This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Wang, Yingying X., G.; Matson, Kevin D.; Prins, Herbert H. T.; Xu, Yanjie; Huang, Zheng Y. X.; de Boer, Willem F.

Title: Risk factors for Lyme disease: a scale-dependent effect of host species diversity and a consistent negative effect of host phylogenetic diversity

Year: 2023

Version: Published version

Copyright: © 2022 The Authors. Published by Elsevier GmbH.

Rights: CC BY 4.0

Rights url: https://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Risk factors for Lyme disease: A scale-dependent effect of host species diversity and a consistent negative effect of host phylogenetic diversity

Yingying X.G. Wang a,b, Kevin D. Matson a, Herbert H.T. Prins c, Yanjie Xu a,d, Zheng Y. X. Huang a,e, Willem F. de Boer a

ABSTRACT

Biodiversity can influence disease risk. One example of a diversity-disease relationship is the dilution effect, which suggests higher host species diversity (often indexed by species richness) reduces disease risk. While numerous studies support the dilution effect, its generality remains controversial. Most studies of diversity-disease relationships have overlooked the potential importance of phylogenetic diversity. Furthermore, most studies have tested diversity-disease relationships at one spatial scale, even though such relationships are likely scale dependent. Using Lyme disease as a model system, we investigated the effects of host species richness and phylogenetic relatedness on the number of reported Lyme disease cases in humans in the U.S.A. at two spatial scales (the county level and the state level) using piecewise structural equation modelling. We also accounted for relevant climatic and habitat-related factors and tested their correlations with the number of Lyme disease cases. We found that species assemblages with more related species (i.e., host species in the order Rodentia) were associated with more Lyme disease cases in humans. Host species richness correlated negatively with the number of Lyme disease cases at the state level (i.e., a dilution effect), a pattern that might be explained by the higher number of reservoir-incompetent species at high levels of species richness at this larger spatial scale. In contrast, a positive correlation was found between species richness and the number of Lyme disease cases at the county level, where a higher proportion of rodent species was associated with higher levels of species richness, potentially amplifying the disease risk. Our results highlight that analysis at a single spatial scale can miss some impacts of biodiversity on human health. Thus, multi-scale analyses with consideration of host phylogenetic diversity are critical for improving our understanding of diversity-disease relationships.

1. Introduction

Understanding the links between host diversity and infectious disease risk (i.e., diversity-disease relationships) is essential under biodiversity loss (Ezenwa et al., 2006; Huang et al., 2016; Johnson et al., 2015; Ostfeld and Keeling, 2012; Pereira et al., 2012). One such relationship, the so-called “dilution effect,” suggests that high host species diversity (e.g., high species richness) within an assemblage reduces disease risk (Keesing et al., 2006; LoGiudice et al., 2003; Ostfeld and Keeling, 2000). Several mechanisms may be behind the dilution effect. For example, in the case of susceptible host regulation, high-diversity assemblages contain more reservoir-incompetent hosts that reduce the abundance of reservoir-competent hosts. With encounter reduction, incompetent hosts reduce the contact rates among reservoir-competent hosts or between hosts and vectors (Keesing et al., 2006). While numerous studies support the dilution effect (Civitello et al., 2015; Huang et al., 2016), its generality remains controversial (Huang et al., 2016; Salkeld et al., 2013; Wood et al., 2014; Young et al., 2013; Zargar et al., 2015).

One limitation of the study of diversity-disease relationships, including the dilution effect, is scale dependency (Huang et al., 2016; Wood and Lafferty, 2013). Indeed, many factors and processes...
governing transmission of multi-host pathogens within host assemblages (such as behaviour, movement, distribution of hosts and vectors, landscape structure, climate, etc. [Estrada-Pena et al., 2014; Rohr et al., 2019]) operate at different scales [Moore et al., 2012], which leads to scale-dependent diversity-disease relationships [Huang et al., 2016; Rohr et al., 2019]. This scale dependency may be especially important in some vector-borne diseases [Huang et al., 2016]. For example, in the case of malaria, the presence of domestic animals around the living quarters of people can attract mosquitoes away from the people and lead to reduced human malaria risk at a local scale. However, these animals may elevate regional mosquito densities, and thus increase average malaria risk at the landscape scale [Dobson et al., 2006]. To date, few empirical studies have examined the effect of host diversity on disease risk across scales.

A second limitation of current studies of diversity-disease relationships relates to the focus on species richness. In addition to species richness, other host assemblage parameters also can be important. For example, it was recently shown that phylogenetic diversity (i.e., host phylogenetic relatedness) can influence disease risk, so studies on diversity-disease relationships should therefore also account for differences in phylogenetic diversity among the studied species [Fountain-Jones et al., 2018; Parker et al., 2015; Wang et al., 2019]. Because closely related host species are more genetically and biologically similar than distantly related ones [Freckleton et al., 2002; Gilbert and Webb, 2007; Webb et al., 2002], related hosts may more readily exchange infectious diseases. For example, shared physiological traits, including immune defence mechanisms [Huang et al., 2013b; Olival et al., 2017; Webb et al., 2002], can equate to reduced molecular, immunological, and ecological barriers that otherwise limit cross-species transmission [Longdon et al., 2011; Vienne et al., 2009]. The role of phylogenetic diversity of host assemblages has been documented in the spread of directly transmitted pathogens, such as plant foliar fungi [Liu et al., 2016; Parker et al., 2015] and avian influenza (Z. Y. X. Huang et al., 2019). However, it is less clear how host phylogenetic diversity shapes vector-borne pathogen dynamics.

Lyme disease, the most commonly reported vector-borne diseases in the temperate zone [Ogden et al., 2009], has increased in incidence and spatial extent in the United States of America (U.S.A.) since 1975 [Bacon et al., 2008]. The organism that causes Lyme disease, the spirochete Borrelia burgdorferi sensu lato, is transmitted to humans through tick bites. The blacklegged tick, Ixodes scapularis, is the primary tick vector in the eastern and central U.S.A. [Barbour and Fish, 1993]. The number of Lyme disease cases in humans is best predicted by the number of infected ticks, which is a product of infection prevalence and density of ticks [Wood and Lafferty, 2013]. Some species (e.g., white-footed mice, Peromyscus leucopus) are reservoir-host species for B. burgdorferi s.l., meaning these species are essential for the persistence and transmission of B. burgdorferi s.l.; other species (e.g., white-tailed deer, Odocoileus virginianus) are not. However, deer, which are fed upon reproducing females in an area might promote the density of questing ticks by providing more blood meals, thereby increasing disease risk [Wood and Lafferty, 2013]. This mechanism is not borne out at the relatively large spatial scale of states in the eastern and central U.S.A., where the number of Lyme disease cases correlates negatively with host species richness [Turney et al., 2014]. So, when it comes to Lyme disease, a model system of the dilution effect, disease diversity relationships and the influence of spatial scales therein are strongly debated [States et al., 2014; Turney et al., 2014].

In addition to host assemblage characteristics, previous studies also identified several biotic and abiotic factors that influence Lyme disease risk [Kilpatrick et al., 2017]. For example, climatic factors can influence Lyme disease risk by influencing ticks: high temperatures can reduce tick densities, and low humidity can reduce tick survival [Diuk-Wasser et al., 2006; Vail and Smith, 2002]. Forest fragmentation might increase the contact rates between humans and ticks, enhancing Lyme disease risk [Tran and Waller, 2013]. Furthermore, forest fragmentation may also influence Lyme disease risk indirectly by affecting host composition. For example, Allan and co-workers found a positive correlation between forest fragmentation and Lyme disease risk. They suggest that habitat fragmentation led to an increased abundance of white-footed mice (P. leucopus), a reservoir-host species of B. burgdorferi s.l. [Allan et al., 2003]. Climate and landscape factors can indirectly impact Lyme disease risk by influencing host assemblages, but these indirect effects are poorly studied.

We investigated the effects of host assemblage characteristics, including both phylogenetic diversity and species richness, and forest fragmentation on Lyme disease risk in the eastern and central U.S.A. We analysed spatial patterns of Lyme disease case numbers at both the county level and the state level because previous studies have revealed the importance of spatial scale in studying diversity-disease relationships [Halliday and Rohr, 2019; Huang et al., 2016; Magnusson et al., 2020; Wood and Lafferty, 2013]. We expected a diluting effect of species richness at smaller spatial scales because biotic interactions are generally stronger at smaller scales [Cohen et al., 2016]. Phylogenetic diversity is expected to be negatively correlated with the number of Lyme disease cases since closely related hosts species (such as an assemblage composed of a higher proportion of Rodentia species) can promote transmission [Fountain-Jones et al., 2018; Parker et al., 2015; Wang et al., 2019]. Our study aims to provide new insights into the dilution effect and the importance of host phylogenetic diversity across scales.

2. Materials and methods

2.1. Disease data

Data on Lyme disease cases in humans in the U.S.A. was obtained from the Centers for Disease Control and Prevention (CDC; http://www.cdc.gov/lyme). These data include the number of Lyme disease cases per year at both the state level (2010-2016) and the county level (2000-2016; Fig. 1). There are two tick species associated with the expansion of Lyme disease in the U.S.A., namely the blacklegged tick (I. scapularis) in the eastern part of the U.S.A. and the western blacklegged tick (Ixodes pacificus) in the western part. We only used data from states and counties with documented occurrences of I. scapularis [Eisen et al., 2016], since B. burgdorferi s.l. prevalence in ticks and the number of Lyme disease cases are higher in the eastern part of the U.S.A.

Some reported zeros (i.e., no cases of Lyme disease reported) may represent false absences (i.e., Lyme disease cases were reported as zero before the first establishment of Lyme disease; We treat zeros after the first establishment of Lyme disease as true zeros). To partly address this potential bias, we first excluded all states or counties without any reported cases during the whole study period, and assumed that Lyme disease has never been established in these counties. For all remaining states or counties, we then removed all reported zeros before the year when at least one Lyme case was reported. The final datasets consisted of 35 states with a total of 285 annual counts of Lyme disease cases and 1213 counties with 9741 annual counts of Lyme disease cases (Fig. 1).

2.2. Host assemblage data

Based on a list of mammal host species of I. scapularis [Turney et al., 2014] and distribution ranges for these species from the IUCN Red List database [IUCN, 2015], we calculated host species richness (Table 1; SR;
total number of mammal host species) per state and per county. By using a recent published phylogenetic tree of mammals (Rolland et al., 2018), we quantified phylogenetic relatedness within the mammal host assemblages by calculating the mean pairwise phylogenetic distance (MPD) per state and per county. Because of the potential correlation between MPD and SR (Webb et al., 2000), we standardised MPD using a null model by shuffling the tip labels of the phylogenetic tree. The resulting standardised MPD (MPD.Z) was not correlated with SR (Swenson, 2014; Wang et al., 2019). Calculations of SR, MPD, and MPD.Z across all the years of the study were conducted in R (version 3.5.0) using functions of pd, mpd, sesmpd respectively in the Picante package (Kembel et al., 2010).

2.3. Data processing of other predictors

Based on the results of a previous study on Lyme disease in the U.S.A (Turney et al., 2014), we also included several covariates that can affect the number of Lyme disease cases (Table 1). Using land cover data from the U.S. Geological Survey (U.S.G.S. Gap Analysis Program, 2011), we calculated per state and per county the total combined area of deciduous and coniferous forests (i.e., habitat area), the preferred habitats of I. scapularis (Ostfeld et al., 1995). We calculated edge density as an index of the fragmentation of these preferred habitats. We calculated the distance to source as the distance of the closest border of each state and county to the closest border of Connecticut, important geographic origin of Lyme disease in North America (Turney et al., 2014). We included this variable to compare our analyses with the results from (Turney et al., 2014). We calculated the mean annual temperature and the mean annual precipitation using the WorldClim version 2.0 database (Fick and Hijmans, 2017). Finally, we accounted for the total area (km²) of each state and county (Area). All data processing and variable calculations were performed in ArcGIS (version 10.5).

2.4. Statistical analyses

To determine the relationships between predictors and the number of Lyme disease cases in humans at the state level, we first fitted a generalised linear mixed model (GLMM) with a Poisson distribution to compare our results with a previous study (Turney et al., 2014). We also used an observation-level random effect to deal with overdispersion of
the data. We carried out a similar analysis at the county level using a GLMM with a negative binomial distribution to account for the overdispersion and with state included as a random factor (since, e.g., Lyme disease prevention strategies might vary by state). For both state- and county-level analyses, we included human population size as an offset. Before performing GLMMs, collinear predictor variables were excluded (Pearson correlation > 0.8; Table S1, S2), and the selected variables were scaled to have a mean of 0 and a standard deviation of 1. We used backward selection to select the best models (i.e., model with least Akaike’s information criterion, AIC). All models were fitted in R (version 3.5.0) using the lme4 package.

Based on the GLMM results, we also fitted piecewise Structure Equation Models (SEMs; electronic supplementary material, Fig. S1) to examine the interdependencies of the predictors and their direct and indirect effects on the number of Lyme disease cases at both the county level and the state level (Zelcheck, 2016). Under this approach, we constructed several regressions. The first GLMM, which had the number of Lyme disease cases as response variable, was a full model that included host species richness (SR), standardised mean pairwise phylogenetic distance (MPD.Z), distance to source, habitat area, edge density, mean annual temperature, and mean annual precipitation as predictor variables. Because habitat factors (i.e., habitat area and edge density) and climatic factors (i.e., mean annual temperature and mean annual precipitation) can also have indirect impacts on the number of Lyme disease cases via host SR, MPD.Z, or both, two additional models were constructed. Here, we used linear mixed models (LMMs) for the two endogenous variables (i.e., SR and MPD.Z) that only included clear mechanistic predictors (Figure S1). For SR, the LMM included habitat size, edge density, mean annual temperature, and mean annual precipitation; for MPD.Z, the LMM included only edge density. An observation-level random effect was included in the state level models, and state was included as a random factor in the county level models. We report the standardised coefficients for each path in each model. We also report marginal $R^2$ values, which measure the variation explained by fixed factors only. We compared AICc values between SEMs with and without certain direct effects and selected the best final model. The overall fit of the piecewise SEMs was evaluated by Fisher’s C statistic, which indicates whether there are any missing paths. All SEMs were fitted in R with the piecewiseSEM package.

To better understand the potential mechanism behind the effect of host species richness (SR) and standardised mean pairwise phylogenetic distance (MPD.Z) on the number of Lyme disease cases, we also constructed linear models to study the link between SR (response variable) and the proportion of reservoir-competent hosts (i.e., Peromyscus spp.) and the proportion of deer species (i.e., Odocoileus spp.) at both the county level and the state level. We also tested the relationship between MPD.Z and proportion of Peromyscus spp. and the proportion of Rodentia species at both the county level and the state level.

3. Results

At the larger spatial scale of analysis, the state level, the GLMM analysis (Table 2) showed that host species richness (SR), standardised mean pairwise phylogenetic distance (MPD.Z), habitat area, distance to source, mean annual temperature, and mean annual precipitation were all negatively correlated with the number of Lyme disease cases. The SEM (Fig. 2a) illustrates these negative effects, while also showing that both mean annual temperature and mean annual precipitation had a significant positive effect on SR and that edge density had a significant negative effect on MPD.Z. The Fisher’s C value for the final SEM was 5.81, indicating that there were no missing paths, and the p value for the final SEM was 0.45. A $p > 0.05$ from a Chi-square test ($χ^2$) indicates optimal fitting of the chosen SEM (Fan et al., 2016; Mulaik et al., 1989).

At the county level, the GLMM analysis (Table 2) showed that the number of Lyme disease cases was positively correlated with SR, habitat area, edge density, and mean annual temperature, but negatively correlated with MPD.Z, distance to source, and mean annual precipitation. The SEM (Fig. 2b) illustrates these correlations, while also showing that mean annual temperature had a significant negative effect on host species richness and that mean annual precipitation and edge density had positive effects on host species richness. The Fisher’s C value of this SEM was 1.11, indicating that no paths were missing. The p value of the final SEM was 0.58, indicating a good fit.

At the state level, SR was negatively correlated with the proportion of Odocoileus spp. (Table S3). At the county level, SR was positively correlated with the proportion of Peromyscus spp. and negatively correlated with the proportion of Odocoileus spp. (Table S3). At both the county level and the state level, MPD.Z was significantly negatively correlated with both the proportion of Peromyscus spp. and the proportion of Rodentia species (Table S4).

4. Discussion

The potential scale dependency of relationships between biodiversity and disease risk has been previously acknowledged (Halliday and Rohr, 2019; Johnson et al., 2015; Magnusson et al., 2020; Wood and Lafferty, 2013), but rarely tested. Moreover, most studies of diversity-disease relationships have focused on host species richness only, ignoring differences in host phylogenetic diversity. Here, we analysed the effects of the composition of host assemblages, including both species richness and phylogenetic diversity, on the number of Lyme disease cases in humans at both the county level and the state level in the U.S.A. In these analyses, we also took into account several climatic and landscape factors. The direction of the effect of host species richness on the number of Lyme disease cases was scale dependent, while host phylogenetic diversity was consistently negatively correlated with Lyme disease cases at both spatial scales.

4.1. Scale-dependent relationships between host species richness and number of Lyme disease cases

The scale dependency of diversity-disease relationships is part of a large scientific debate (Halliday and Rohr, 2019; Johnson et al., 2015; Magnusson et al., 2020; Wood and Lafferty, 2013). For example, Halliday and Rohr (2019) studied 205 diversity-disease relationships and concluded that negative relationships generally occur at small scales (<100 km²) where species interactions are strong, and that positive relationships generally occur at regional scales (>1,000,000 km²). However, Magnusson et al. (2020) studied diversity-disease relationships across spatial scales from 38 studies and found significant negative
diversity–disease relationships at small (10 m-10 km), intermediate (10-2,000 km), and large (> 2,000 km) scales. Moreover, even for Lyme disease, a model system of the dilution effect, the influence of spatial scales is strongly debated (States et al., 2014; Turney et al., 2014). For example, Wood and Lafferty (2013) suggested that the dilution effect may occur at small spatial scales (e.g., communities in forests). However, Turney et al. (2014) found that species richness reduced cases of Lyme disease at larger spatial scale (i.e., the state level).

Our analyses suggested a scale-dependent effect of host species richness. Specifically, a positive diversity-disease relationship at the county level, and a negative relationship at the state level. Some previous studies indicate that biodiversity can sometimes amplify disease risk, even at small spatial scales, when reservoir-competent hosts increase with increasing biodiversity (Halliday et al., 2017; Halliday and Rohr, 2019). In our study, the positive relationship at the county level between host species richness and the proportion of *Peromyscus* spp. (i.e., reservoir-competent host species) supports this mechanism. Higher species richness at the county level was associated with relatively more reservoir-competent host species, which can thereby increase the number of infected ticks and the number of Lyme disease cases in humans. Similar positive relationships between host species richness and disease risk have been reported for other vector-borne disease. For example, in the case of malaria, more animals support elevated mosquito densities, which increase regional malaria risk (Dobson et al., 2006). To fully understand these relationships, it is important to know more about species densities, and tick densities, and *Borrelia* prevalence levels.

At the state level, we found a strong negative relationship between host species richness and the number of Lyme disease cases. This result agrees with an earlier analysis (of the same database) that did not account for host phylogenetic diversity (Turney et al., 2014). At the state level, the proportion of *Odocoileus* spp., which are incompetent for *Borrelia* but important for tick reproduction, decreases with increasing overall host species richness. A decrease in the proportion of *Odocoileus* spp. could reduce overall tick abundance and thereby the number of Lyme disease cases in humans (Wood et al., 2013), which could explain the dilution effect. A dilution effect at the larger (i.e., state level) spatial scale might result from greater spatial heterogeneity preventing the spread of the disease by limiting the movement of host species (Estrada-Peña et al., 2014) or by limiting the contact rate between ticks and host species. More analyses that incorporate data on densities and movements of reservoir-competent hosts species are required to better understand the underlying causal mechanisms of the dilution effect at Fig. 2. Path diagram of a piecewise Structural Equation Model (SEM) showing the direct and indirect effects of predictors on the number of Lyme disease cases at the state level (a) and the county level (b). Solid red arrows represent positive effects ($p < 0.05$), solid black arrows represent negative effects ($p < 0.05$), and dotted grey arrows represent non-significant effects ($p > 0.05$). We report the path coefficients as standardised effect sizes next to arrows.
larger spatial scales.

In this study, we found a strong effect of species richness on Lyme disease cases in humans at both the county level and the state level. Results of our study do not necessarily conflict with previous studies because the current study was carried out at two intermediate spatial scales (e.g., median size of state = 120,740 km² and median size of county = 1,600 km²). Our results hold two important messages about the strength and direction of diversity-disease relationships. First, the impact of species richness on disease risk can be strong even beyond the local level. Second, the direction of the effect of species richness on disease risk depends on the relationship between species richness and the proportion of reservoir-competent host species (e.g., Peromyscus spp.) and proportion of reservoir-incompetent species (Odocoileus spp.). For example, disease risk increases when reservoir-competent hosts increase with species richness.

4.2. Consistent negative relationships between host phylogenetic diversity and number of Lyme cases

At both the county level and the state level, host phylogenetic diversity (as indexed by standardised mean pairwise phylogenetic distance, MPD.Z) was negatively correlated with the number of Lyme disease cases. In other words, assemblages characterised with low standardised mean pairwise phylogenetic distance (i.e., high phylogenetic relatedness) have a larger number of Lyme disease cases. This result suggests that, as expected, highly related assemblages are composed of relatively more host species that are competent for both B. burgdorferi s.l. and its vector tick. These closely related host species share biological traits, including some immunological defences (Gilbert and Webb, 2007; Longdon et al., 2011). Borrelia burgdorferi s.l. is more likely to survive and circulate in species assemblages composed of closely related host species. In addition, related host species sometimes share habitats or habitat components that offer broadly suitable conditions (Fountain-Jones et al., 2018; Huang et al., 2014; McCoy et al., 2013). Such biological and ecological ‘overlaps’ might increase contact rates between ticks and suitable hosts, facilitating the completion of the tick life cycle, increasing tick densities, and promoting the transmission of B. burgdorferi s.l.

In addition, our results showed that the impact of MPD.Z on Lyme disease cases is not scale-dependent. Transmission of pathogens is often a complex process, influenced by host-vector, host-pathogen, and vector-pathogen interactions (Estrada-Peña et al., 2014). Pathogens tend to be able to infect closely related host species (Gilbert and Parker, 2016; Streicker et al., 2010). This host-pathogen process is likely independent of spatial scale.

Host-vector interactions might be influenced by spatial scales, as distribution of species or composition of species vary over spatial scales (i.e., the county level versus the state level). However, we detected negative relationships between the standardised mean pairwise phylogenetic distance (MPD.Z) and both the proportion of Peromyscus spp. and the proportion of Rodentia at both the county level and the state level. Communities that contain more closely related species also have a relatively higher number of reservoir-competent host species. In such cases, ticks are more likely to feed on reservoir-competent hosts, and thus promote transmission of B. burgdorferi s.l.

4.3. Edge density and habitat area have scale-dependent direct and indirect influences on the number of Lyme disease cases

Edge density, which we used as proxy for forest fragmentation, had both direct and indirect positive effects on the number of Lyme disease cases at the county level, whereas it had only an indirect effect at the state level.

At the county level, higher edge density of deciduous and coniferous forests was linked to a higher risk of Lyme disease, which is in agreement with previous Lyme disease studies carried out at the county level in U.S.
scales. Future studies of diversity-disease relationships should take into account not only host species richness but also host phylogenetic diversity in order to continue advancing our understanding of this timely research topic, which spans the interests of fundamental and applied ecologists and conservation biologists.

**Funding**

This work was supported by the National Natural Science Foundation of China (31870400, 32271605), Chinese Scholarship Council (No. 201506190134), the Program of Natural Science Research of Jiangsu Higher Education Institutions of China (16KJB1800013) and the Priority Academic Programme Development (PAPD) of Jiangsu Higher Education Institutions.

**Data accessibility**

The numbers of cases of Lyme disease at both the county level and the state level were obtained from the United States Centers for Disease Control (http://www.cdc.gov/lyme). Land cover data was obtained from the United States Geological Survey (USGS https://www.usgs.gov/land-resources/eros/lulc). The climate data are available from WorldClim (http://www.worldclim.org/). Biodiversity data and R code for analysis are available from Supplementary material.

**CRediT authorship contribution statement**

Yanjie X.G. Wang: Conceptualization, Methodology, Software, Formal analysis, Data curation, Validation, Writing – original draft, Writing – review & editing. Kevin D. Matson: Conceptualization, Validation, Supervision, Writing – original draft, Writing – review & editing. Herbert H.T. Prins: Resources, Writing – review & editing. Yanjie Xu: Formal analysis, Visualization, Writing – review & editing. Zheng Y.X. Huang: Conceptualization, Software, Validation, Writing – review & editing. Funding acquisition. Willem F. de Boer: Supervision, Conceptualization, Validation, Writing – original draft, Writing – review & editing.

**Declaration of Competing Interests**

We have no competing interests.

**Data availability**

I have shared links to the data I used in Data accessibility.

**Acknowledgements**

We thank Helen J. Esser for her valuable suggestions.

**Supplementary materials**

Supplementary material associated with this article can be found in the online version, at doi:10.1016/j.tbd.2022.102073.

**References**

