, making it difficult to interpret the information in the network without first isolating the major patterns. To find biogeographic clusters in the network that we wish to study, we must identify these patterns. As biologists, we would like to identify the boundaries between biogeographic units, which will be encoded as major topological features in the network. Moreover, we would like to capture most of the information about specific relationships with broad brushstrokes.

Network community detection is a methodological process that does exactly this: identifies the major topological features of networks<sup>39</sup>. Though many algorithms have been proposed to do this task<sup>40</sup>, the *map equation* is an excellent candidate for biogeography because of its accuracy<sup>40</sup>.

The map equation approach minimizes as an objective function the theoretical limit of a description length of a random walk on the network, where nodes are aggregated into community structures. The following story provides some intuition about this process. A scientist looks at a random locality. She then randomly chooses a taxon found at that locality. Next, she randomly selects a locality that the taxon she chose is found in. She now chooses another taxon from the new locality, and repeats this process forever. In a network with emergent biogeographic features, she will likely spend long bouts of this process within biogeographic units, which we will refer to as BUs. Specifically, she can only switch to a new biogeographic unit when she chooses a taxon that is not endemic to the biogeographic unit she is currently in. If she would like to communicate a list of the localities and taxa she chose, it would save her time to communicate a list of the major biogeographic units (which contain taxa and localities) she visited, rather than each individual locality and taxon.

Here we write the basic form of the map equation in the notation of the occurrence matrix. However, for more in-depth discussion and derivation, we refer the reader to a map equation tutorial<sup>41</sup>, or the original paper<sup>13</sup>.

This bipartite network is unweighted and undirected, all links are equal, and each link is symmetric. We refer to the total localities in the network as  $A$ , and the total number of taxa as  $T$ . In undirected networks, the probability that the scientist visits any node (locality or taxa) in her infinite random walk is the number of links that node has, divided by two times the number of links in the entire network. We refer to the number of links multiplied by 2 in the network as  $L$ , defined as

$$L = 2 \sum_{i=1}^A \sum_{j=1}^T M_{ij} \quad (2)$$

We multiply the number of links by two because each link is symmetric, so we need to count each link twice. The frequency with which she visits locality  $i$  in her random walk is therefore

$$p_A^i = \frac{\sum_{j=1}^T M_{ij}}{L}, \quad (3)$$

while the frequency with which she visits taxon  $j$  is

$$p_T^j = \frac{\sum_{i=1}^A M_{ij}}{L}. \quad (4)$$

As cartographers, we would like the scientist to convey her random walks as concisely as possible. This is an optimization problem – out of all possible partitions, we must choose the one that allows her to convey her random walk the most concisely. A partition is a division of the network into BUs, such that each node is uniquely assigned to a BU. The best partition will be the optimal compression of the major patterns in the bipartite network. To evaluate these partitions, we must express the frequency with which she switches between BUs. The frequency with which she leaves BU  $m$  is the number of links that lead from nodes in BU  $m$  to nodes outside BU  $m$ , written  $n_e^m$  and divided by the total number of links

$$p_e^m = \frac{n_e^m}{L}, \quad (5)$$

while the frequency with which she is inside BU  $m$  is

$$p_s^m = \frac{2n_s^m}{L}, \quad (6)$$

where  $n_s^m$  is the total number of links between nodes in BU  $m$ , multiplied by 2 because each link must be counted twice. These probabilities are sufficient to express the extended form of the map equation<sup>41</sup>

$$L(M) = \left( \left( \sum_m p_e^m \right) \log \left( \sum_m p_e^m \right) - 2 \sum_m (p_e^m) \log(p_e^m) + \sum_m (p_e^m + p_s^m) \log(p_e^m + p_s^m) - \left( \sum_{j=1}^T p_T^j \log p_T^j + \sum_{i=1}^A p_A^i \log p_A^i \right) \right) \quad (7)$$

where  $L(M)$  is the amount of information required to convey an infinite random walk on partition  $M$ . All logarithms listed are in base 2, because information is measured in units of bits<sup>42</sup>. The fourth term is independent of the partition  $M$ , while the first three terms change based on the proposed partition. To seek the best partition among many, we used an algorithm that is freely available online at (<http://www.mapequation.org>). An applet that illustrates the concept behind the map equation is available online (<http://www.mapequation.org>).

**Partition robustness.** The algorithm that minimizes the map equation returned largely equivalent partitions across random seeds (7.8–7.92 bits, 1.3–1.33% compression). Choice of partition did not change our results.

**Geographic range null model.** We use a modification of the approach outlined in Payne and Finnegan<sup>23</sup> to test the relative magnitude of geographic range selectivity between BUs. What follows is a “fields of bullets” null model, where each genus is exposed to a globally homogeneous but province-level kill probability,  $\text{Pr}(\text{die})$ . We use this approach because it generates a correlation between survivorship and geographic range. In this model, the genera that are in more provinces have more chances to survive an indiscriminate culling. Such an assumption generates a correlation between geographic range and survivorship. The probability that a genus goes extinct is the probability that it dies in every province

$$\text{Pr}(\text{extinction}|n_i) = \text{Pr}(\text{die})^{n_i}, \quad (8)$$

where  $n_i$  is the number of provinces genus  $i$  is in. We presume that genus survival is binary, surviving if it evades the kill probability once or more in any combination of regions

$$\text{Pr}(\text{survive}|n_i) = 1 - \text{Pr}(\text{extinction}|n_i). \quad (9)$$

The null model is parametrized by  $\text{Pr}(\text{die})$ , therefore the value must be inferred from data. We use the least squares estimate of  $\hat{\text{Pr}}(\text{die})$  given the observed survivorship data

$$\hat{\text{Pr}}(\text{die}) = \arg \min_{\text{Pr}(\text{die})} \left( \sum_i (X_i - (1 - \text{Pr}(\text{die})^{n_i}))^2 \right), \quad (10)$$

where  $X_i$  is the survival outcome of a genus, one for survival, zero for extinction. Here,  $\arg \min$  is used to indicate a minimization procedure across  $\text{Pr}(\text{die})$ , such that the  $\text{Pr}(\text{die})$  with the least squared error is chosen. Caution should be used when applying this least squares estimate to data that do not have many broadly distributed taxa, for if too few broadly distributed taxa are in the data the outcome may be overfit to those few broadly distributed taxa. However, the Raup and Jablonski dataset has taxa from all geographic range levels, so the model is not overfit. The least squares estimate  $\hat{\text{Pr}}(\text{die})$  for the Raup and Jablonski dataset was 0.805. While changes in this estimate could change the expectation of the model, our latitudinal gradient result is robust to changes of this probability. The observed survival percentage  $Y_{\text{obs}}$  for a group of taxa is therefore

$$Y_{\text{obs}} = \frac{1}{m} \sum_i X_i, \quad (11)$$

where  $m$  is the total number of taxa under consideration. We can also calculate the expected survival percentage  $Y$  from the geographic ranges of taxa in the group under consideration (for us, a BU). The random variable  $Y$  will be normally distributed when the number of taxa is large enough, allowing us to use the standard normal distribution function to calculate the probability that the observed extinction percentage was generated by the null model – which assumes a province-level kill probability that is the same across provinces. The expected value of  $Y$  is

$$E[Y] = \frac{1}{m} \sum_i E[X_i] \quad (12)$$

$$= \frac{1}{m} \sum_i (1 - \hat{\text{Pr}}(\text{die})^{n_i}), \quad (13)$$



while the variance is

$$\text{Var}(Y) = \sum_i \text{Var}\left(\frac{1}{m} X_i\right) \quad (14)$$

$$= \frac{1}{m^2} \sum_i \left(1 - \widehat{\text{Pr}}(\text{die})^{m_i}\right) \left(\widehat{\text{Pr}}(\text{die})^{m_i}\right). \quad (15)$$

In the main text, we used this geographic range null model to compare the observed BU extinction percentage to the expected BU extinction percentage, or  $Y_{\text{obs}}$  versus  $Y$ . The error bars for Figure 2 from the main text were calculated assuming that the observed BU extinction percentage was binomially distributed (the variance above plus the observed binomial variance).

- Erwin, D. H. Lessons from the past: Biotic recoveries from mass extinctions. *Proceedings of the National Academy of Sciences*. **98**, 5399–5403 (2001).
- Jablonski, D. Mass extinctions and macroevolution. *Paleobiology*. **31**, 192–210 (2005).
- Raup, D. M. Biological extinction in earth history. *Science*. **231**, 1528–1533 (1986).
- Wilson, G. P. *et al.* Adaptive radiation of multituberculate mammals before the extinction of dinosaurs. *Nature*. **483**, 457–460 (2012).
- Kiessling, W. & Baron-Szabo, R. C. Extinction and recovery patterns of scleractinian corals at the Cretaceous-tertiary boundary. *Palaeogeography, Palaeoclimatology, Palaeoecology*. **214**, 195–223 (2004).
- Smith, A. B. & Jeffery, C. H. Selectivity of extinction among sea urchins at the end of the cretaceous period. *Nature*. **392**, 69–71 (1998).
- Jablonski, D. Extinction and the spatial dynamics of biodiversity. *Proceedings of the National Academy of Sciences*. **105**, 11528–11535 (2008).
- Harnik, P. G., Jablonski, D., Krug, A. Z. & Valentine, J. W. Genus age, provincial area and the taxonomic structure of marine faunas. *Proceedings of the Royal Society B: Biological Sciences*. **277**, 3427–3435 (2010).
- Valentine, J. W., Foin, T. C. & Peart, D. A provincial model of Phanerozoic marine diversity. *Paleobiology*. **4**, 55–66 (1978).
- Rex, M. A., Crame, J. A., Stuart, C. T. & Clarke, A. Large-scale biogeographic patterns in marine mollusks: a confluence of history and productivity? *Ecology*. **86**, 2288–2297 (2005).
- Valentine, J. W. & Jablonski, D. Origins of marine patterns of biodiversity: some correlates and applications. *Palaeontology*. **53**, 1203–1210 (2010).
- Meyers, L. A., Newman, M. & Pourbohloul, B. Predicting epidemics on directed contact networks. *Journal of Theoretical Biology*. **240**, 400–418 (2006).
- Rosvall, M. & Bergstrom, C. T. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*. **105**, 1118–1123 (2008).
- Roopnarine, P. D. Extinction cascades and catastrophe in ancient food webs. *Paleobiology*. **32**, 1–19 (2006).
- Araújo, M. B., Rozenfeld, A., Rahbek, C. & Marquet, P. A. Using species co-occurrence networks to assess the impacts of climate change. *Ecography*. **34**, 897–908 (2011).
- Paleobiology Database. <http://www.paleodb.org>
- Schulte, P. *et al.* The Chicxulub asteroid impact and mass extinction at the Cretaceous-Paleogene boundary. *Science*. **327**, 1214–1218 (2010).
- Krug, A. Z., Jablonski, D. & Valentine, J. W. Signature of the end-Cretaceous mass extinction in the modern biota. *Science*. **323**, 767–771 (2009).
- Jablonski, D. *Extinction in the fossil record*. In: R. M. May and J. H. Lawton, eds., *Extinction rates*. Oxford, Oxford University Press (1995).
- Raup, D. M. & Jablonski, D. Geography of end-Cretaceous marine bivalve extinctions. *Science*. **260**, 971–973 (1993).
- Supplementary information data are available as supporting material.
- Jablonski, D., Roy, K., Valentine, J. W., Price, R. M. & Anderson, P. S. The impact of the pull of the Recent on the history of marine diversity. *Science*. **300**, 1133–1135 (2003).
- Payne, J. L. & Finnegan, S. The effect of geographic range on extinction risk during background and mass extinction. *Proceedings of the National Academy of Sciences*. **104**, 10506–10511 (2007).
- Heim, N. A. & Peters, S. E. Regional environmental breadth predicts geographic range and longevity in fossil marine genera. *PLoS One*. **6**, e18946 (2011).

- Steuber, T., Mitchell, S. F., Buhl, D., Gunter, G. & Kasper, H. U. Catastrophic extinction of caribbean rudist bivalves at the Cretaceous-Tertiary boundary. *Geology*. **30**, 999–1002 (2002).
- Robertson, D. S., McKenna, M. C., Toon, O. B., Hope, S. & Lillegraven, J. A. Survival in the first hours of the Cenozoic. *Geological Society of America Bulletin*. **116**, 760–768 (2004).
- Pierazzo, E. & Artemieva, N. Local and global environmental effects of impacts on earth. *Elements*. **8**, 55–60 (2012).
- Arens, N. C. & West, I. D. Press-pulse: a general theory of mass extinction? *Paleobiology*. **34**, 456–471 (2008).
- Tobin, T. S. *et al.* Extinction patterns,  $\delta^{18}\text{O}$  trends, and magnetostratigraphy from a southern high-latitude Cretaceous - Paleogene section: links with Deccan volcanism. *Palaeogeography, Palaeoclimatology, Palaeoecology*. **350**, 180–188 (2012).
- Vamosi, J. C. & Vamosi, S. M. Extinction risk escalates in the tropics. *PLoS One*. **3**, e3886 (2008).
- Boucot, A. J. Does evolution take place in an ecological vacuum? II. *Journal of Paleontology*. **57**, 1–30 (1983).
- Sepkoski, J. J. A compendium of fossil marine animal genera. *Bulletins of American Paleontology*. **363**, 1–560 (2002).
- Jablonski, D. & Raup, D. M. Selectivity of end-Cretaceous marine bivalve extinctions. *Science*. **268**, 389–391 (1995).
- Hallam, A. *Atlas of palaeobiogeography*. Elsevier Science and Technology (1973).
- Kreft, H. & Jetz, W. A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*. **37**, 2029–2053 (2010).
- Wright, P., Cherns, L. & Hodges, P. Missing molluscs: field testing taphonomic loss in the Mesozoic through early large-scale aragonite dissolution. *Geology*. **31**, 211–214 (2003).
- Cherns, L. & Wright, V. P. Missing molluscs as evidence of large-scale, early skeletal aragonite dissolution in a Silurian sea. *Geology*. **28**, 791–794 (2000).
- James, N. P., Bone, Y. & Kyser, T. K. Where has all the aragonite gone? *Journal of Sedimentary Research*. **75**, 454–463 (2005).
- Fortunato, S. Community detection in graphs. *Physics Reports*. **486**, 75–174 (2010).
- Lancichinetti, A. & Fortunato, S. Community detection algorithms: A comparative analysis. *Physical Review E*. **80**, 056117 (2009).
- Rosvall, M., Axelsson, D. & Bergstrom, C. T. The map equation. *European Journal of Physics*. **178**, 13–23 (2009).
- Shannon, C. E. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*. **5**, 3–55 (2001).

## Acknowledgments

We thank David Jablonski for guidance with the Paleobiology Database Maastrichtian invertebrates. We benefited from a methods discussion with Charles Marshall, and comments from Douglas Erwin. Additionally, we thank David Jablonski and David Raup for making their bivalve dataset publicly available on the Paleobiology Database.

## Author contributions

DAV, EBH, CAES and GPW conceived the project. DAV, GPW, EBH, CAES and CTB wrote the manuscript. DAV and EBH created the figures. DAV performed the analysis. All authors were involved in interpreting the data and contributed to the manuscript preparation.

## Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**License:** This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

**How to cite this article:** Vilhena, D.A. *et al.* Bivalve network reveals latitudinal selectivity gradient at the end-Cretaceous mass extinction. *Sci. Rep.* **3**, 1790; DOI:10.1038/srep01790 (2013).