# Towards a networks based approach to biomonitoring

Clare Gray

School of Biological and Chemical Sciences

Queen Mary University of London

A thesis submitted to the University of London for the degree of Doctor of Philosophy

February 2016

# Statement of originality

I, Clare Gray, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature:

Date: 12/2/16

# Details of collaboration and publications:

**Chapter one**: This text is adapted from work was done alongside many co-authors, which was published in Journal of Applied Ecology in 2014. Myself and Guy Woodward conceived the original idea, I organised a meeting and invited co-authors, I then organised contributions from co-authors and edited them into the final document. All co-authors contributed towards the final text. The full list of co-authors is: Clare Gray, Donald J. Baird, Simone Baumgartner, Ute Jacob, Gareth B. Jenkins, Eoin J. O'Gorman, Xueke Lu, Athen Ma, Michael J. O. Pocock, Nele Schuwirth, Murray Thompson and Guy Woodward. Full permission to reproduce the text and figures here has been received.

See here for the online version: http://onlinelibrary.wiley.com/doi/10.1111/1365-2664.12300/abstract

**Chapter two**: This chapter was published in Food Webs in 2015, and appears here with appropriate permissions from the publisher. I conceived the original idea alongside Guy Woodward. Myself, Dan Perkins and Lawrence Hudson then developed these ideas further and designed the analysis. Lawrence Hudson wrote the R code for the function itself. I did all the analysis and wrote the manuscript. All co-authors contributed towards the final text. The full list of co-authors is: Clare Gray, David H. Figueroa, Lawrence N. Hudson, Athen Ma, Dan Perkins and Guy Woodward. Full permission to reproduce the text and figures here has been received.

See here for the online version:
http://www.sciencedirect.com/science/article/pii/S2352249615300045

**Chapter 3**: This work was done with data belonging to the Upland Waters Monitoring Network. I conceived the original idea alongside Guy Woodward and Athen Ma. I developed it, conducted the analysis and wrote the manuscript. It is planned that this chapter will be submitted to Advances in Ecological Research in early 2016. The full list of co-authors is: Clare Gray, David McElroy, Xueke Lu, Athen Ma, Don Monteith,

Ewan Shilland and Guy Woodward.

**Chapter 4**: This work was done alongside David Bohan, Órla McLaughin and Sandrine Petit from Institut National de la Recherche Agronomique in Dijon, France. I conceived the original idea, alongside David Bohan and Órla McLaughlin, we then developed that idea and the analysis. I conducted the analysis and wrote the manuscript with input from David Bohan. All co-authors contributed towards the final text. It is planned that this chapter will be submitted to Current Biology in early 2016. The full list of co-authors is: Clare Gray, Athen Ma, Órla McLaughin, Sandrine Petit, Guy Woodward, David Bohan.

**Chapter 5**: This was a large project involving many organisations. Guy Woodward and Murray Thompson conceived the original sampling design. I led the fieldwork and much of the lab work for the food web analysis, whilst supporting Murray Thompson with the remaining fieldwork. He and I processed and identified the invertebrates from September 2013 and March 2014, I supervised students (Laura Palmer, Amy West, Marie-Clare Danner and Masamichi Fujisawa) who processed the invertebrates from September 2014 and March 2015. Carl Sayer, Laura Rutland and Scot Warren processed and identified the diatom samples from September 2013, March 2014 and September 2014. Murray Thompson processed the leaf litter bags from September 2013 and March 2014, I then supervised Amy West who processed the bags from September 2014 and March 2015. I supervised students Laura Palmer and Cara Patel to process the Chlorophyll samples. I constructed the food webs and did all the analysis presented here. The data from September 2013 was published by Freshwater Biology in 2015 (Appendix C). I contributed the food web analysis and text directly to the manuscript (Appendix C) although I was not the lead author.

**Appendix A:** This work has been done alongside collaborators Samraat Pawar and Katharina Brink from Imperial College London, Ian Donohue from Trinity College Dublin and Xueke Lu, Athen Ma and Guy Woodward. I conceived and developed the original idea with conversations with each of the collaborators. Katharina Brink

provided R code to measure the Mutual Information of the food webs.

**Appendix B:** This manuscript has been submitted to Nature Climate Change, at the time of writing we are responding to a second round of reviewers' comments. I have worked alongside Xueke Lu and Athen Ma since the beginning of this project, I am joint first author with Xueke Lu. Athen Ma, Xueke Lu and I have met regularly to discuss the analysis and writing up of this project, Xueke Lu has done all the analysis and I wrote the initial manuscript draft. All co-authors contributed to the final submitted draft. The full list of co-authors is: Xueke Lu, Clare Gray, Lee Brown, Mark Ledger, Alexander Milner, Raúl Mondragón, Guy Woodward and Athen Ma.

**Appendix D:** This work was published as a chapter in the book 'Aquatic Functional Biodiversity: Ecological and Evolutionary Approaches' and was done alongside co-authors. I organised a meeting and invited co-authors, I then organised contributions from co-authors and edited them into the final document. All co-authors contributed towards the final text. The full list of co-authors is: Clare Gray, Iliana Bista, Simon Creer, Benoit Demars, Francesco Falciani, Don Monteith, Xiaoliang Sun and Guy Woodward. Full permission to reproduce the text and figures here has been received.

# Abstract

Effective monitoring of the environment for anthropogenic impacts is essential for managing and conserving ecosystems, especially in the face of global climate change and an ever increasing human population. Yet current biomonitoring schemes are grounded in species or trait based approaches, and lack the tools required to deal with the effects of stressors on species and their interactions in complex natural systems. Ecological networks can offer new insights into ecosystem degradation by explicitly considering the interactions between species, adding value to current taxonomically constrained schemes.

Here, I develop a formalisation of a method for constructing ecological networks from species lists and trophic information harvested from the primary literature (Chapter 2). I then use this method to augment traditional biomonitoring data with information on the interaction between species to build large collections of food webs (Chapters 3-5). I apply novel network analysis methods from complex network research to examine the substructure of these networks. In Chapter 3, I find that the structure, and substructure, of freshwater food webs are fundamentally altered by hydrochemical stress (Appendix A). Chapter 4 demonstrates that the structure of agricultural food webs are linked to the delivery of beneficial pest control services, potentially allowing those services to be enhanced through management of food web structure. Finally, in Chapter 5 I use more detailed food web data to investigate how freshwater food webs are impacted by a catastrophic pesticide spill, how the indirect effects propagate through the food web, and how the structure of the community and ecosystem functioning recover over time.

The findings presented herein demonstrate that ecological networks constructed from routine biomonitoring data can be a useful tool for understanding the impacts of stressors on ecological communities. Considering the interactions between species is vital if we are to fully understand, and mitigate against the negative effects of global climate change on biodiversity.

# Acknowledgements

Firstly, I'd like to thank my supervisors, Dr Athen Ma and Professor Guy Woodward for initially giving me the opportunity to take on this project, and then supporting and guiding me through it. Additionally, they have given me many opportunities to get involved in some fantastic side projects which have really allowed me to get as much out of my time as a PhD student as I could. Mark Trimmer acted as my panel chair throughout my PhD, as well as took over as surrogate supervisor when I needed it, and he always provided me with useful feedback and interesting analysis ideas during our panel meetings.

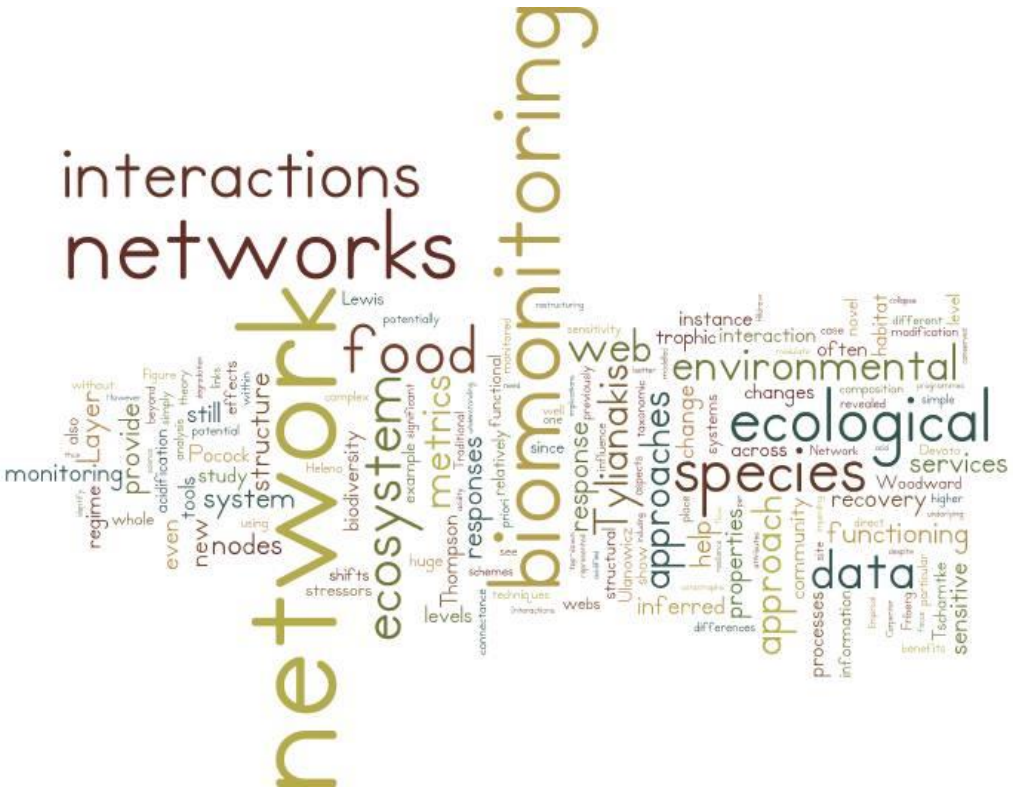
The Woodward lab group (both past and present) have been unfailingly supportive and helpful, Dan Perkins, Eoin O'Gorman, Órla McLaughlin, Murray Thompson and Becca Kordas were always there to talk through ideas and problems with me. Xueke Lu has patiently answered my silly questions without fail. So many, many people have helped with the Kennet project and made the rigorous fieldwork schedule possible, in particular Marie-Claire Danner, Joseph Huddart, Manon Czuckermand, Laura Palmer, Laura Rutland, Amy West, Masamichi Fujisawa and Cara Patel. Thanks also to the many people I have bothered for help under the guise of a collaboration; David Bohan, Ewan Shilland, Don Monteith, Lawrence Hudson, Samraat Pawar, Ian Donohue and Katharina Brink.

Most importantly, I need to thank my parents, Roger and Susan Gray, and extended family, without whose support and encouragement over the years none of this would have been possible. Lastly, special thanks go to Alex, who has patiently endured the many grumps, moans, antisocial working hours, and the car (and me) smelling of fish. Thank you Alex, congratulations, you survived the bear trap!

This work was generously supported by a Queen Mary University of London Studentship, and the Freshwater Biology Association.

# Table of Contents

# 1 | Ecological networks: the missing links in biomonitoring science[1]




[1] Gray, C., Baird, D.J., Baumgartner, S., Jacob, U., Jenkins, G.B., O'Gorman, E.J., Lu, X., Ma, A., Pocock, M.J.O., Schuwirth, N., Thompson, M. & Woodward, G. (2014). Ecological networks: the missing links in biomonitoring science. *Journal of Applied Ecology*, **51**, 1444–1449.

## 1.1 Biomonitoring in the 21st century

Biomonitoring programmes were first developed in the wake of the Industrial Revolution to measure the effects of environmental stressors on the natural world. They focus on measuring the biological response to environmental perturbations, as biota integrate the changes in environmental quality over time into one measure (i.e. presence, absence or composition), unlike chemical data which reflects only one point in time. Most current programmes are taxonomically constrained and monitor changes in biodiversity, although increasingly, aspects of ecosystem functioning are being incorporated. A range of indices have been developed which score taxa on their known sensitivity to particular stressors (e.g. Balloch, Davies & Jones. 1976; Wright, Furse & Armitage 1993; Bonada *et al.* 2006; Murphy *et al.* 2013). Changes are assessed against a baseline level relative to a reference or idealised level (e.g. targets for restoration or acceptable levels of a response variable for that place and time).

This approach of assessing the biota of a site with respect to a 'reference' condition now underpins many biomonitoring schemes across Europe (e.g. Simpson *et al.* 2005; Murphy *et al.* 2013) and other parts of the world (Simpson & Norris 2000). However, pre-industrial (i.e. pre 1800) target conditions for many habitats no longer exist, or are very rare; many of European freshwater habitats are impacted (Friberg *et al.* 2011; Malaj *et al.* 2014) and many grasslands worldwide have experienced eutrophication driven by emissions and nitrogen deposition (Clark & Tilman 2008). Furthermore preindustrial states are very difficult to model with confidence (Battarbee *et al.* 2005). Additional challenges are provided by global climate change, since the reference conditions themselves may be shifting (Pauly 1995; Bennion *et al.* 2011). Unfortunately, this makes assigning appropriate reference conditions for biomonitoring problematic.

Due to the paucity of baseline data, current biomonitoring schemes are still unable to diagnose many perturbations, often there is also a generally poor understanding of the underlying ecological mechanisms governing an ecosystems response to

environmental change (Friberg *et al.* 2011). Newly emerging environmental threats, such as the many facets of climate change, pose new challenges for biomonitoring schemes. Thus, there is a growing need to determine how best to assess the impact of these emerging stressors, both in isolation and in combination. Also, the structural biodiversity-centric focus of traditional methods (e.g. Wright, Furse & Armitage 1993; Metcalfe-Smith 1996; Murphy *et al.* 2013) now needs to be augmented with more explicitly functional measures (e.g. Young, Matthaei & Townsend 2008), to provide complementary insights into the impacts of stressors (Woodward *et al.* 2012b). Incorporating species interactions into biomonitoring approaches may help to overcome many of the limitations of current biomonitoring approaches, and provide a new template for ecosystem monitoring.

## 1.2 The advantages of incorporating species interactions into biomonitoring schemes

Traditional biomonitoring schemes have focused on presence/absence or abundance of taxa (network 'nodes') across environmental gradients, while ignoring the network of pairwise interactions ('links') between them (Friberg *et al.* 2011). Such taxonomic grounding limits its ability to generalise beyond the characteristic biota of a given region or system. For instance, when assessing the ecological status of European rivers, huge effort has been devoted to harmonising approaches and data across member states, forcing practitioners to resort to complex statistical inter-calibration (see Birk *et al.* 2013). However, network approaches are not reliant on the taxonomy of the nodes *per se*, and so, in theory, can be used to compare emergent topologies of networks irrespective of biogeographical differences in species composition.

Environmental legislation increasingly requires both the structural and functional attributes of a particular community to be considered (e.g. the Water

Framework Directive; European Commission 2000), but the latter are often still missing or inferred, despite increasing calls for them to be embedded in ecological assessments. Network approaches can help address this gap as many ecosystem processes and the services they provide depend on interactions between taxa (Thompson, Dunne & Woodward 2012). Interactions between these network nodes influence biodiversity and ecosystem functioning (Kremen 2005; Thompson, Dunne & Woodward 2012) and a system's sensitivity to environmental change (Tylianakis, Tscharntke & Lewis 2007). Changes in network structure can provide clues to altered dynamics and ecosystem functioning.

A classic example of food web interactions determining alternative outcomes of both structural and functional responses to environmental stressors comes from shallow lakes. Here, catastrophic regime shifts are triggered by extreme nutrient concentrations, but in intermediate conditions trophic cascades in the food web can flip the ecosystem from one stable state to another, even in the absence of additional environmental change (Scheffer & Carpenter 2003). Ecological hystereses, whereby community recovery is modulated by the biota and not simply the reverse trajectory of the response to an impact (Scheffer & Carpenter 2003), highlight how the network of species interactions that underpin critical processes and services (such as clean water, or fisheries) can influence both the internal dynamics of the system and its resilience to environmental change (e.g. Thompson, Dunne & Woodward 2012). A good example of this is that of Broadstone Stream (**Figure 1**), as the food web of this small stream recovered from the effects of acidification, counterintuitive patterns emerged. The community response did not simply show a straightforward reverse of the trajectory of the response to acidification, and invertebrate numbers actually declined as pH rose. These system-level responses only made sense when viewed in the context of the food web: the declines in invertebrate numbers coupled with a succession of invasions of progressively larger predators, represented increasing top-down effects and the resultant restructuring of the mass-abundance scaling properties

of the network even though the prey assemblage composition remained relatively constant. Traditional biomonitoring techniques could not explain this ecological response because they lacked the key ingredient: species interactions within the food web.

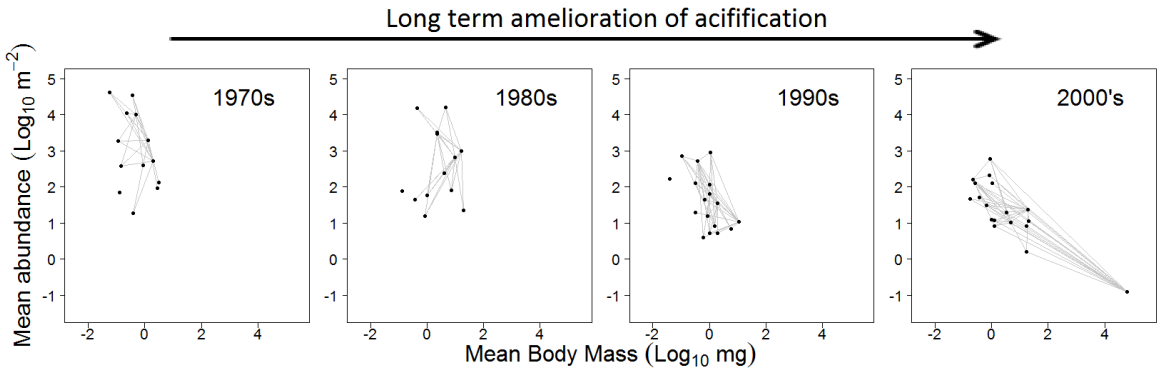



**Figure 1**. Broadstone stream food webs plotted in 'trivariate' space; as species abundance versus body mass data, with links between nodes representing trophic interactions. The abundance of invertebrates declines despite improving environmental conditions, as top-down effects intensify. Redrawn from Layer et al. (2011).

Keystone species can be identified through a network approach (e.g. Jordán 2009), helping to focus monitoring efforts towards those that are ecologically most significant, since highly connected species often determine network stability and vulnerability to cascading secondary extinctions (Dunne, Williams & Martinez 2002b). Similarly, a network approach can also help improve efficiency by identifying and tracking those species or interactions that are most sensitive to change: thus, keystone and indicator nodes could help provide novel early warning systems for detecting impending regime shifts or catastrophic ecosystem collapse (Aizen, Sabatino & Tylianakis 2012).

A network approach can help to reveal the complicated direct and indirect effects of stressors on an ecological community, beyond the simple loss or gain of species. For example, when freshwaters are acidified and specialist herbivores are excluded, generalist herbivore–detritivore species occupy their niche space, slowing

their re-establishment (e.g. Layer, Hildrew & Woodward 2013). Translocation experiments have shown that these acid tolerant consumers can perform just as well, if not better, in the absence of interactions with more acid sensitive species in the network, suggesting they are not simply acidophilous. Empirical and modelling work has provided some evidence that generalist acidified networks are more robust than their counterparts at higher pH: i.e. ecological inertia within the food web may be modulating biological recovery as acidity ameliorates (Layer *et al.* 2010b; Layer, Hildrew & Woodward 2013).

Network analysis has also revealed how another major environmental stressor – drought – leads to a top-down erosion of stream food webs: large and rare species high in the web are especially sensitive and overall ecosystem functioning is compromised due to severely impaired biomass fluxes through the network (Ledger *et al.* 2013). The complex interconnected consequences of environmental stress for a particular system can thus only be fully understood from a network perspective, allowing *a priori* predictions to be made and appropriate management strategies to be developed. Ecotoxicology could also benefit from taking this more system-based approach, as different pest control agents (insecticides, herbicides, fungicides) will affect different trophic levels and compartments in the food web, with ramifications that ripple far beyond the intended targets or other species with acute sensitivity to the poison. Monitoring the network as a whole would help detect these potentially critical indirect and often unanticipated effects (e.g. Baird *et al.* 2001).

Consideration of the interactions between species when assessing a community's response to perturbations can provide a deeper insight into the mechanisms governing those responses (Scheffer & Carpenter 2003; Woodward *et al.* 2010a; Friberg *et al.* 2011). An understanding of the characteristics of a 'healthy', unperturbed community might remove the need for a 'reference condition' approach to biomonitoring altogether, allowing the limitations associated with that approach, such as shifting baselines, to be circumvented.

## 1.3 Incorporating ecological networks into biomonitoring schemes

Although potentially useful, network-based approaches must still overcome some significant challenges, particularly in terms of gathering data on interactions. In some cases biomonitoring data are explicitly interaction-based, e.g. monitoring pollinators by collecting individuals from flowers (as in Kremen, Ullman & Thorp 2011; Pocock, Evans & Memmott 2012) but, on the whole, direct monitoring of the interaction itself is currently too labour intensive to be practical in routine biomonitoring schemes (Hegland *et al.* 2010). For example, to characterise the interactions in a traditional food web it is necessary to examine many hundreds of guts for each consumer species (e.g. Woodward, Speirs & Hildrew 2005). If a networks based approach to the widespread biomonitoring of the natural world is to be adopted, then the efficiency with which ecological networks can be built must be dramatically increased.

Where directly observing interaction data is impractical, one approach is to augment monitoring data by inferring interactions based on prior knowledge. Such inferences are especially valuable where assemblages across trophic levels are routinely monitored, e.g in aquatic systems (fish, macroinvertebrates and algae in freshwaters and whole fish assemblages in the sea). Interactions can be added from previously observed interactions, e.g. from data papers (e.g. Brose *et al.* 2005; Barnes 2008) and online resources, such as the Interaction Web Database (http://www.nceas.ucsb.edu/interactionweb/index.html) or the Database of Insects and their Food Plants (http://www.brc.ac.uk/dbif/). For instance, Mulder and Elser (2009) constructed a set of 22 food webs from biomonitoring data and published trophic interactions to show how chemical soil properties influence network structure and hence soil processes and services. Quantitative networks can be created from these known interactions based on simple rules (e.g. Chapter 4; Pocock, Evans & Memmott 2012). Where historic data exist (e.g. the UK Upland Waters Monitoring

Network; Kernan *et al.* 2010) networks could even be inferred by hindcasting back through time (Chapter 3).

Such inferred networks have potential limitations, however, as they ignore possible behavioural differences in species between systems, (i.e. preferential feeding depending on which resources are available) and unexpected or state-specific changes in networks (e.g. those pre-empting regime shifts) could go undetected. Notwithstanding these caveats, the potential benefits are substantial, as the parameterisation of networks using simple allometric scaling rules could ultimately allow interaction strengths or energy fluxes to be inferred and stability or productivity to be modelled dynamically (e.g. Appendix A; Berlow *et al.* 2009; Layer *et al.* 2010b; Tang, Pawar & Allesina 2014). This would provide a currently missing system-level link between structure and (inferred) functioning. Inferring networks from the vast amounts of biomonitoring data already in existence would bring the benefits of ecological network science into aspects of biomonitoring, while circumventing the huge effort required to construct each network anew from direct observation. To achieve this, a systematic tool for the automated construction of large volumes of ecological networks is required (Chapter 2; Gray *et al.* 2015b).

Just as the goals and aims of biomonitoring differ from site to site, the type of network monitored is likely to also vary, as the ecosystem services and functions they provide are prioritized differently from place to place. There is huge scope for further development in this area, for example, in understanding the extent to which networks can withstand restructuring before the goods and services, which they provide become impaired (e.g. Chapter 5; Tylianakis *et al.* 2010; Thompson, Dunne & Woodward 2012). Some systems show clear signs in their network structure of impending regime shifts which have consequences for ecosystem functioning (e.g. Rawcliffe *et al.* 2010), whereas other networks experience significant network rearrangements without affecting some network metrics (Raffaelli & Friedlander 2012). Thus the interpretation of network data will depend upon the type of system

being monitored as well as the desired ecosystem goods and services.

## 1.4 Network metrics informing ecological research

Traditional ecological network research has focussed on the analysis of basic network metrics such as linkage density (L/S; where L is the number of links, and S the number of nodes), connectance ($L/S^2$), generality (number of resources per consumer), vulnerability (number of consumers per resources) of nodes, food chain length and proportion of basal, intermediate and top nodes (Thompson *et al.* 2012). How these basic metrics vary with the size of the network and with environmental gradients has been extensively studied (Briand 1983; Briand & Cohen 1984; Morris *et al.* 2014). Increasingly though as the quality of the underlying food web data have improved, and analysis methods have advanced, these metrics are proving to be heavily influenced by sampling effort and insufficient to answer the types of ecological question under investigation (e.g. Goldwasser & Roughgarden 1997; Heleno, Devoto & Pocock 2012; Morris *et al.* 2014).

It is important that any metrics used for analysis are rooted in ecological theory, as those which have been traditionally used are, as well as appropriate for addressing the hypothesis in question. Linkage density and connectance are both measures of how well connected a community is, highly connected communities might be more resilient to perturbations because redundant interactions might protect the community from secondary extinctions (Dunne, Williams & Martinez 2002b; Thébault & Fontaine 2010). As such, connectance has been proposed as an important and holistic biological indicator (Gilbert 2009). However, a meta-analysis revealed that there is no evidence that connectance is related to conservation value (Heleno, Devoto & Pocock 2012).

Food chain length indicates the number of times energy has passed from a consumers diet into consumer biomass (Figure 2), between a basal species and a top consumer in

a food web (the 'trophic level'; Williams & Martinez 2004). Food chain length can be measured in a number of ways, Levine's (1980) prey-averaged trophic height is a commonly used variant, and is equal to 1 + the mean trophic level of all the consumer's resource. The calculation and ecological meaning of food chain length and trophic heights is challenged by the ubiquity of omnivory, cannibalism and mutual predation. There are theoretical constraints on the length of food chains, Elton (1927) predicted that trophic levels are limited to be fewer than six, others have suggested that food chains found in natural systems are shorter than you would expect by chance (Pimm 1980; Lawton 1989; Yodzis 1989; Williams & Martinez 2000). However it is unknown to what extent under-sampling has contributed to these findings (Lawton 1989; Huxham, Raffaelli & Pike 1995; Goldwasser & Roughgarden 1997; Marcogliese & Cone 1997).

Generality is a measure of the breadth of a consumer's diet. If a consumer has a specialist (i.e. narrow) diet and is reliant on few resource nodes, then it might be expected to be vulnerable to extinction as the loss of only a few species from the food web might leave it with no resources. Equally if a consumer has low generality it is likely to have low topological importance as it mediates few energy pathways through the food web. Vulnerability is a measure of how many consumers are feeding on a

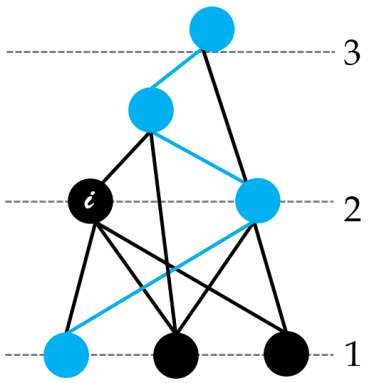



**Figure 2.** An imaginary food web. An example food chain of length three is highlighted in blue. Trophic height (Levine 1980) is shown to the right. The generality of node *i* is three, as it consumes three resource nodes. Its vulnerability is one as it is preyed upon by one consumer.

particular resource species (Figure 2), and therefor how important that node is for the flow of biomass through the food web. Generality and vulnerability scores can be normalised to the size of the food web, allowing values to be compared across different systems, or the standard deviation to be calculated to compare the variability of those scores. Again, sampling effort has been found to strongly influence these metrics, complicating their comparison between different systems (Lawton 1989; Huxham, Raffaelli & Pike 1995; Goldwasser & Roughgarden 1997; Marcogliese & Cone 1997).

Metrics from engineering have begun to enter ecological research, such as measures of network efficiency. The efficiency of a network is a measure of how reachable nodes are from any other node in the network, hence it builds upon the more simplistic measure of connectance (Figure 3). This method reveals information about the substructure of networks, it is more sophisticated than connectance which provides information about the density of connections averaged across the whole network, rather it is a descriptor of how well distributed these interactions are (Figure 3). This method has been applied to measure the global and local (i.e. node specific) efficiency of neural networks (Latora & Marchiori 2001), but is yet to be applied to

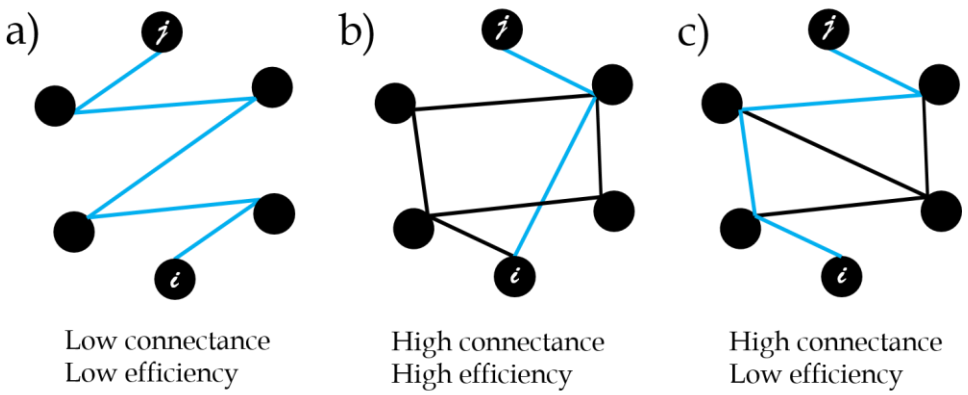



**Figure 3.** Some example networks with high and low global efficiency and connectance. The shortest path between nodes *i* and *j* are highlighted. Connectance is insensitive to the distribution of links in a network, such that both b) and c) have the same connectance score. However all the nodes in b) are within two links of one another, yielding a high efficiency score, whilst the shortest path between nodes *i* and *j* in c) is four, yielding a low efficiency score.

ecological networks. This method is related to the 'small-world' phenomenon, networks which display 'small-world' characteristics have shorter path lengths between nodes than would be expected for a network of that size (Figure 4), which has important implications for the spread of perturbations through the network (Watts & Strogatz 1998; Montoya & Solé 2002). More traditionally used metrics (such as connectance) cannot capture this property. As such the nodes within a network with high efficiency are more highly connected than expected, and are likely to be robust to species and link loss.

The application of advanced complex network analysis techniques to ecological networks is an exciting new avenue of research, many of these new tools may prove useful in increasing our understanding of the structure and dynamics of natural communities, allowing us to make predictions, and design conservation strategies

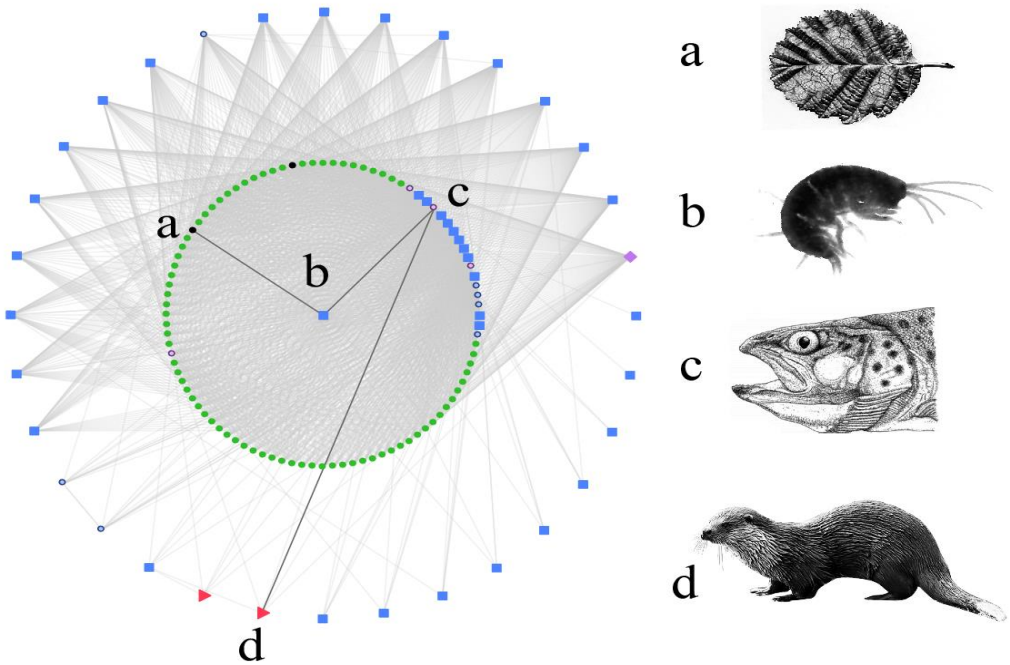



**Figure 4.** An example of a food web with small-world properties. This highly connected food web is dominated by short path lengths between resources and consumers, an example of which is highlighted; a = coarse particulate organic matter (e.g. leaf litter), b = *Gammarus pulex*, c = *Salmo trutta*, d = *Lutra lutra*. The two concentric circles of nodes represent the shortest food web distances to or from *G. pulex*; all nodes in the network are within two links *G. pulex*. Symbols for nodes represent different trophic elements: green circles = producers, blue squares = macroinvertebrates, purple diamonds = vertebrate ectotherms, red triangles = endotherms, black circles = abiotic resources. Reproduced from Thompson *et al.* (2015).

more effectively. Substructural analysis of ecological networks is in its infancy, the analysis of large collections of replicated ecological networks using these novel tools has the potential to provide a far deeper understanding of the response of ecological networks to environmental change. Interdisciplinary collaboration will continue to allow the flow of ideas and novel metrics from other applications of network science, including biomedical research, social networks and information theory, into ecology (e.g. Ulanowicz 2004) to yield ever more sophisticated tools: the challenge now is to adopt and adapt these novel informatics approaches in a well-informed way to add value to biomonitoring.

## 1.5 Aims and thesis structure

The main aim of this project was to apply novel network analysis techniques, taken from other fields of complex network research, to examine the suitability of ecological networks, in particular food webs, as a biomonitoring tool for understanding the impacts of anthropogenic stressors on the environment. To do this, I first developed a function in the R statistical programming language to automate the process of constructing food webs from species lists and trophic information harvested from the literature. I then use this function to combine routine biomonitoring data with information about the trophic interactions between species to build unprecedentedly large collections of food webs spanning environmental gradients. I then analyse these collections using methods taken from complex network research to determine if their structure was sensitive to those environmental gradients.

Chapter 2 – 'Joining the dots: an automated method for constructing food webs from compendia of published interactions'. Here I developed a novel R function which automates the construction of food webs from taxonomic lists, and a dataset of trophic interactions. While researchers have used this approach before, it is typically done by hand, and without a clear output which allows the source of each interaction to be traced. This R function provides an output which can be published alongside the food web stating the source of each interaction, and any diet assumptions which might have been made. I then tested the performance of this method against the traditional method of constructing food webs through analysis of gut contents, as well as some models which could be used to predict food web structure, and found that it matched and often outperformed those models. I then used this method to construct the food webs analysed in Chapters 3-5.

Chapter 3 – 'The recovery of freshwater food webs from the effects of acidification'. Here I augmented typical biomonitoring data which had been collected over the last 24 years in order to monitor the recovery of 23 lake and stream sites from the effects

of acidification, and construct 451 food webs. I examined these food webs to assess how their structure had recovered over time, and what the principal hydrochemical determinants of food web structure were. A first for food web research, I measured the global efficiency of the networks to make inferences about the connectivity of the food webs as they respond to hydrochemical stress.

Chapter 4 – 'Food web topological plasticity disrupts the provision of ecosystem services'. Here I used biomonitoring data, which was collected to assess the impacts of genetically modified crops on farmland biodiversity, to construct a collection of 374 carabid beetle food webs. Carabid beetles are known to regulate the abundance of weed seeds in arable fields, and so provide a pest control service to farmers. I examined the structure of these food webs to assess if the presence of alternative resources, here gastropod prey, interfered with the interactions between carabid consumers and their weed seed resources, and disrupted this pest control service.

Chapter 5 – 'The recovery of a freshwater food web from a catastrophic pesticide spill'. Here I present a study of 8 food webs constructed from samples taken from the River Kennet in Wiltshire, UK, on which there was a major spill of the pesticide chlorpyrifos in 2013. These food webs are more detailed than those built from routine biomonitoring data, and contain mass and abundance data for each node allowing more detailed network analysis to be performed. I examined how the efficiency of energy transfer through these food webs was affected by, and recovered from the pesticide spill. I applied complex network analysis techniques to examine the sub-structure of these food webs, I measured the core size of these food webs. I linked these changes in food web structure to changes in the ecosystem functions they support, and how that recovered over time.

# 2 | Joining the dots: an automated method for constructing food webs from compendia of published interactions[1]


[1] Gray, C., Figueroa, D.H., Hudson, L.N., Ma, A., Perkins, D. & Woodward, G. (2015). Joining the dots: an automated method for constructing food webs from compendia of published interactions. *Food Webs*, **5**, 11–20.

## 2.1 Summary

Food webs are important tools for understanding how complex natural communities are structured and how they respond to environmental change. However their full potential has yet to be realised because of the huge amount of resources required to construct them *de novo*. Consequently, the current catalogue of networks that are suitable for rigorous and comparative analyses and theoretical development still suffers from a lack of standardisation and replication.

Here, we present a novel R function, `WebBuilder`, which automates the construction of food webs from taxonomic lists, and a dataset of trophic interactions. This function works by matching species against those within a dataset of trophic interactions, and 'filling in' missing trophic interactions based on these matches. We also present a dataset of over 20,000 freshwater trophic interactions, and use this and four well-characterised freshwater food webs to test the method.

The `WebBuilder` function facilitates the generation of food webs of comparable quality to the most detailed published food webs, but at a fraction of the research effort or cost. Furthermore, it matched and often outperformed a selection of predictive models, which are currently among the best, in terms of capturing key properties of empirical food webs. The method is simple to use, systematic and, perhaps most importantly, reproducible, which will facilitate (re-) analysis and data sharing. Although developed and tested on a sample of freshwater food webs, this method could easily be extended to cover other types of ecological interactions (such as mutualistic interactions).

## 2.2 Introduction

Characterising food webs (networks representing trophic interactions between species) and other ecological networks (networks which represent any type of ecological interaction, such as pollination) can help us understand and, ultimately predict multispecies systems' responses to changes in environmental conditions (Tylianakis *et al.* 2010; Thompson *et al.* 2012). Food webs can reveal subtle but important changes in the biotic interactions that underpin ecosystem functioning, stability, and resilience to perturbations - higher-level phenomena that cannot be inferred from studying the nodes (i.e., species or populations) alone (Thompson *et al.* 2012; Gray *et al.* 2014).

Despite the many advantages of a network-based approach to ecology, significant challenges need to be overcome, particularly in terms of gathering interaction data. Interactions occur between individuals and data are often collected at this level: for example, via collection, rearing and identification of every leaf miner, and subsequent leaf miner parasitoid along a transect to build herbivore-parasitoid networks (Memmott, Martinez & Cohen 2000; Macfadyen *et al.* 2011), or through dissecting and identifying consumer gut-contents via microscopy (Layer, Hildrew & Woodward 2013). Such laborious methods require substantial investment of time and resources, and it can take many thousands of lab hours to characterise just one food web, which even then may still be undersampled for links between its rarer members (see Table 1; e.g. Woodward *et al.* 2005; Olito & Fox 2014). Many hundreds or thousands of individuals of each species are often needed to fully characterise the full set of feeding links within a food web (e.g. Ings *et al.* 2009), which is rarely practical given the financial and time restraints of research funding. In addition, such comprehensive sampling is often destructive and can impose undesirable disturbance on study systems. Consequently, empirical food webs are often incompletely described and constructed from relatively small sample sizes (Kaiser-Bunbury *et al.* 2010; Layer, Hildrew & Woodward 2013). This limits the conclusions that can be

drawn and the number of comparable food webs that are available both across and within studies (Briand 1983; Bascompte *et al.* 2003; Olesen *et al.* 2007) although exceptions to this exist (Bascompte *et al.* 2003; Cohen & Mulder 2014). Most studies still have patchy and differing levels of sampling effort and taxonomic resolution, making meta-analyses difficult or even inappropriate: the ability to construct large numbers of realistic, comparable food webs across multiple systems would, therefore, help realise the true potential of network approaches (Gray *et al.* 2014).

**Table 1.** Methods for constructing food webs, with their advantages and disadvantages.

| Method | Advantages | Disadvantages | Examples |
|---|---|---|---|
| Observation of evidence of interaction (e.g. feeding trials or gut contents analysis) | High confidence in links produced. | Very slow and labour-intensive.<br>Rare interactions are often missed.<br>Interaction type is biased by the method employed, e.g. the prey of suctorial predators cannot be determined through gut contents analysis. | Woodward *et al.* (2005)<br>Macfadyen *et al.* (2011)<br>Henson *et al.* (2009)<br>Ledger *et al.* (2012) |
| Extrapolating from previously published interactions (e.g. `WebBuilder` function) | Fair confidence in links produced.<br>Rare interactions can be included.<br>Interactions from multiple studies determined through different methods can be easily incorporated.<br>Low effort and quick. | Reliant on the quality of the data contained within the reference dataset.<br>Can only be used to construct 'cumulative' or 'summary' food webs, i.e. temporal or spatial changes in feeding behaviour cannot be incorporated. | Hall & Raffaelli (1991)<br>Goldwasser *et al.* (1993)<br>Havens (1993)<br>Piechnik *et al.* (2008)<br>Pocock *et al.* (2012)<br>Layer *et al.*(2013)<br>Cohen & Mulder (2014)<br>Strong & Leroux (2014) |
| Predictive models | Ecological rules and theory can be incorporated.<br>Low effort and quick. | Require prior knowledge of the structure of the food web in order to optimize parameter values.<br>Many perform poorly at predicting individual interactions, even when food web structure is predicted well. | Cohen *et al.* (1985)<br>Williams & Martinez (2000)<br>Petchey *et al.* (2008)<br>Allesina & Pascual (2009)<br>Allesina (2011)<br>Olito & Fox (2014) |

Ecological networks are often constructed by incorporating species interactions from the published literature (Table 1) and many food webs are constructed entirely

in this manner (Goldwasser & Roughgarden 1993; Havens 1993; Cohen & Mulder 2014; Strong & Leroux 2014), while other food webs contain a blend of observational and extrapolated data (Pocock, Evans & Memmott 2012; Layer, Hildrew & Woodward 2013). By filling in 'missing' trophic interactions to a given species list, the implicit assumption is made that, if a given pair of species have been observed to interact at one site, they will interact in the same way at other sites where they co-occur (at least in terms of a feeding link between the species being realised, or not). Food webs built through this method are often referred to as 'summary' or 'cumulative' food webs as they represent all potential interactions (of a particular type, for instance trophic interactions within a food web) between species of a particular community, rather than a snapshot in time. As such, food webs built through this method are unsuitable for detecting changes in species feeding behaviour across sites or over time, but are highly effective for detecting broad macro-ecological trends such as changes in food web structure across environmental gradients (Piechnik, Lawler & Martinez 2008; Mulder & Elser 2009; Layer *et al.* 2010b).

This approach can be taken further, by assigning interactions of species on the basis of taxonomic similarity: *i.e.*, species within the same genus are assumed to have identical links if a link has been established through direct observation for at least one congener (Goldwasser & Roughgarden 1993; Layer *et al.* 2010b). This process is often used when constructing summary food webs for species the interactions of which have not been fully characterised (e.g., as revealed from yield-effort curves) to minimise potential biases arising from under-sampling, i.e. including *only* observed links would otherwise significantly underestimate food web complexity, especially among the rarer and/or more obscure taxa (Woodward et al 2010). Recent work (Eklof *et al.* 2012) has provided justification for this approach, by highlighting the strong influence that taxonomy has in determining the structure of food webs. Thus, given the prevalence of undersampling in even relatively well-described food webs, dietary data extrapolated from the literature and generalised taxonomically can potentially

produce far more complete and realistic summary food webs than those that rely solely on observations made in a particular locale.

Despite the prevalence of these methods for constructing summary food webs in the literature (Goldwasser & Roughgarden 1993; Havens 1993; Layer *et al.* 2010a; Pocock, Evans & Memmott 2012; Cohen & Mulder 2014; Strong & Leroux 2014), there is still no standard method for inferring feeding interactions, resulting in inconsistencies among studies, even within the same ecosystem type. This is especially problematic because authors rarely state explicitly which links have been observed or extrapolated, or the source from which they have been drawn, or how closely the previously published interactions match those reported in their particular study, making replication impossible and preventing other researchers from scrutinising published interactions fully (but see Strong & Leroux 2014).

Recent research has sought to develop predictive models of the structure of ecological networks (Rohr *et al.* 2010; Eklof *et al.* 2012; Gravel *et al.* 2013; Olito & Fox 2014). Simple rules based on ecological theory have been used to model and predict the structure and topology of food webs, the most successful of which include deterministic models based on information on species' body sizes, for example the 'Difference', 'Ratio', and 'Difference/Ratio' models (Allesina 2011) and the Allometric Diet Breadth Model (ADBM; Petchey *et al.* 2008) which incorporates allometric scaling and optimal foraging parameters. Whilst these models have been developed primarily to advance ecological theory, they provide a possible tool through which food webs could be built *de novo* in order to address questions about network structure across environmental gradients or scales. However, to achieve their best performance (proportion of correctly predicted links) these models require some prior knowledge about the number of links in the network. For instance, for the models mentioned above a researcher is required to go through a parameter optimisation procedure, by fixing the number of links, values of constants and exponents can be derived, by maximizing the number of links correctly predicted. When constructing a network for

the first time for a particular system, a researcher would be required to fix the number of links to an expected value which would bias the network structure towards that which the researcher expected to find.

Additionally, recent work (Olito & Fox 2014) has highlighted that while predictive models might perform well at predicting metrics of network structure, they tend to perform poorly at predicting pairwise interactions (Vázquez, Chacoff & Cagnolo 2009; Verdú & Valiente-Banuet 2011; Sáyago *et al.* 2013; Vizentin-Bugoni, Maruyama & Sazima 2014), so whilst they may predict network structural metrics well, they are doing so for the wrong reason as the underlying biological mechanisms have not been fully incorporated into the predictive models (Petchey *et al.* 2011). To the best of our knowledge, the models used here have not, up until now, been used to predict network structure *de novo*, as this is not the scenario for which they were developed.

Given the limitations of constructing food webs from observation of interactions or predictive models, we need an automated, repeatable and reliable method of building local food webs that can be applied across studies and, ultimately, different ecosystem and network types. Here, we introduce a method, the `WebBuilder` function that assembles food webs by systematically assigning links for taxa based upon a given set of user-defined rules applied to a dataset of known trophic interactions. We provide an implementation of our method for the R statistical modelling language (R Core Team 2013), building upon the methods and data structures provided by the Cheddar R package (Hudson *et al.* 2013). We tested the method on four highly resolved freshwater food webs which have had their interactions characterised through gut contents analysis, as these represent some of the most complete food webs described to date, as a test case for our proof-of-concept. Specifically, our key aims were to:

1. Collate a dataset of trophic interactions in a standard format to act as an example system in which to test this method.

2. Automate the process of constructing food webs from this reference dataset in a repeatable and reliable manner.
3. Compare the performance of this method with the structure of food webs with 'known' interactions, i.e. those which have been built through observation of the interactions.
4. Compare the performance of this method with another way of predicting food web structure; the ADBM, Difference, Ratio and Difference/Ratio models.

## 2.3 Methods

### 2.3.1 Dataset of trophic interactions

We collated a dataset of 20,823 pairwise trophic interactions among species (or the next highest level of resolution available, usually genus), from 51 different data sources, most of which were primary literature (Table A.1, online supporting material). It contains trophic interactions between primarily UK freshwater species, including 203 producer taxa, 593 invertebrate taxa, 24 fish taxa, 10,348 producer-animal links, 9,531 animal-animal links and 944 detritus-animal links. When the necessary data were not available in the original publications, we contacted the authors directly, where possible, to obtain the raw data. The taxonomy of every resource and consumer has been standardised through the Global Names Resolver (http://resolver.globalnames.biodinfo.org/) using the Global Biodiversity Information Facility dataset. For every resource-consumer link the taxonomy (species, genus, subfamily, family, order, class) of both is given, along with life-stage information, if relevant, and a literature reference for the source of the link. This dataset builds upon the collection assembled by Brose et al (2005), and to the best of our knowledge, represents the largest standardised collection of trophic links for freshwater organisms. This dataset is available to download at https://sites.google.com/site/foodwebsdataset/ (doi: 10.5281/zenodo.13751) and is

designed to be easily updated by the iterative addition of new data (details of how to submit new data to the dataset are given on the website), allowing its content to improve over time, in an analogous manner to molecular-based bioinformatics datasets. New data will be subjected to a quality assurance procedure prior to inclusion in the dataset. Specifically, all taxa will be parsed through Global Names Resolver (http://resolver.globalnames.biodinfo.org/) using the Global Biodiversity Information Facility dataset. Additionally the data will be eyeballed for irregularities. It is anticipated that these data will exist as an open access resource, and as such the community of researchers who access it will report any errors they find so they can be double checked and removed. New iterations of the dataset can be produced, hosted on the webpage alongside the original, and assigned a new DOI, allowing researchers to cite exactly which version of the dataset they have used for their research, allowing analyses to be repeated using identical versions to those cited in a given study, if required in the future.

### 2.3.2 The `WebBuilder` function

The method of constructing ecological networks by extrapolating from previously published interactions is implemented in a new R function - `WebBuilder` (see online supporting material for code). The user is required to provide the following; firstly a list of taxa (i.e. nodes) in the community of interest (step I,

Figure **5**), this data can be gathered from multiple sources and could be in the form of survey or biomonitoring data. Secondly, for each node, the minimum level of taxonomic generalisation (explained below; step II), and the taxonomic classification of each node (step III). Lastly a registry – a dataset of known trophic interactions, including taxonomic classification (step IV), an example of which is published here, but which can also be created by the user or obtained elsewhere. It is recommended that the user resolve the taxonomy of their taxa list and registry using the same procedure so as to ensure that taxa are matched correctly, if the user were using the

registry provided here they would need to parse their taxa list through Global Names Resolver (http://resolver.globalnames.biodinfo.org/) using the Global Biodiversity Information Facility dataset. The function searches the registry for every possible combination of resource-consumer interactions (for N taxa there are $N^2$ possible trophic interactions) which match the provided taxa list given the specified level of taxonomic generalisation.

The minimum level of taxonomic generalisation determines the taxonomic rank at which matches are made, thus generalising the resources or consumers of the candidate node to the species, genus, subfamily, family, order etc level, as specified in the input (step II,

Figure **5**). For instance, a researcher might decide to ascribe the level of taxonomic generalisation of 'genus' to the mayfly *Baetis fuscatus*, allowing it to be matched with the more commonly studied species *Baetis rhodani* in the dataset, and take on the

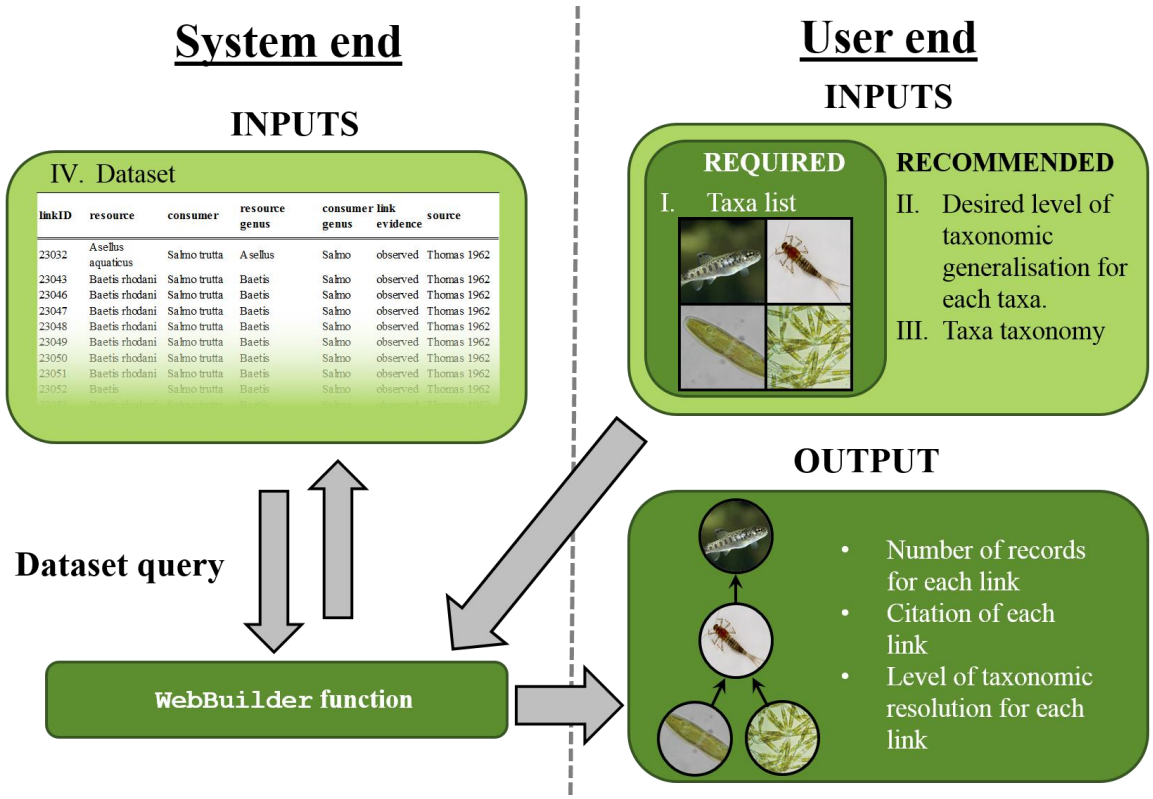



**Figure 5**. A simplified workflow demonstrating the `WebBuilder` function. For a workable example see online supporting material.

appropriate feeding interactions of that species, i.e. those which include taxa also present on the provided taxa list (see the first Scenario in Figure 6). This level of taxonomic generalisation is selected based on knowledge of a candidate node's trophic interactions in relation to its sister taxa (i.e. if all members of the same taxonomic unit can be assumed to have the same trophic interactions or not), and this can be tailored depending on the resource/consumer status of the node. For example, consumers of the larvae of the non-biting midge subfamily Tanypodinae tend to be trophic generalists and would likely consume other larvae of the family Chironomidae, while it is not likely that the resources of Tanypodinae larvae (which are predominantly predatory) would be shared by all Chironomidae larvae (many of which are grazers or filter feeders). Hence it would not be appropriate to assign the trophic generalisation level 'family' to both resource and consumer interactions of Chironomidae. Instead a researcher might ascribe the 'resource method' for Tanypodinae as 'family', but the 'consumer method' as 'subfamily' (see the second Scenario in Figure 6). The function output contains references to the original empirical links, the number of matches that were found and the taxonomic level at which those matches were found, so links can be additionally screened and scrutinised *post hoc,* and analysis can be repeated easily because the function output contains the necessary information. Example R code is supplied (see online supporting material).

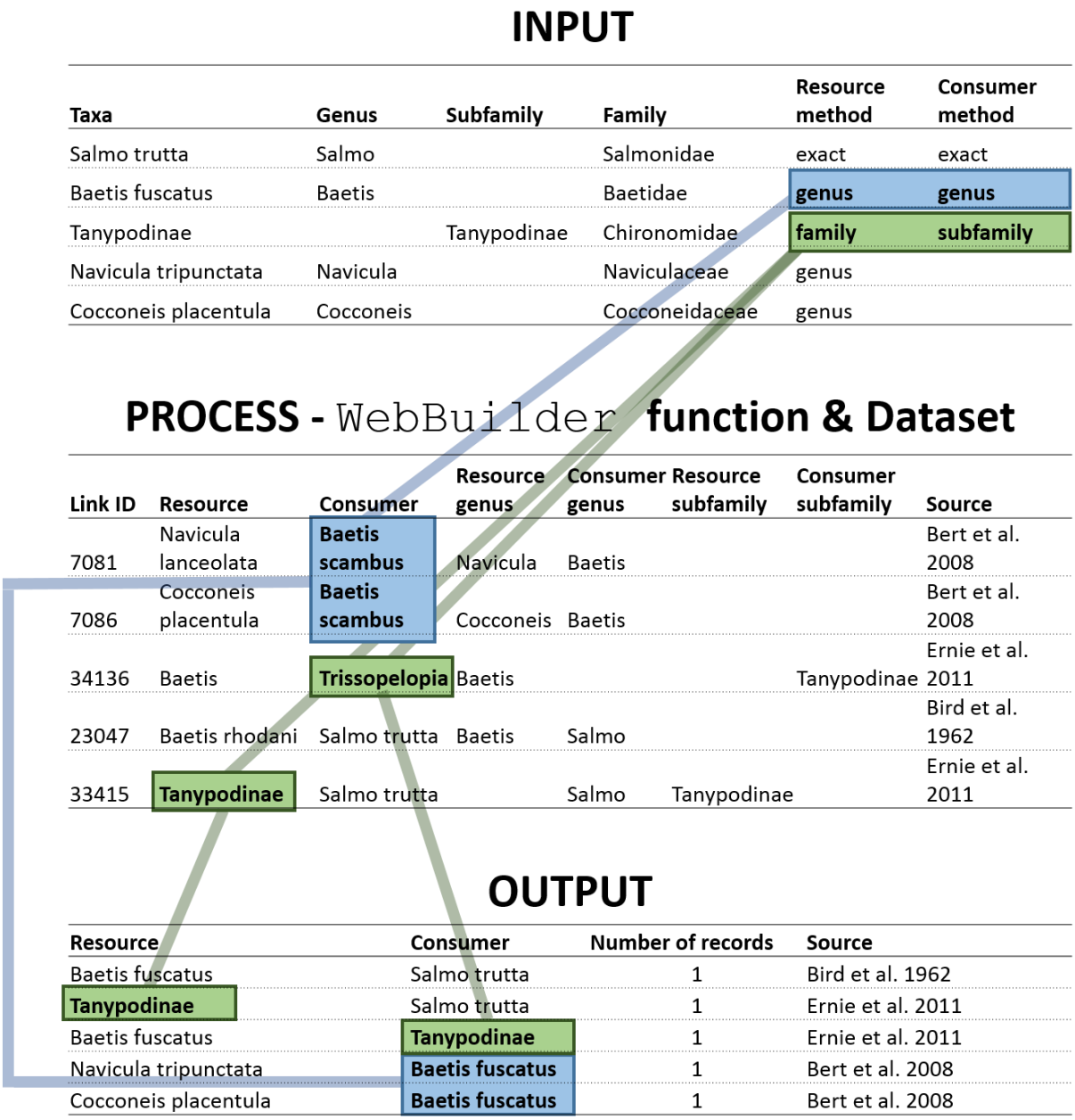



**Figure 6**. An example of inputs and outputs for the `WebBuilder` function. Two different scenarios are highlighted. Firstly in blue the taxa *Baetis fuscatus* is generalised to the genus level for both its consumer and resource links, this allows it to be matched with *B. scambus* in the registry and the *Navicula tripunctata* – *B. fuscatus*, and *Cocconeis placentula* – *B. fuscatus* links to be included in the output. Secondly in green, the taxa Tanypodinae are generalised to the family level for its resource links and subfamily level for its consumer links, allowing it to matched with all entries in the registry with the subfamily Tanypodinae and the Tanypodinae – *Salmo trutta*, and *Baetis fuscatus* – Tanypodinae links to be included in the output.

### 2.3.3 Comparing the `WebBuilder` function with empirical food webs

The `WebBuilder` function was validated on a collection of highly-resolved stream food webs which have had their trophic interactions characterised through direct observation; Broadstone Stream (Woodward *et al.* 2010b), Afon Hirnant (Woodward *et al.* 2010b; Gilljam *et al.* 2011), Tadnoll Brook (Edwards *et al.* 2009) and the summary food web for the replicated four reference Mill Stream side-channels (Ledger *et al.* 2012; Woodward *et al.* 2012a). The replicates for the Mill Stream data were aggregated to aid comparison with the other food webs, which were all constructed as a single summary food web. The Broadstone and Afon Hirnant food webs contained only trophic interactions between macro-invertebrates, the Tadnoll food web contained interactions between macro-invertebrates and fishes and the Mill Stream data contain interactions between macroinvertebrates, algae and detritus. When the `WebBuilder` function was used to generate the empirical food webs, in turn each respective local dataset was first removed from the global dataset, so each food web was generated in the absence of its own link information (to remove circularities).

The performance of the `WebBuilder` function was evaluated by calculating the True Skill Statistic (TSS; Allouche, Tsoar & Kadmon 2006). This statistic was used as it can be seperated into its component parts to provide information on the types of differences between the empirical and generated food webs, and builds upon the most commonly used metric which is simply the proportion of links correctly generated (Petchey *et al.* 2008; Woodward *et al.* 2010b; e.g. Allesina 2011). This statistic was chosen over likelihood based approaches because we were not interested so much in the efficiency of these predictive models, more the biological realism of the generated food webs (Petchey *et al.* 2011). The TSS is calculated from the following formula:

$$TSS = (ad - bc)/[(a + c)(b + d)]$$

where $a$ is the number of links which were correctly generated by the function (the True Positives Rate; TPR), $b$ the number of links generated by the function but not

observed empirically, *c* the number of links not generated by the function but were observed empirically and *d* the number of links neither generated by the function nor observed empirically. TSS score values range from -1 to 1, where a score of -1 represents a generated food web that is the inverse of the empirically observed one (no observed empirical links are seen in the generated food web, and every non-link in the empirical food web is present in the generated food web), and 1 representing a generated food web having the exact same links as the empirically observed one.

Each empirical food web was generated using the level of taxonomic generalisation considered most appropriate (see online supporting material), this mostly consisted of exact and genus level matches although some family and order matches were used. To test how the generated food webs compared to their empirical counterparts, a series of network metrics were calculated; number of links (L), linkage density (L/S; where S is the number of nodes), connectance (C, where $C=L/S^2$), generality (the average number of resources per consumer), vulnerability (the average number of consumers per resource), and proportion of top, intermediate and basal nodes (with cannibalistic links removed). The difference between the generated and empirical network metric was tested with paired Wilcoxon signed rank tests.

To test how the quality of the generated food webs varied with dataset size, the dataset was randomly subsampled, in sequential steps of 5% from 5-100%, of the original dataset size, and then used to generate each food web. Each subsample size was repeated five times and each empirical food web was generated in the absence of its own food web data as above, to remove circularities. For each node within each network, the same level of taxonomic generalisation was used as above.

To test how the quality of the pairwise interactions generated by the `WebBuilder` function varied with the level of taxonomic generalisation, each food web was built using exact, genus, family or order taxonomic generalisation for all nodes. The degree (the number of links into or out of a particular node), generality, and vulnerability for every node in the generated food web was compared with that

in the empirical network. The difference between the two for every node was recorded so that a positive score represented interactions 'missed' by the `WebBuilder` function, and a negative score represented 'extra' interactions not found in the empirical food web. The distribution of these scores gives an indication of how well the `WebBuilder` function predicted pairwise interactions across the whole network: i.e., if, on average, it tended to 'miss' more interactions, or tended to pick up 'extra' interactions. To test if the mean was different from zero (indicating no difference in the quality of pairwise interactions between the generated and empirical food webs) a one sampled t-test was used.

### 2.3.4 Comparing the `WebBuilder` function to theoretical food web models

The performance of the `WebBuilder` function was compared with examples of some of the best-performing predictive models currently available: the 'Difference', 'Ratio' and 'Difference/Ratio' models (Allesina 2011) as well as the Allometric Diet Breadth Model (ADBM; with 'ratio' handling time, Petchey *et al.* 2008). The 'Difference', 'Ratio' and 'Difference/Ratio' models all generate food web links on the basis of body size, (either the difference between consumer body size and resource body size, the ratio between the two, or the difference multiplied by the ratio). The ADBM builds on this and incorporates allometries of body size and foraging behaviour of individual consumers to model food web structure (see Petchey *et al.* 2008; Allesina 2011 for more detailed explanations). Detritus nodes were first excluded from the Mill Stream food web because these nodes had no body size or abundance data. For the 'Difference', 'Ratio' and 'Difference/Ratio' models two parameters required optimisation, $a$ and $b$. For the ADBM we used parameter values for the mass to attack rate constant ($a$), resource mass to attack rate exponent ($a_i$) and consumer mass to attack rate exponent ($a_j$) from the literature (Rall *et al.* 2012) rather than through parameter optimisation as in Petchey *et al.* (2008), so as to simulate a situation for which the `WebBuilder` function was designed, where food webs are being generated for the first time with

no prior knowledge of the system other than the species richness. For two parameters (mass to handling time constant, *h.a;* mass to handling time critical ratio, *b*) we were unable to find information in the literature with which to value these parameters, so went through the process of optimisation. For all models this was achieved by constructing food webs with a range of values for each parameter, and selecting those food webs which had a number of links that was within the range set by the `WebBuilder` function, i.e. if the `WebBuilder` function generated *K* links, and there were *L* empirical links and *K-L=t* we selected all possible solutions within the range *L-t:L+t*, to make the comparison with the `WebBuilder` function fair. Note that for some food webs the difference between *L* and *K* was large, leading to large variation in the food web sizes generated by these models. Indeed for the Afon Hirnanlt food web this range fell below zero, and so the range was arbitrarily set to be the same proportional size as that of the Tadnoll food web, which had the next highest range. Parameter optimisation was conducted without using the connectance of the empirical food webs, hence although the same data have been used, results will vary from previous publications. Prior knowledge of the connectance of food webs would not be possible if a food web were being built *de novo*, so here we are using these models in a different way from their original application.

## 2.4 Results

### 2.4.1 Comparing the `WebBuilder` function with empirical food webs

When we constructed food webs from random subsets of the dataset, the quality (as measured by TSS scores) of the generated food webs improved as the number of records in the dataset increased, allowing more complete resource and consumer interactions to be ascribed to each taxa (Figure 7). The strength of this relationship was food web specific, for instance Broadstone and Afon Hirnant did not continue to improve beyond a dataset size of about 25%. These food webs are

relatively simplistic compared to Tadnoll and Millstream, and so the `WebBuilder` function reached its optimum performance when generating these food webs with a fraction of the total dataset. Tadnoll and Millstream did not reach their asymptotes suggesting that more data are needed to improve upon the quality of their generated food webs.

The level of taxonomic generalisation for each node was important for the quality of the generated food web; if the taxonomy of a given node list was generalised too far (typically beyond the family level) then the ascribed links became unrepresentative and the food web become over-connected resulting in an increased FPR and lower TSS score (Figure A.1, online supporting material). At the scale of

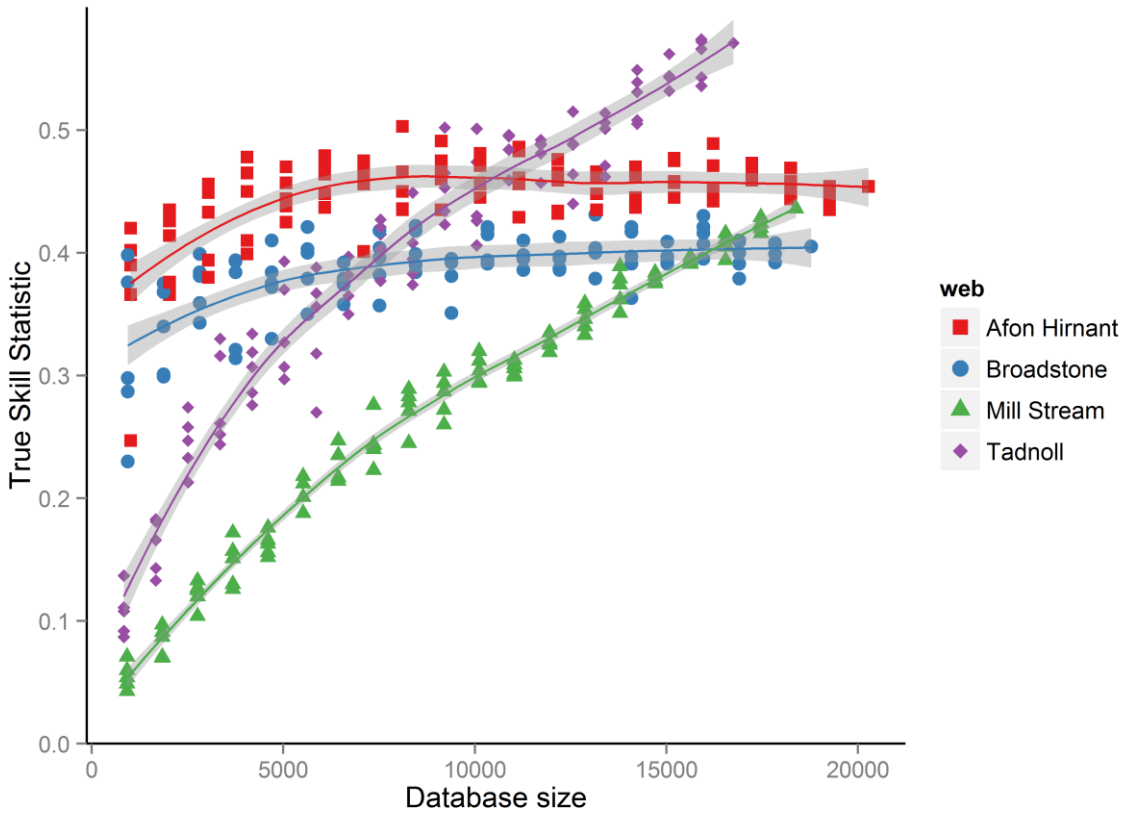



**Figure 7.** The quality of the generated food web increases with the size of the dataset. Fluctuations in the TSS score are caused by changes in the component parts of the TSS, i.e. while the TPR may increase as the dataset size increases, other metrics such as the FPR might also increase, causing the total TSS to fall (see methods). Lines are fitted using a LOESS smoother (Cleveland et al. 1992), grey shading indicates the 95% confidence intervals.

individual trophic interactions, the difference in degree, generality and vulnerability was generally positive when matching taxa exactly or at the genus level, and becomes progressively more negative as the taxonomic generalisation increased, indicating that the `WebBuilder` function was 'missing' links when matching nodes exactly or at the genus level, and including progressively more links the further the taxonomy was generalised (Figure 8). For Afon Hirnant and Tadnoll there was no significant difference in the generality of consumers between the generated and empirical food webs when taxa were matched at the genus level, and there was no significant difference in vulnerability of resources for the Tadnoll food web when matched at the genus level. This suggests that matching taxa at the genus level for these food webs produces the most 'accurate' pairwise interactions. For all other food webs and levels of taxonomic generalisations the generated links were different from that of the empirical food web (Figure 8).

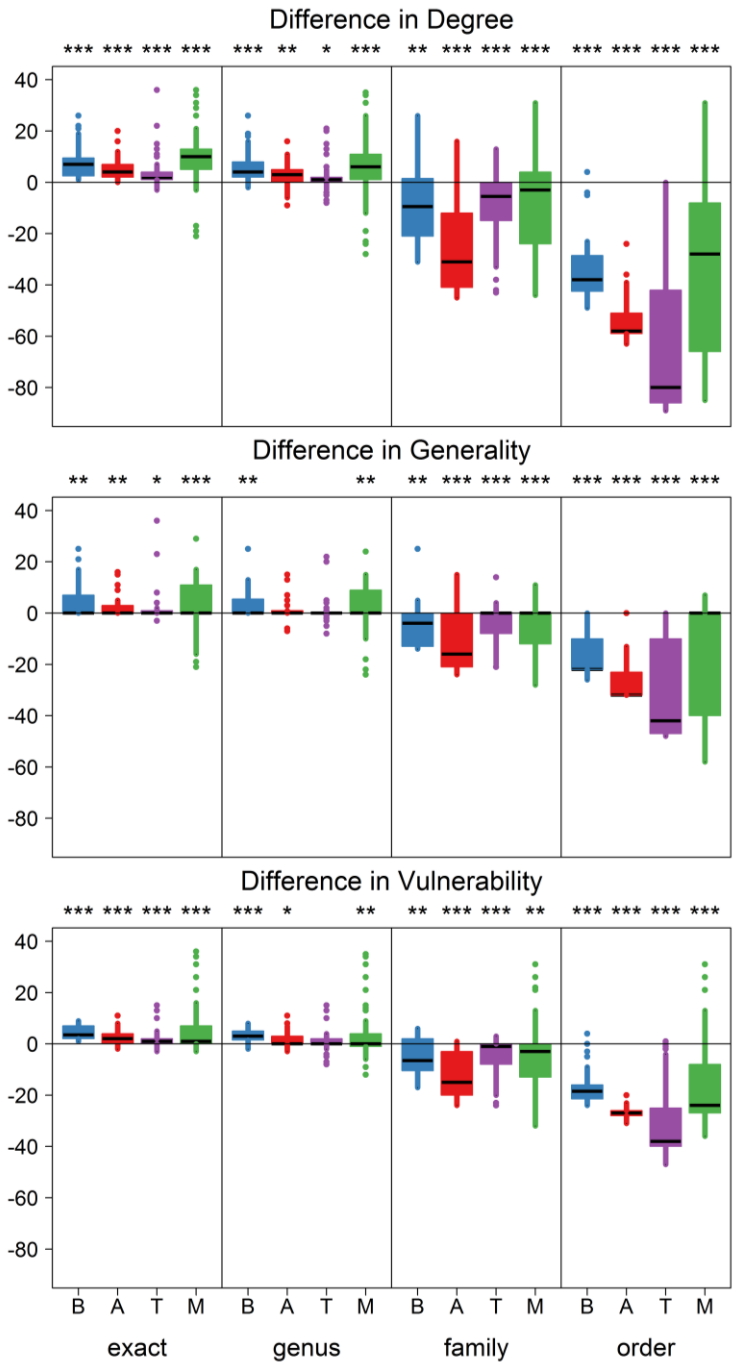



**Figure 8**. Box plots showing the changes in generated trophic interactions as the level of taxonomic generalisation is varied. The difference in degree (top), generality (middle) and vulnerability (bottom) of individual nodes between the generated and empirical food web, thus the sample size reflects the number of nodes in the empirical food web. Positive values represent links which were 'missed' by the `WebBuilder` function, while negative values represent additional links not found empirically. Box plots are colour coded: Broadstone (B; blue), Afon Hirnant (A; red), Tadnoll (T; purple) and Mill Stream (M; green).Stars indicate if the mean is different from zero (one sample t-test) and indicate if the generated trophic interactions are different from that of the empirical food web, 0.05 >= p > 0.01 = *, 0.01 >= p > 0.001 = **, p <= 0.001=***.

The occurrence of each food web's nodes in the dataset is given in Table 2. The coverage of these within the reduced dataset (food web data were removed from the dataset when used to generate the food web for that site) varied between 1,497 (Broadstone) and 6,704 occurrences (Mill Stream) (Table 2). Even at the family level some nodes from Broadstone, Afon Hirnant and Mill Stream were not represented in the dataset, meaning that those nodes needed to be generalised further still for the `WebBuilder` function to generate their links. These nodes tended to be rare taxa which were poorly represented in the dataset. The generated food webs (see online supporting material for the generated trophic links) had similar network metrics to the empirical food webs (Table 3), although the proportion of top nodes was consistently lower in the generated food webs, and the proportion of intermediate and basal nodes was consistently higher. All generated food web metrics were found to be not significantly different to that of their empirical counterparts (paired Wilcoxon signed rank test, p=>0.05). The TSS ranged from 0.405 (Broadstone) to 0.571 (Tadnoll Brook). Two food webs (Broadstone and Afon Hirnant) contained nodes that were found to have predatory links in the empirical food web but were not predicted to have any by the `WebBuilder` function, and *vice versa* many nodes distributed across all the food webs were predicted to have consumer links but were not found to have any empirically (Figure 9).

**Table 2.** The representation of the food web taxa within the full dataset and partial dataset (i.e., diet data gathered from a food web were excluded from the generation of its own inferred food web).

| Food web | Number of appearances in dataset | | Percentage of nodes appearing in partial dataset at each taxonomic level | | |
|---|---|---|---|---|---|
| | Full data set | Partial data set | exact | genus | family |
| Broadstone | 2,196 | 1,497 | 81% | 84% | 94% |
| Afon Hirnant | 2,945 | 2,266 | 72% | 79% | 92% |
| Tadnoll | 12,405 | 4,314 | 84% | 91% | 100% |
| Mill Stream | 11,545 | 6,704 | 87% | 96% | 97% |

**Table 3.** The number of links (L), linkage density (L/S, where S=number of nodes), the connectance (C, where C=L/S2), Generality, Vulnerability, proportion of top, intermediate and basal species of the empirical and generated food webs. The performance of the WebBuilder function (relative to the original empirical food web) is summarised by the TSS statistic (which gives an overall measure of performance), and TPR (the proportion of links correctly generated). All food web metrics for the generated food webs were found to be similar to that of their empirical counterparts (paired Wilcoxon signed rank test, p = >0.05)

| Network | | L | L/S | C | *Generality* | *Vulnerability* | *Top* | *Intermediate* | *Basal* | TSS | TPR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Broadstone | empirical | 124 | 4.43 | 0.158 | 13 | 4.33 | 0.68 | 0.29 | 0.04 | 0.405 | 0.376 |
| | generated | 194 | 6.93 | 0.247 | 9.84 | 7.48 | 0.04 | 0.64 | 0.25 | | |
| Afon Hirnant | empirical | 93 | 2.82 | 0.085 | 7.67 | 4 | 0.58 | 0.12 | 0.24 | 0.454 | 0.24 |
| | generated | 250 | 7.58 | 0.23 | 15.87 | 8.5 | 0 | 0.45 | 0.39 | | |
| Tadnoll | empirical | 169 | 2.91 | 0.05 | 9.33 | 3.23 | 0.69 | 0.21 | 0.1 | 0.571 | 0.372 |
| | generated | 285 | 4.91 | 0.085 | 12.45 | 5.59 | 0.07 | 0.31 | 0.53 | | |
| Mill Stream | empirical | 680 | 9.19 | 0.124 | 16.98 | 14.15 | 0.46 | 0.19 | 0.35 | 0.436 | 0.531 |
| | generated | 639 | 8.64 | 0.117 | 16.63 | 11.7 | 0.19 | 0.32 | 0.41 | | |

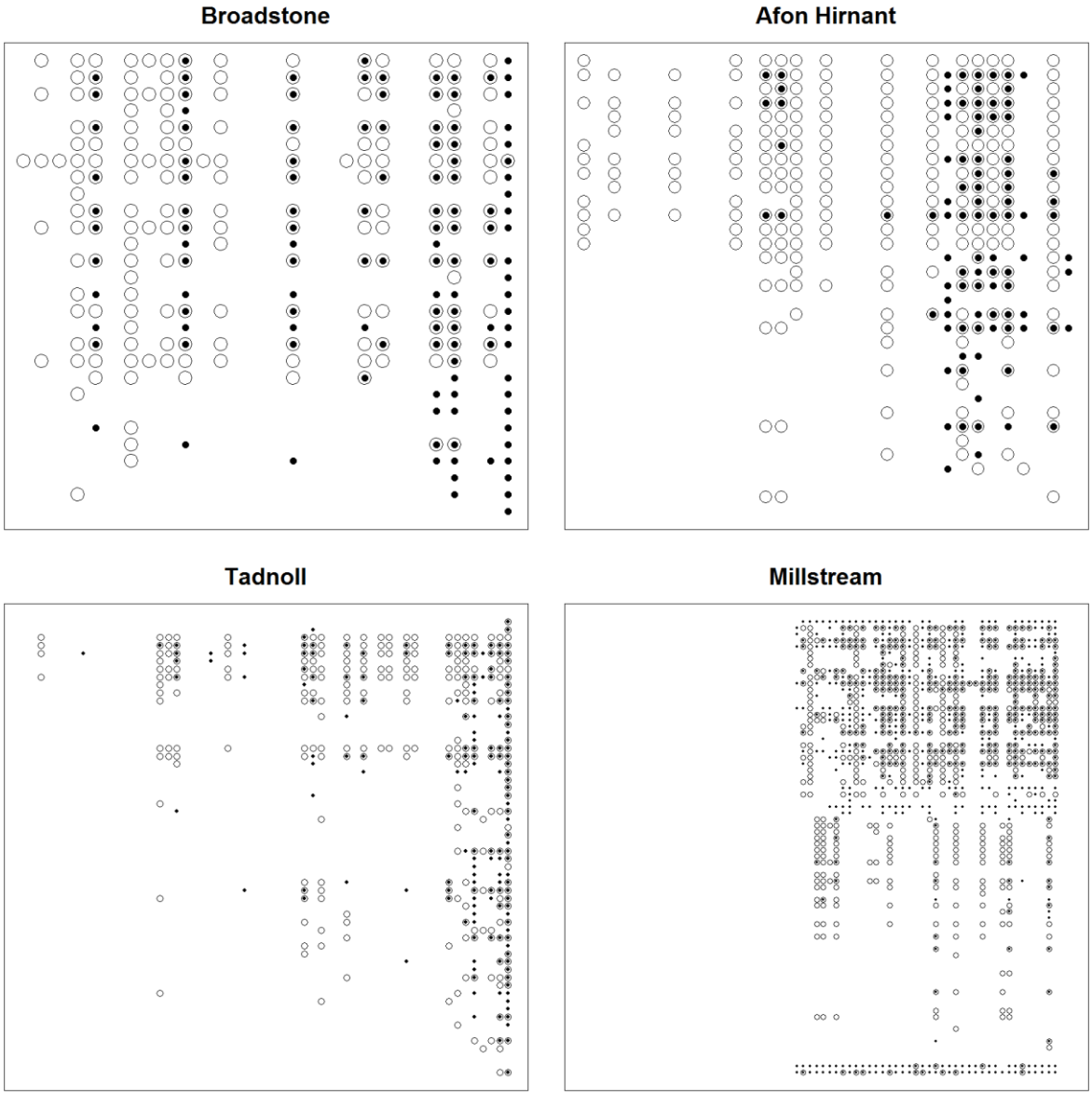



**Figure 9**. Predation matrices for the empirical food webs compared to those generated by the AR method. Nodes are ordered by increasing body mass. A trophic link is represented by a point indicating that the taxon in that column consumes the taxon in that row. Links generated by the `WebBuilder` function are represented by empty circles, and those found empirically are represented by smaller, filled circles.

### 2.4.2 Comparing the `WebBuilder` function method to theoretical food web models

The percentage of links correctly predicted (TPR) by the ADBM ranged from 12-43%, for the Difference model it was 0-46%, Ratio model it was 0-51% and the Difference/Ratio it was 0-54% (Figure 10). All webs generated by the `WebBuilder` function had higher TPR and TSS scores than the median values for the Difference, Ratio and Difference /Ratio models (Figure 10). In general the `WebBuilder` function had higher TPR and TSS scores than the ADBM, however the TPR score for Tadnoll and Afon Hirnant were similar to the median ADBM TPR score for that particular food web (as opposed to the overall median). Additionally the TSS score for Tadnoll generated by the `WebBuilder` function was similar to the median ADBM TSS score (Figure A.2, online supporting material).

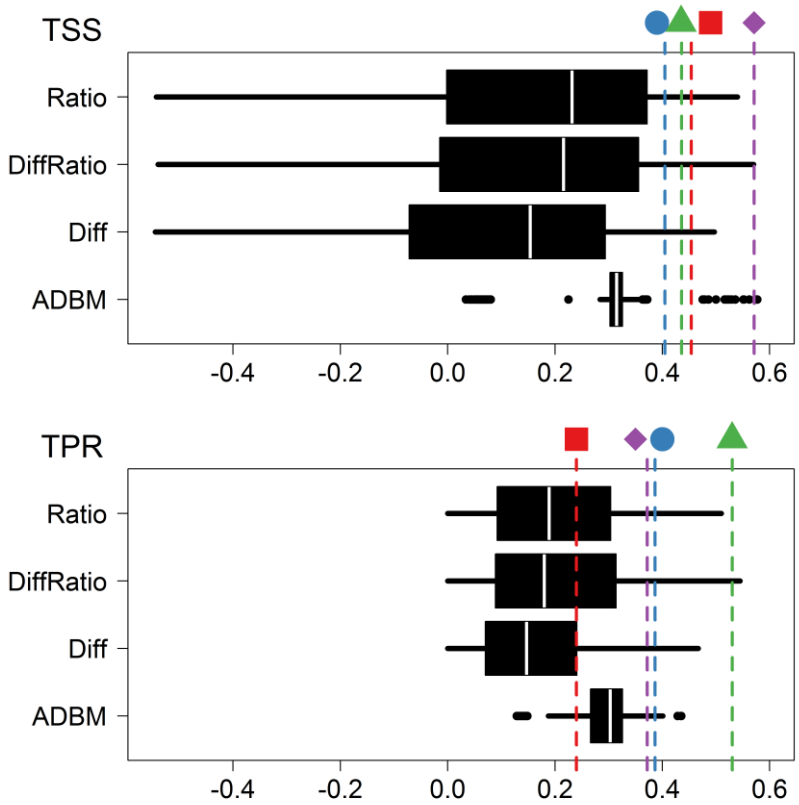



**Figure 10**. Box plots showing the performance of the `WebBuilder` function compared to the ADBM, Difference, Ratio and Difference/Ratio models. The performance of the `WebBuilder` function is plotted as four vertical lines, one for each of the empirical food webs; Broadstone ●, Afon Hirnant ■, Tadnoll ◆ and Mill Stream ▲. The TSS score (top panel) gives an overall measure of the performance of the predictive method relative to empirical food webs, and varies between 1 (a generated food web that is exactly the same as the empirical food web) and -1 (a generated food web which is the exact inverse of the empirical food web). The TPR (True Positives Rate; bottom panel) is the proportion of generated food web links that were also found empirically, and varies between 0 (no links generated correctly) and 1 (all links generated correctly). A box plot of each set of values is given, indicating the range, quartile ranges and median of each set of values. For the `WebBuilder` function only the individual scores for the four food webs are shown, for all others there are too many generated scores to be shown individually; ADBM (n=508), Difference (n=32,025), Ratio (n=43,638) and Difference/Ratio (n=41,602).

## 2.5 Discussion

### 2.5.1 Strengths and weaknesses of the `WebBuilder` function

Here we have demonstrated a systematic and reproducible method for building ecological networks from compilations of previously observed interactions. The `WebBuilder` function facilitates comparability across studies, re-analysis and data sharing. Although developed in the context of freshwater food webs, given its simplicity and generality the `WebBuilder` function could be easily applied to other systems, such as terrestrial food webs or even mutualistic networks. Plenty of other datasets already exist which could be exploited similarly to produce comparable, reproducible networks from marine and terrestrial systems (e.g., Barnes 2008; Database of Insects and their Food Plants; http://www.brc.ac.uk/dbif/).

The `WebBuilder` function is an effective tool for constructing summary ecological networks for the first time. The overall performance (TSS) of the `WebBuilder` function exceeded that of the ADBM, Difference, Ratio and Difference /Ratio models. The proportion of correctly predicted links (TPR) was similar to or exceeded that of the ADBM. The ADBM cannot predict links for nodes that have no body-size information – either because it is not known or because the concept is meaningless for the node, such as detrital resources. This problem does not apply to the `WebBuilder` function. The ADBM, Difference, Ratio and Difference/Ratio models have been used to generate the food webs presented here before, and have performed better than we have achieved here (Petchey *et al.* 2008; Woodward *et al.* 2010b; Allesina 2011), however to achieve this accuracy the generated food webs were constrained to have the same connectance as the empirical food webs, an approach not available when building a food web for the first time. Indeed there were instances here that the TPR and TSS of modelled food webs exceeded that of the `WebBuilder` function, but from the range of possible food webs generated by these models, it is impossible to select the most 'accurate' one without knowledge of the expected number of links. The `WebBuilder` function does not rely on prior knowledge of the

food web, only on the correct identification of the nodes, thus reducing biases and restrictions.

It is perhaps unfair to compare the performance of the `WebBuilder` function to that of the ADBM, Difference, Ratio and Difference/Ratio models due to the inherent differences in the mechanisms through which they operate, and indeed it is not our intention for this exercise to be taken as a criticism of these alternative approaches. Rather, we have compared them here in order to place the `WebBuilder` method in the broader context of some of other more widely-used predictive methods currently available. Comparing our approach with the performance of the ADBM essentially represents a test of, and a means of improving, our understanding of the mechanistic theory behind these trophic interactions. Comparing the food webs produced by our approach with empirical food webs represents a test of the quality of the underlying data held within the dataset of trophic interactions. The `WebBuilder` function should be used as tool with which to construct large collections of food webs with which to test our understanding of food web structure across environmental gradients. The `WebBuilder` function is particularly suited to constructing food webs for data-poor systems, e.g., where there is no information available about the abundance or body size of nodes, with the only information being a list of species present. Clearly the ADBM or other predictive models would not be suited to these conditions, as they were never designed to work in this way. The `WebBuilder` function, however, would be able to generate reasonably realistic food webs if given a reference dataset of relevant trophic interactions. The `WebBuilder` function is adaptive, and can be improved upon over time; for instance, by increasing the size and coverage of the dataset of interactions. Hence it requires a substantial amount of data to perform well, unlike the predictive models analysed here. These types of methods can be viewed as complementary: a researcher might use both in conjunction in order to harness the advantages of both to better predict food web structure: indeed we envisage combining the `WebBuilder` function and other

predictive approaches in parallel to build and understand food webs.

The four food webs presented here are among the most highly resolved and complete freshwater food webs published to date, yet the links are still under-sampled for many nodes (Woodward *et al.* 2010b), due to methodological issues and logistical constraints on sampling effort. The `WebBuilder` function can help to overcome these issues. Firstly, it can take many hundreds of individuals to characterise a species diet (Ings *et al.* 2009) and thus the interactions between rare consumers and rare resources are often under-sampled. The `WebBuilder` function helps to overcome this as rare interactions need only be observed once in the dataset of previously published interactions in order to be incorporated into applicable food webs as they are constructed: i.e., potentially the "global diet" of a species is held within the dataset, and can be expanded in future data collections. Secondly, the method of observing interactions often limits the types of interactions which can be characterised; for instance, the prey of suctorial predators (which are especially common in terrestrial ecosystems) cannot be identified through traditional gut contents analysis, but if characterised through other means (e.g. laboratory trials or molecular sequencing) they can be included in the dataset and incorporated into generated food webs. For instance, two suctorial predators in the Broadstone food web (*Platambus maculatus* and *Bezzia sp.*) did not have their guts analysed for predatory links in the original study (Woodward, Speirs & Hildrew 2005) and so had been previously excluded from the food web (e.g. Petchey *et al.* 2008; Woodward *et al.* 2010b), these nodes would have been predicted to prey upon other species by the `WebBuilder` function. This is due to the `WebBuilder` function generalising the taxonomy of these nodes, and their subsequent appearances in the dataset, as other studies have characterised the diets of these taxa. Some links in the dataset of trophic interactions were known from just a single data source, e.g. *Cordulegaster boltonii* as a consumer of *Nemurella pictetii* is known only from the Broadstone food web. Therefore, when we excluded self-referrential diet data, the `WebBuilder` function reconstruction of Broadstone did not

predict a trophic link between *C. boltonii* and *N. pictetii*. We have not quantified how often this effect occurred. As with other open-source datasets, anomalies will be ironed out as the dataset is enriched with more observations as it grows, and its coverage will improve over time.

Besides constructing food webs *de novo*, the `WebBuilder` function could be used to standardise a collection of networks gathered from different sources prior to analysis. This would effectively standardise the sampling effort for included interactions (although not for species richness or taxonomic resolution) and would remove spatially or temporally explicit interactions (or lack thereof). If the analysis was concerned with the structure of summary food webs from different locations and habitat types then this might be an appropriate first step.

### 2.5.2 Future Directions

The realism of links generated by the `WebBuilder` function could be addressed by assessing the number of times a particular interaction appears in the dataset, as well as the number of times an interaction could have occurred but did not (i.e. species found at the same site but not found to interact). If a particular interaction has been observed many times across many systems, it is probably reasonable to assume it also occurs at other sites where those species co-exist. However, if it has only been observed rarely, or at a site with very different characteristics than the one in question (for instance contrasting environmental conditions, or significantly different community assemblages) this assumption might not be so reasonable. As the size of the dataset continues to grow, evaluation of whether links are realised or not will improve over time.

The `WebBuilder` function is designed to construct summary food webs, and ignores potential behavioural shifts of species, hence it is unsuitable for constructing temporally or spatially explicit food webs. Additional data such as abundance

information could be used to weight interactions, this would, for instance, reduce the weight of interactions between rare species reducing their influence on food web structure and increasing the realism of the resulting food web. There is an increasing body of literature detailing the importance of weak and strong interactions within networks (de Ruiter, Neutel & Moore 1995; Berlow *et al.* 2004; Vázquez *et al.* 2007) and a multitude of methods already exist for determining interaction strengths in food webs (see Berlow *et al.* 2004) some of which can be employed alongside the `WebBuilder` function. Thus, despite the 'coarse' nature of food webs built in this way there is much potential for their use in ecological research, and by combining them with models such as those presented here potential mismatches arising from behavioural shifts could be highlighted.

It would be straightforward for the underlying code of the `WebBuilder` function to be extended to incorporate a range of traits that could influence the realisation of potential trophic interactions, other than phylogeny, such as life stage or body size. For instance, within freshwater food webs body size is an important determinant of trophic interactions, and food web structure predicted using body size alone may be more accurate than those predicted using phylogeny alone (Woodward *et al.* 2010b). This could further increase the realism of the constructed food webs and hence their wider applicability and usefulness.

This dataset of trophic interactions was collated to test the performance of the `WebBuilder` function when predicting the structure of the four empirical UK freshwater food webs used here. It would be straightforward to extend the coverage of this dataset by augmenting it with data collected from other geographic regions. If this dataset is used to construct food webs in the future researchers will need to use their discretion to decide how applicable it is to their system. For instance, this initial version of the dataset does not provide good coverage of lentic species, or species from across Europe or other parts of the world. However, interaction data are being published at a rapidly accelerating rate (Ings *et al.* 2009) and this can be used to form

an iterative feedback process, improving data quality over time; the presence of links predicted by the `WebBuilder` function can therefore be tested evermore rigorously in the future. Identifying underrepresented nodes in the dataset will help target further research more cost effectively: e.g., a great deal is known about the diet of a handful of often economically valuable species in the dataset (for instance, brown trout, *Salmo trutta* appears >3,000 times), but very little is known about many others. Additionally, technologies such as those provided by recent advances in molecular sequencing will improve the efficiency of trophic interaction detection (Clare 2014) and therefore the volume of data which can be incorporated into the dataset. We actively encourage researchers with suitable data to contribute them to this dataset. Exciting initiatives such as Global Biotic Interactions (Poelen, Simons & Mungall 2014), by incorporating necessary information such as the method through which an interaction was determined, could provide a global, open source repository of interaction data which the `WebBuilder` function could access through R. As more of these unknown links become known, nodes will not need to be generalised taxonomically in order to find matches in the dataset, the links generated will more closely match the known links for those species and therefore the quality of the ecological food webs generated by the `WebBuilder` function will improve.

### 2.5.3 Conclusions

We have demonstrated that the food webs generated here are comparable to empirically observed food webs and exceeded the accuracy of other potential methods of predicting freshwater food webs. This method could be used to build vast numbers of ecological networks from data that already exists, such as routine biomonitoring data which is collected in huge volumes in many parts of the world (e.g., Dutch soil biomonitoring data have recently been used to build a large collection of food webs; Cohen & Mulder 2014). Producing collections of replicable networks is vital for advancing ecological network research beyond the largely unreplicated case-study

approach that has dominated to date: the `WebBuilder` function approach presents a new robust and repeatable method that helps move us considerably closer to that goal.

# 3 | The recovery of freshwater food webs from the effects of acidification




## 3.1 Summary

Recent work has shown that consideration of the structure of ecological networks, such as food webs, can be vital for a full understanding of how ecological communities respond to environmental change. Our understanding of how the structure of food webs responds to acidification is hindered by small sample sizes and a lack of replication.

We use a uniquely large and replicated collection of 451 freshwater food webs, constructed with data from the UK Upland Waters Monitoring Network, to investigate the changes in network structure that accompany recovery from acidification. We assess if these food webs are suitable for addressing these research questions by quantifying the extent to which they are undersampled through species and link accumulation curves. From each food web we measured a range of network metrics and used these to assess how the structure of the food webs has changed over time at each site. There was no congruence between those sites exhibiting clear chemical recovery trends and evidence of change in their network structure. However when the food webs were modelled at the regional (UK) scale, food web generality, vulnerability and network efficiency decreased with increasing acidity, while node redundancy increased with acidity. Many acidity related variables, such as $SO_4$, pH, dissolved organic carbon, labile aluminium, acid neutralising capacity Ca, $NO_3$ and Cl were identified as drivers of community structure, while only $NO_3$ was found to drive changes in network structure.

These findings, which support previous work done using a far smaller collection of food webs, indicate that community and food web structure are fundamentally altered by acidity. There may be an inherent stability to acidified food webs, which may be limiting biological recovery, however further investigation is required.

## 3.2 Introduction

As we move further into the 6th mass extinction event, a deeper understanding how complex systems respond to environmental change and recover from perturbations is of crucial importance (Pimm *et al.* 1995). Previously work has shown that biological recovery from perturbation does not necessarily follow a reversal in the trajectory of decline (Scheffer & Carpenter 2003; Feld *et al.* 2011; Murphy *et al.* 2014). Species interactions confound attempts to scale up predictions made from individuals or populations to the whole-community or ecosystem level (Ings *et al.* 2009; Thompson *et al.* 2012). The structure of the network created by the interactions between species determines the stability of that community and thus modulates its resistance and robustness to environmental change. Indeed, often it is the interactions between species that are suggested as mechanisms which delay or alter the trajectory of recovery (Scheffer & Carpenter 2003). Therefore it is necessary to consider these interactions when attempting to assess the consequences of environmental change on communities.

Acidification of freshwaters is caused by atmospheric pollution, such as sulphur dioxide, which is deposited in the environment and subsequently washed into freshwater systems, or taken up by moisture in the atmosphere to become 'acid rain' (Driscoll *et al.* 2001). Acidification has profound ecological impacts, including the loss of many acid-sensitive species from all trophic levels (e.g. Round 1990; Rosemond *et al.* 1992; Sayer, Reader & Dalziel 1993). Increased surface water concentrations of inorganic aluminium, which becomes more soluble in acidified soils, is toxic to many species, in particular salmonid fishes (Sayer, Reader & Dalziel 1993) and a range of macroinvertebrate taxa.. Controls on acidic emissions in Europe came into force in 1983 through the United Nations Economic Commission for Europe (UNECE) Convention on Long Range Transboundary Air Pollution (LRTAP) with the specific aim of reducing the impact of acid deposition on soils, vegetation and surface waters. Since this point there has been a dramatic reduction in the emissions of $SO_2$ and $NO_x$

gasses across Europe.

The Upland Waters Monitoring Network (UWMN, formerly the Acid Waters Monitoring Network) was set up in 1988 to assess the chemical and biological recovery of surface waters in the UK. The network comprises 23 stream and lake sites distributed across acid sensitive, base poor geology regions of the UK (Figure 11). The sites were chosen for their vulnerability to acidification. They are generally distributed across the upland areas of the west coast of the UK where precipitation, and fluxes of sulphur and nitrogen have tended to be high, and are predominantly located in regions with base-poor geology and, hence, are particular susceptible to acidification. The sites also include lakes and streams in areas overlying acid sensitive geologies in regions receiving relatively little acid deposition, such as north- western Scotland (Patrick *et al.* 1991). The design of the Network, sampling methodology and analytical protocols are provided by Patrick *et al.* (1995).

The acidity of most UWMN sites has declined significantly since the onset of monitoring (Monteith *et al.* 2014). Evidence for biological responses to chemical improvement varies between sites (Murphy *et al.* 2014) with only half of the sites showing significant trends. Similar, "sluggish", biological recovery has also been reported elsewhere (e.g. Arseneau *et al.* 2011). Several hypotheses have been put forward to explain the lag in biological recovery (Yan *et al.* 2003; Monteith *et al.* 2005), including: dispersal limitations, ocasional acid episodes and food web dynamics which might resist the re-establishment of more acid-sensitive species. The long distance dispersal abilities of freshwater macroinvertebrates is now known to be sufficient to recolonise UWMN sites and so cannot be the mechanism preventing biological recovery (Masters *et al.* 2007; Hildrew 2009). Likewise, while many of the stream sites experience episodic drops in pH which might hinder recovery (Evans, Monteith & Harriman 2001), lakes are far less prone to dramatic fluctuations in hydrochemistry, and both streams and lakes show limited biological recovery. Additionally, across the network, the pH during more acidic episodes has fallen more

rapidly than has the average tendency (Monteith *et al.* 2014), and thus occasional acidic events is not considered to be limiting biological recovery (Monteith *et al.* 2005). It has been proposed that the dynamics of the food web itself may be limiting biological recovery. Generalist herbivore/detritivores macroinvertebrates are known to inhibit the return of acid-sensitive specialist algal grazers (Ledger & Hildrew 2005; Layer, Hildrew & Woodward 2013), and dynamic modelling has revealed that acidified food webs are more robust over time, suggesting that they might resist re-invasion (Layer *et al.* 2010b). Redundancy within networks is a property which may provide resistance to perturbations, and hence may be a property of more stable systems (Naeem 1998; Solé *et al.* 2003; Peralta *et al.* 2014). Likewise, some food webs have been found to display 'small-world' properties, namely have shorter path lengths between nodes than expected (Watts & Strogatz 1998; Montoya & Solé 2002), which influences the rate at which perturbations propagate (Montoya, Pimm & Solé 2006). Network efficiency, which is a measure of how well connected a network is, was measured to make inferences about a networks 'small-world' properties (Latora & Marchiori 2001).

Food webs are a representation of the structure and functioning of communities which, in turn, can regulate their sensitivity to environmental change (Ings *et al.* 2009; Thompson *et al.* 2012). In particular, food web complexity and the distribution of interaction strengths are key determinants of stability, influencing how a community responds to environmental stress (May 1972; McCann 2000). For example, food web size, linkage density and trophic height in acid sensitive waters all decrease with exposure to lower pH (Layer *et al.* 2010b, 2011). Acidified food webs are also smaller, simpler and have lower average interaction strengths than more circumneutral freshwater food webs (Woodward & Hildrew 2002; Layer *et al.* 2010b). To date, assessment of temporal dynamics of food webs in response to the amelioration of acidification has been restricted by the resources required to map them.

Here we use community data collected by the UWMN, coupled with an

understanding of key trophic interactions gathered from the literature, to construct an unprecedentedly large collection of food webs. We use this collection to investigate how these complex networks re-assemble as the community recovers from acidification. Specifically, we address the following questions:

1) Has there been directional change in network structure over the past 25 years at UWMN sites? Do such changes indicate recovery from acidification?
2) What are the major environmental determinants of community composition and network structure?
3) Do acidified food webs have greater redundancy and efficiency, as would be expected for more stable networks?

## 3.3 Methods

### 3.3.1 Sites

The UWMN consists of 11 stream and 12 lakes distributed across the UK (Figure 11): full site descriptions and sampling methodologies are provided in Patrick, Monteith & Jenkins (1995) and Kernan *et al.* (2010). Water chemistry, epilithic diatom, macroinvertebrate and fish sampling began in spring 1988 and continued uninterrupted at most sites up to 2012, except for access restrictions to a few sites during a foot-and-mouth disease outbreak in 2001, and some isolated adverse weather conditions in other years (see Kernan *et al.* 2010). The sites are distributed along a latitudinal gradient across the UK, which can be interpreted as a proxy for the degree of acid deposition that each site was exposed to at the onset of monitoring as those sites at high latitudes were exposed to relatively little acid deposition whilst those sites at lower latitudes tended to be more heavily acidified (Patrick *et al.* 1991). One lake site, Loch Coire nan Arr, was affected by damming that increased water levels and was replaced in 2001 by Loch Coire Fionnaraich which has comparable characteristics.

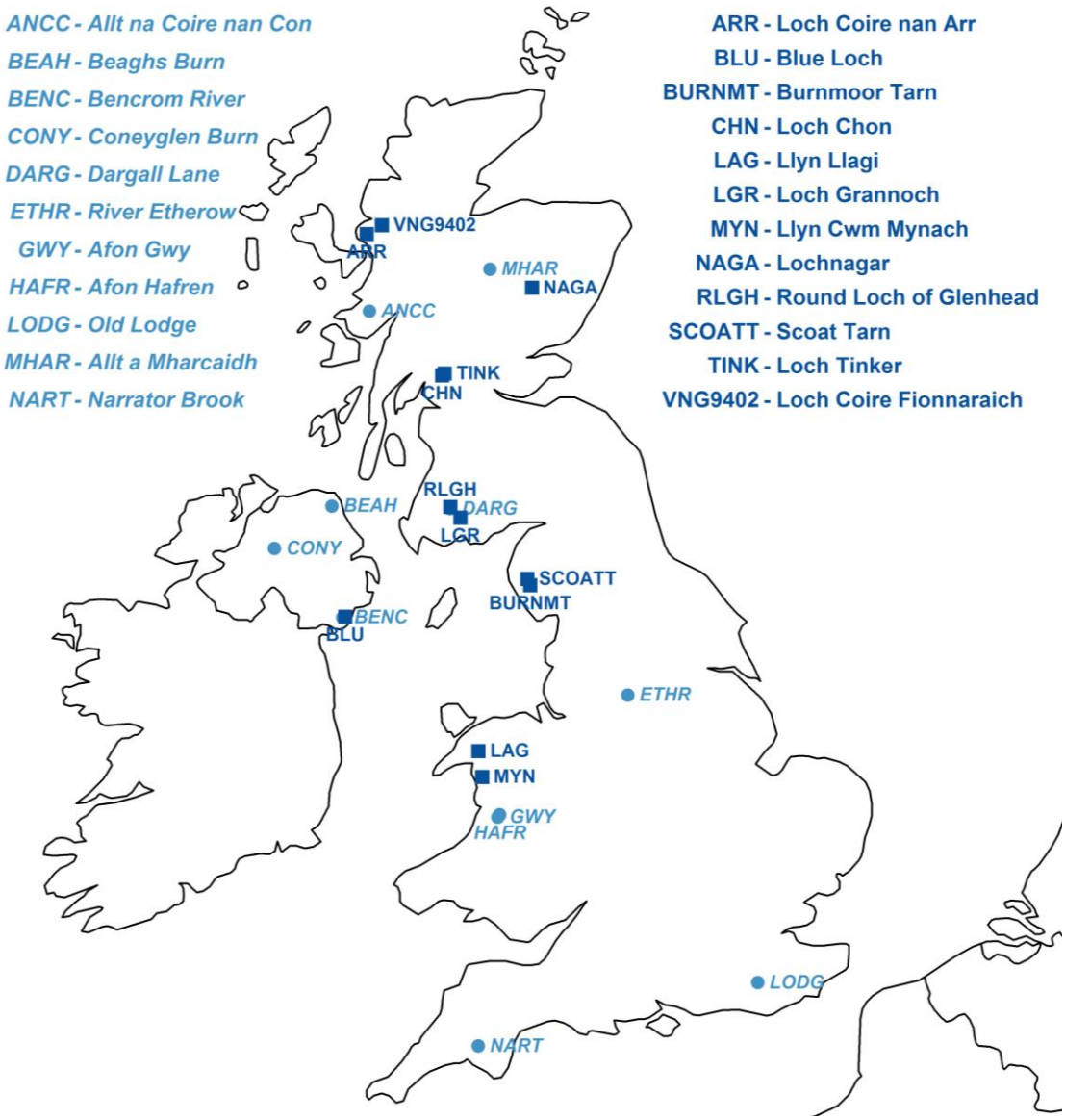



**Figure 11**. The Upland Waters Monitoring sites, consisting of 11 lakes (dark blue squares) and 12 streams (light blue circles).

### 3.3.2 Hydrochemistry

Hydrochemistry samples were taken monthly from stream sites, and quarterly from lake sites. All dip samples were collected in acid-rinsed bottles. A large number of chemical variables were recorded at each site, for more details see Kernan *et al.* (2010) and Monteith *et al.* (2014). In total 14 variables considered to be key drivers or indicators of acidification (Monteith *et al.* 2014) were used here; pH, Gran Alkalinity,

$H^+$, Conductivity, nitrate ($NO_3$), non-labile aluminium, soluble aluminium, labile aluminium, Dissolved Organic Carbon (DOC), sodium (Na+), sulphate ($SO_4^{2-}$), calcium ($Ca^{2+}$) and $Cl^-$. Acid Neutralising Capacity (ANC) describes the ability of water to resist acidification by a strong acid, and is calculated here from DOC and labile aluminium concentrations, as well as alkalinity, see Kernan *et al.* (2010) for more details.

Except for pH, we used the annual arithmetic mean of all hydrochemical data as summary statistics for each site. Annual average pH was calculated by first converting pH to $H^+$ concentration, calculating the annual arithmetic mean, and then converted back to pH. In addition, minimum recorded pH and ANC were used to produce annual hydrochemical summary statistics for each sites.

### 3.3.3 Biota

Benthic diatom, macroinvertebrate and fish populations were sampled annually from 1988-2012. Benthic diatoms were sampled by selecting five cobble sized stones at a depth below that of minimum flow in streams, or the permanently submerged littoral zone in lakes. The stones were taken from discrete locations - upstream, middle and downstream of a surveyed 50 m reach in streams, or three or four surveyed locations around the shore of lakes, with areas close to inflow or outflow streams being avoided. Epilithic diatoms were removed by brushing into a clean funnel and plastic vial then preserved in Lugol's Iodine immediately. Samples were prepared using standard techniques (Battarbee *et al.* 2001) and examined by light microscopy at x1000 magnification. Three hundred valves were counted from each sample and identified to species level.

Macroinvertebrates were sampled by taking five separate one minute kick samples using a standard hand net (300µm mesh) from riffle sections of streams and the dominant littoral habitat of lakes. Using a white tray, halogen lamp and fine forceps, all invertebrates were picked out and preserved with 70% Industrial Methylated Spirit. With the exception of Diptera, Oligochaeta and Bivalva, taxa were

identified to species level. Diptera were further identified to family level and Bivalvia to genus level. All taxa were counted.

Annual electric fishing surveys were employed to assess the abundance of salmonid populations at each stream site and at the outflow streams immediately downstream from each lake site. The presence of any non-salmonid fish species was recorded although no abundance data were collected. Fishing occured between mid-September and mid-October each year. The sampling procedure used three 50m reaches distributed across 500m of the stream or lake outflow, Each 50m reach was fished using stop nets and electric fishing apparatus. The fishing was repeated in each reach three times, or more if no clear drop off in numbers occurred.

### 3.3.4 Food web construction

Binary food webs, in which species (nodes) and links are described in terms of their presence/absence in each year at each site were constructed for all sites in all years for which there was complete biological and hydrochemical data, this resulted in the production of 451 food webs in total. Feeding links between species were inferred from published literature, and filled in for each network using the `WebBuilder` function (Chapter 2; Gray *et al.* 2015) and associated dataset of trophic interactions, in R (R Core Team 2013). This method is based on the assumption that all feeding links between specific pairs of species that have been reported previously would be realized wherever and whenever both species co-exist at a study site (Hall & Raffaelli 1991; Martinez 1991; Layer *et al.* 2010b; Pocock, Evans & Memmott 2012). In some instances, due to a paucity of trophic interaction data, feeding links were assigned on the basis of taxonomic similarity.

### 3.3.5 Network metrics

A range of food web metrics were calculated from each food web. The number of nodes in each network was measured as the total number of connected species. Mean trophic height of each food web was calculated using the method of Levine (1980) and defined as 1 plus the mean trophic level of a consumer's resources,

averaged across all consumers. The maximum trophic height of each food web was defined in the same way, except that the maximum value across all consumers was taken. Mean generality ($G$; number of resources per consumer) and mean vulnerability ($V$; number of consumers per resource) of each network was calculated. Additionally for each taxon $k$, normalised $G$ and $V$ were calculated:

$$G_k = \frac{1}{L/S}\sum_{i=1}^{S} a_{ik} \quad (1)$$

$$V_k = \frac{1}{L/S}\sum_{j=1}^{S} a_{jk} \quad (2)$$

Where $S$ is the number of nodes and $L$ the number of links in a food web. $a_{ik} = 1$ if taxon k consumes taxon i (otherwise $a_{ik} = 0$), and $a_{jk} = 1$ if taxon k is being consumed by taxon j (otherwise $a_{jk} = 0$). Mean $G_k$ and $V_k$ in any given food web equal 1, making their standard deviations, which give an indication of the variability in $G$ and $V$ respectively across a network, comparable across networks of different size. These metrics were all calculated using the R package cheddar (Hudson *et al.* 2013).

The global efficiency (Latora & Marchiori 2001) of a network describes the 'reachability' of each node by any other node, and is a measure of the overall connectivity of the network. The global efficiency of each network was calculated as follows:

$$E = \frac{1}{S(S-1)}\sum_{i \neq j \in G} \frac{1}{d_{ij}} \quad (3)$$

Where $d_{ij}$ is the shortest path length between node $i$ and $j$, using the sna R package (Butts 2013).

The proportional redundancy of each network was calculated by grouping nodes into trophic species (i.e. nodes with common resources and consumers) and then calculated as follows:

$$Redundancy = 1 - \frac{T}{S} \quad (4)$$

Where $T$ is the number of trophic species within the network. Redundancy was calculated using functions from the cheddar package (Hudson *et al.* 2013) in R.

### 3.3.6 Statistical data analysis

All statistical data analysis was done in R version 3.1.1 (R Core Team 2013). Mann–Kendall trend tests determined whether there were significant trends in mean pH, ANC, DOC, labile aluminium, and network metrics over time at each site. We used $\chi^2$ contingency tests to assess the extent to which sites that exhibited clear chemical recovery trends also showed evidence of change in their network structure (Murphy *et al.* 2014). For each chemical variable, and for each network metric, we counted the number of sites (out of 23) that exhibited (a) a trend in both, (b) a biological but not a chemical trend, (c) a chemical but not a biological trend, and (d) with neither trend. The $\chi^2$ test assessed whether the distribution of sites across these four categories was different to that due to random chance.

Principal Component Analysis (PCA) was performed on the water chemistry data of each site. Yearly mean (or minimum) values for key hydrochemical variables were centred to zero and scaled by their standard deviations, and sample scores on the first PC axis (PC1) extracted for use as a proxy for water chemical stress. Each network metric was regressed against PC1, and any trend assessed with Generalised Linear Mixed Effects models. Site type (lake or stream) was fitted as a fixed effect, and any potential interactions with PC1 were assessed on the basis of stepwise model

simplification and model AIC. For each model, site and year were used as random effects, but a range of random effects structures were investigated for each response variable, the best model was selected on the basis of AIC.

To assess the principal hydrochemical determinants of community structure, a distance-based Redundancy Analysis (RDA) model was used in step-wise model selection of hydrochemical variables based on their P-values and AIC scores. The community matrix was constructed from the diatom, invertebrate and fish data, which were counts of each species at each site in each year. As three different sampling methodologies had been used to characterise the biota, the Wisconsin double standardization was used; the abundance values were first standardized by each species maximum score, and then by sample total, and by convention multiplied by 100 (Bray & Curtis 1957). Bray-Curtis dissimilarity scores were used. The hydrochemical variables used here were further selected to minimise co-linearity, in total 8 variables (yearly mean values) were used; pH, $SO_4$, Dissolved Organic Carbon (DOC), labile aluminium, Acid Neutralising Capacity (ANC), Ca, $NO_3$ and Cl. In the step-wise model selection procedure, first each hydrochemical variable was used as the sole constrained explanatory variable within the RDA model, and the explained variation by each model was recorded. Secondly the variables were ranked by the explained variation in each of the constrained ordinations. Variables were then sequentially added to the model, at each stage the significance of each variable was assessed using Monte Carlo permutations and the variation explained by the remaining variables was recalculated and the variables re-sorted by this value. Variables were sequentially added to the model in this manner until the next best variable no longer significantly improved the model. Comparison of variables is based on AIC criteria and p-values from Monte Carlo permutation test (n=199). Finally the explanatory power of each of the variables in the final model was assessed using Permutational multivariate analysis of variance (PERMANOVA) with 9999 permutations. To determine the principle environmental determinants of network

structure, the step-wise model selection procedure was repeated using a dataframe of network metrics in the place of the community matrix and Euclidean distances. All ordination analysis was performed using the vegan package in R (Oksanen *et al.* 2015). Variance partitioning was used to assess the variation in community or food web structure explained by the first four (or fewer) significant environmental variables.

## 3.4 Results

### 3.4.1 Has there been directional change in network structure?

Trends in hydrochemical variables varied across all sites. Several showed significant increasing trends in average annual pH (13 out of 23; Figure C Appendix E), Acid Neutralising Capacity (18 out of 23; Figure D Appendix E), Dissolved Organic Carbon (20 out of 23; Figure E Appendix E) and significant decreasing trends in labile aluminium (14 out of 23; Figure F Appendix E) suggesting that at least partial recovery from acidification has occurred at most sites (Monteith *et al.* 2014). Some sites showed significant increasing trends in mean trophic height (8 out of 23; Figure G Appendix E), vulnerability (8 out of 23; Figure H Appendix E), and its standard deviation (7 out of 23; Figure I Appendix E), significant decreasing trends in redundancy (11 out of 23; Figure J Appendix E), standard deviation in generality (10 out of 23; Figure K Appendix E) and efficiency (6 out of 23; Figure L Appendix E). Generality increased in two sites, and decreased in four others (Figure M Appendix E), likewise maximum trophic height increased in one site, and decreased in two others (Figure N Appendix E). These mixed trends were unrelated to the sites severity of acidification at the beginning of monitoring. Indeed, $\chi^2$ tests revealed that there was no congruence between those sites exhibiting chemical and biological recovery (Table A Appendix E).

### 3.4.2 What are the environmental determinants of community and food web structure?

When food web data were analysed together at the regional (UK) scale, network metrics were related to the degree of environmental stress that the food web was exposed to (Figure 12). Generality, vulnerability and efficiency decreased with increasing environmental stress (low pH) whilst redundancy increased with increasing environmental stress (low pH).

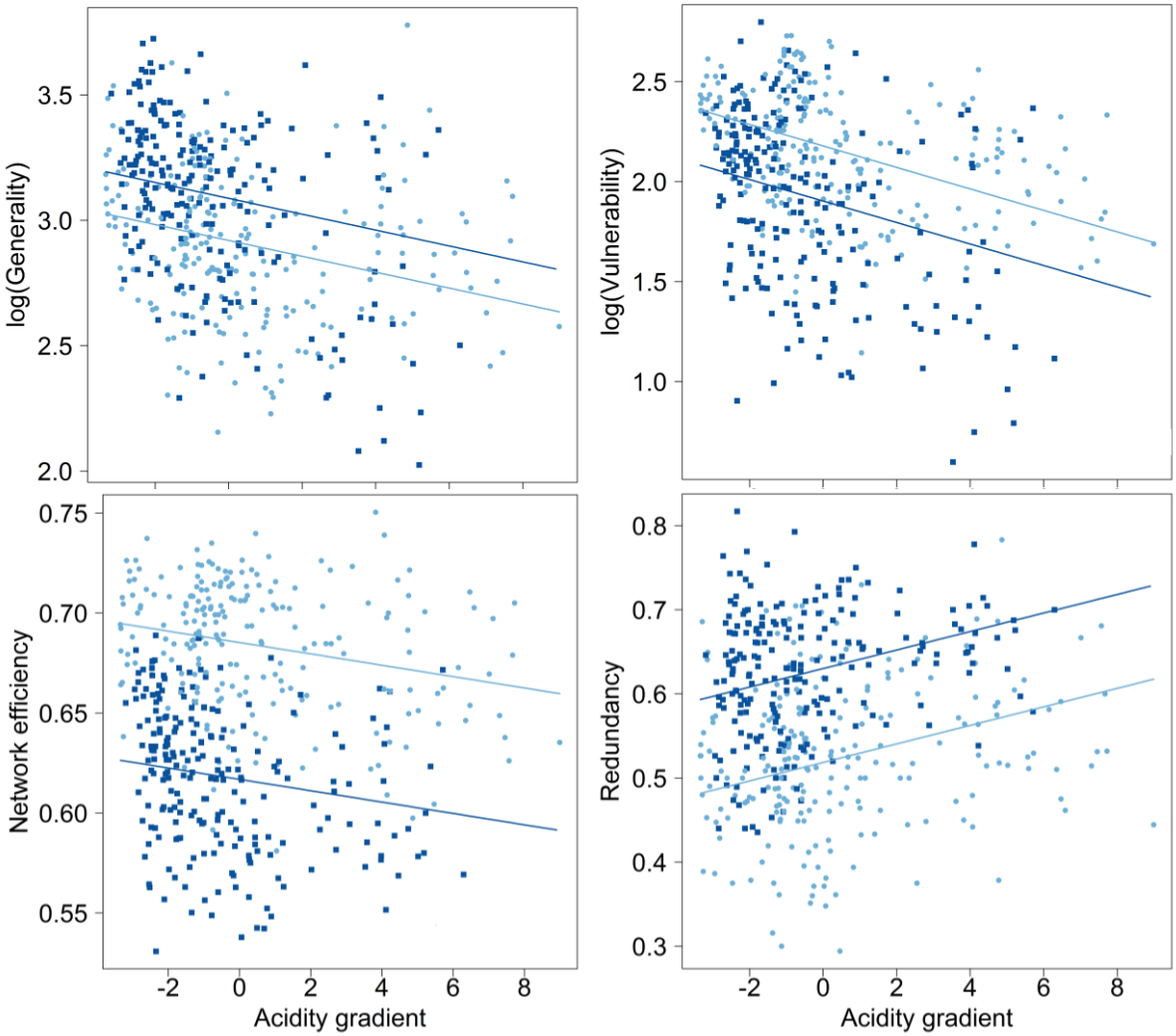



**Figure 12**. Network metrics vary with environmental stress. The acidity gradient is PC1 extracted from the PCA of the water chemistry data, and is strongly related to pH, ANC & labile aluminium, such the x-axis can be interpreted as increasing environmental stress from left to right. Lines indicate fitted values from GLMM where $p < 0.05$.

The RDA model testing of the effects of environmental variables on community structure was significant (p=0.001; Figure 13), the constrained component explained 16% of the variation in community structure while the conditional component (site and year) explained 50% of the variation. All 8 environmental variables were included in the model structure through step-wise model selection, after PERMANOVA they all had a significant effect on community structure (Table 4). Variance partitioning determined that $SO_4$ explained 2.02% of the variance in community structure, pH 3.7%, DOC 1.8% and labile aluminium 0.93%.

The RDA model testing of the effects of environmental variables on food web structure was significant (p=0.008; Figure 14), the constrained component ($NO_3$) explained 0.3% of the variation in community structure while the conditional component (site and year) explained 61% of the variation. Among the 10 environmental variables only $NO_3$ had an effect on food web structure (Table 5).

**Table 4**. Effects of hydrochemical variables on community composition determined through PERMANOVA with 9999 permutations, with site and year fitted as conditional variables. All variables are yearly averages. Bold p-values indicate significance at α = 0.05.

| Variable | d.f. | SS | Pseudo-F | p-value |
|---|---|---|---|---|
| SO4 | 1 | 0.36 | 2.0535 | 0.001 |
| pH | 1 | 0.251 | 1.4272 | 0.002 |
| DOC | 1 | 0.334 | 1.9003 | 0.001 |
| L_Al | 1 | 0.273 | 1.5568 | 0.001 |
| ANC | 1 | 0.278 | 1.5853 | 0.002 |
| Ca | 1 | 0.244 | 1.3883 | 0.004 |
| NO3 | 1 | 0.412 | 2.3487 | 0.001 |
| Cl | 1 | 0.274 | 1.561 | 0.003 |
| Residual | 412 | 72.315 | | |

**Table 5**. Effects of hydrochemical variables on food web metrics determined through PERMANOVA with 9999 permutations, with site and year fitted as conditional variables. All variables are yearly averages. Bold p-values indicate significance at α = 0.05.

| Variable | d.f. | SS | Pseudo-F | p-value |
|---|---|---|---|---|
| $NO_3$ | 1 | 0.01186 | 2.6749 | **0.017** |
| Residual | 419 | 1.85714 | | |

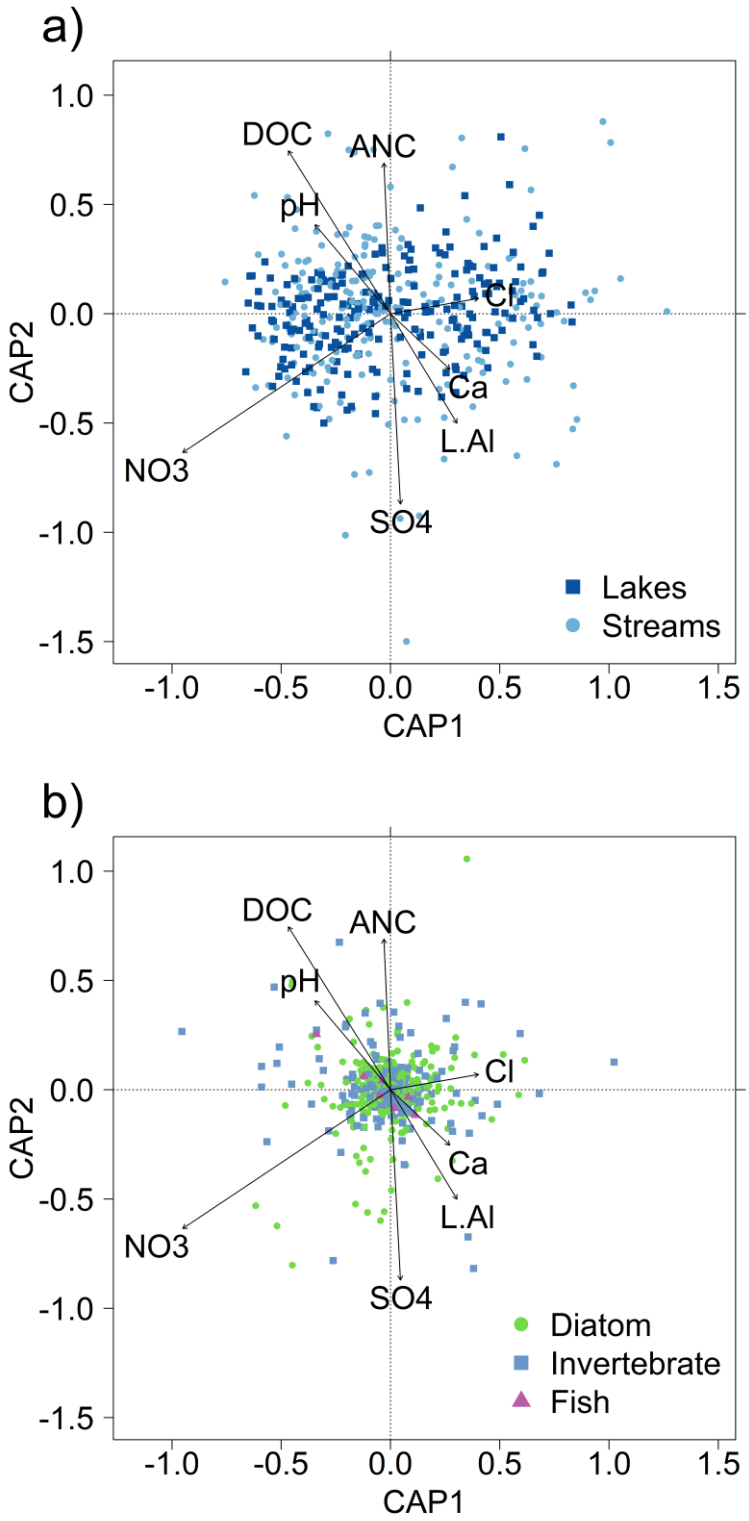



**Figure 13**. Distance-based Redundancy Analysis with $SO_4$, pH, DOC, labile aluminium (L_Al), ANC, Ca, $NO_3$ and Cl fitted as constrained variables and site and year as conditional variables. Site scores (a) and species scores (b) are shown.

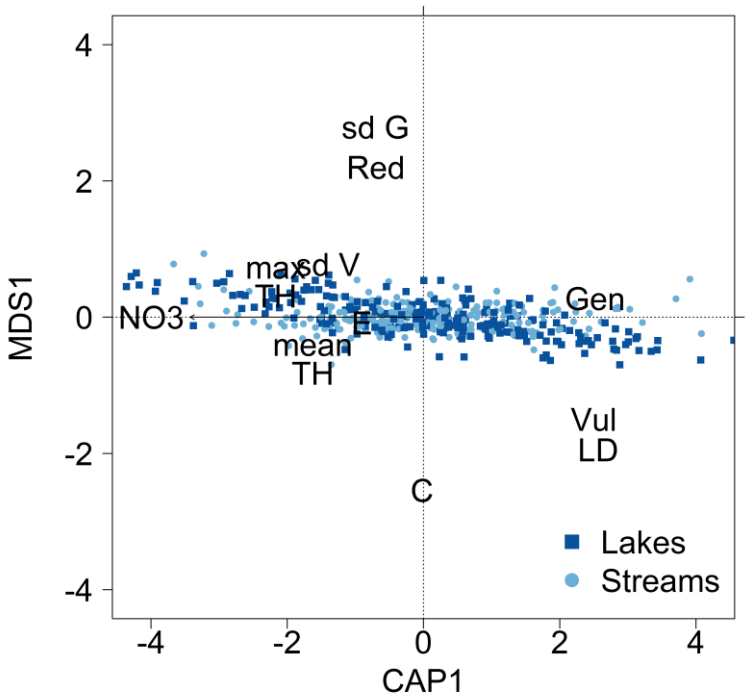



**Figure 14**. Site scores from Distance-based Redundancy Analysis with $NO_3$ fitted as a constrained variable and site and year as conditional variables. maxTH and meanTH are mean and max trophic height respectively. sd.G and sd.V are the standard deviation in normalised generality and vulnerability scores respectively. E is network efficiency and S is the number of nodes.

## 3.5 Discussion

It is clear that scale is an important consideration when assessing food web data. When the current data were analysed at the site scale, site scale sources of variation such as individual site characteristics or sampling error swamped any trends in food web structure over time. For instance weather conditions immediately before or during the time of sampling were largely uncontrolled for. Additionally the acidity gradient that each site is exposed to is small relative to the gradient of the whole dataset. When the data were aggregated and analysed at the regional (UK) scale then significant trends emerged.

There were instances of contradictory trends within the network for instance Afon Gwy and Old Lodge were recovering in terms of their hydrochemistry, but were yet to show a trend in their invertebrate community composition. Conversely, Coneyglen Burn's invertebrate community had experienced significant turnover, but its hydrochemistry did not shown a significant time series trend (Kernan *et al.* 2010). Extreme events, and the small sampling window for each site has caused some sites to lose and regain their significant trends in biota recovery over time (Monteith & Evans 1998, 2005; Kernan *et al.* 2010). Results from the analysis of the aggregated data are easier to interpret although attributing these trends to specific hydrochemical drivers becomes more difficult.

### 3.5.1 Chemical, community and food web recovery across the network

Our results provide clear evidence, in concordance with other published studies, for reductions in the acidity of acidified waters across the UWMN sites (Monteith *et al.* 2014), which is consistent with other international assessments of trends in acidified waters (Stoddard *et al.* 1999; Evans *et al.* 2001; Skjelkvåle *et al.* 2005).

Average annual pH and ANC increased in almost all of the historically acidified sites, indeed, Kernan *et al.* (2010) who used the full dataset (rather than annual averages) found that seventeen of the 22 sites they studied (Loch Coire

Fionnaraich was not included) showed significant increases in pH, and eighteen showed significant increasing trends in ANC. However $SO_4$ concentrations across the network in 2012 remained several times higher at most sites than those found at the least impacted sites (Monteith *et al.* 2014), so clearly there is further to go in terms of chemical recovery at these sites. Labile aluminium has also fallen dramatically across the UWMN sites, although in 2012 the concentration was still far higher than is typical for sites unimpacted by acidification (Monteith *et al.* 2014).

The increasing trends in DOC in UWMN sites is mirrored in other industrialised regions of the northern hemisphere (De Wit *et al.* 2007; SanClements *et al.* 2012) and have been ascribed to an increase in the solubility of soil organic carbon as a consequence of reductions in acid deposition (Monteith *et al.* 2007; Clark *et al.* 2011; Evans *et al.* 2012). This means that although there has been a marked reduction in strong mineral acids (e.g. $SO_4$) this has been partially buffered by an increase in weak organic acids (DOC), and thus the expectation for recovery of pH has been modified (Evans *et al.* 2008). However there is increasing evidence that DOC provides an important ecosystem service as it helps to protect waters from acidification (Monteith *et al.* 2014), and increasing DOC may well be part of a natural chemical recovery trajectory.

The general increasing mean trophic height of food webs over time at each of the sites is to be expected from what we know about how these systems respond to de-acidification; under acidification species are lost throughout the food web but top predators such as fish (Henriksen, Fjeld & Hesthagen 1999) and many predatory macro-invertebrates are especially vulnerable (Layer *et al.* 2011). As these acid-sensitive species re-colonise feeding chains will lengthen (Woodward & Hildrew 2001) and the average trophic height of the food web as a whole will increase. All of the sites which experienced this lengthening of food chains were also increasing in their pH, although not all sites increasing their pH also increased their mean trophic height. This trend was not detected in the aggregated data across the stress gradient,

suggesting that other environmental factors, other than pH (for instance DOC which was closely related to PC2), may be impacting mean trophic height and complicating the pH/trophic height story.

Along the hydrochemical stress gradient, vulnerability and generality both decreased (Figure 12), although no trend in generality overtime in the individual sites was detected, and vulnerability tended to increase over time at individual sites. However there were only eight increasing vulnerability trends out of a possible 23, and many sites showed complex patterns over time (Figure H Appendix E). Those sites increasing their vulnerability were not those recovering from acidification (Table A Appendix E) hence it seems reasonable to conclude that the analysis over the stress gradient is more conclusive than the site by site analysis. The decreasing generality and vulnerability with increasing hydrochemical stress is consistent with the theory that as an acidified system recovers, specialist consumers and larger top predators re-colonise (Woodward & Hildrew 2001; Layer, Hildrew & Woodward 2013) which causes a reduction in the mean number of consumers per resource species (vulnerability), and mean number of resources per consumer (generality). There are known to be marked differences in the feeding patterns of primary-consumers across a pH gradient, the species richness of both algal resources and primary consumers increase with increasing pH, the composition of functional feeding groups within the primary consumers switch from a generalist herbivore-detritivore dominated system at low pH, to a more diverse community including specialist herbivore primary consumers (Ledger & Hildrew 2005; Layer, Hildrew & Woodward 2013). This coupled with the appearance of other acid-sensitive invertebrate species (such as the mayflies *Baetis sp.* and *Caenis sp.*, or the snail *Radix balthica*) and salmonid fish at high pH explains the decreasing trend in vulnerability and generality with decreasing environmental stress. Previous work has found that generality and vulnerability increased with pH (Layer *et al.* 2010b), but used a far smaller sample size than that used here.

Network efficiency is a measure of how 'reachable' each node is from every other node, and as such it is strongly related to how well connected a network is. Decreasing network efficiency with increasing hydrochemical stress (Figure 12) indicates that food webs under more stress are less well connected across the whole network, for instance there may be pockets of species which are poorly connected to other species, such that the average shortest path length between all pairs of nodes is increased. The addition of top predators such as salmonid fish to the system (Woodward & Hildrew 2001) may explain the increased efficiency of less stressed food webs. Consumers in freshwater systems tend to be highly generalist engulfing predators which will consume anything within a given size range. The addition of these generalist interactions between top predators and a swathe of invertebrates within the appropriate size range may well increase the reachability between those resource nodes, as well as link together different feeding pathways (i.e. allochthonous vs autochthonous).

Food webs under more hydrochemical stress contained proportionally more redundant feeding pathways than their counterparts under less stress, the proportion of 'trophic species', nodes feeding on and being fed on by the same species is larger at low pH. This is congruent with the increase is specialist consumers as acidity ameliorates. Additionally, food webs at low pH tend to have few species and few links (Layer *et al.* 2010b), making the scope for unique feeding pathways small.

The 'ecological inertia' of these food webs is an often cited mechanism to explain the delay in biological recovery (Lundberg, Ranta & Kaitala 2000; Ledger & Hildrew 2005; Kernan *et al.* 2010; Layer *et al.* 2010b), whereby food webs under acidified conditions are dynamically stable and resistant to re-colonisation by acid-sensitive species. Townsend *et al.* (1987) measured the persistence of 27 stream invertebrate communities across a pH gradient, and found that those communities from the most acidified sites were the most persistent. Likewise Layer *et al.* (2010b) used dynamic modelling to determine the robustness of stream food webs to species

extinctions, and found that food webs from more acidified conditions are more robust. Here we provide some evidence in support of this theory, redundancy is an important feature engineered into (non-ecological) stable systems, providing robustness against node loss, this translates into biological systems and leads us to predict that food webs with greater redundancy amongst its nodes might be more robust to simulated species removal. In biological systems redundancy increases the reliability of ecosystem functioning (Naeem 1998; Peralta *et al.* 2014). Here we found that food webs from acidified waters had higher redundancy amongst their nodes suggesting that they might provide more reliably ecosystem functioning rates (Naeem 1998; Peralta *et al.* 2014), and might be more robust. However, contrasting this we found that more acidified food webs had lower global efficiency. The efficiency of a network is closely related to its small world properties, highly efficient networks also exhibit small world properties. Ecological networks with small world properties can be relatively stable (Solé & Montoya 2001; Dunne, Williams & Martinez 2002a; however, see Appendix A). These contrasting results warrant further investigation to reveal if acidified food webs are more stable, or, if they are more stable in some regards and not others. For instance, acidified food webs may be more persistent (the strength of perturbation required to change a community; Pimm 1984) and at the same time less robust to perturbation (Appendix A).

### 3.5.2 Environmental drivers of community & food web structure

The chief drivers of community structure were to be expected as they were all either key drivers of ($SO_4$, Ca, $NO_3$, Cl) or respond to (pH, ANC, DOC, labile aluminium) changes in acidity. That these were not found to be the main drivers of network structure is surprising, especially since they are clearly related to changes in network structure (Figure 12). Many of the network metrics were relatively unrelated to $NO_3$ (those aligned with the first unconstrained axis, such as redundancy, standard deviation in generality, mean trophic height, efficiency) suggesting that the hydrochemical variables analysed here were not, at least in isolation, the principal determinants of network structure, rather their combined effect on the acidity of

freshwater systems had significant implications for food web structure.

### 3.5.3 Food web construction

The use of inferred feeding links in food web studies has been criticised on the basis that they might over estimate diet breadth, and fail to detect behavioural differences between sites (Hall & Raffaelli 1997; Raffaelli 2007). However we believe that the use of 'summary' food webs, which include the full complement of known possible tropic interactions can still be a useful tool for understanding community dynamics. Indeed, given the limitations of summary food webs, they are more likely to be insensitive to environmental change rather than reveal erroneous trends. Hence we believe that the trends revealed here are real, and warrant further examination perhaps by building food webs in a more empirical manner, such as through gut contents analysis (as in Woodward, Speirs & Hildrew 2005; Layer *et al.* 2010b).

Another potential limitation to the food webs produced here is that they don't include the full freshwater community, in particular the meiofauna and top predators such as the European Dipper (*Cinclus cinclus)* or Otter (*Lutra lutra)* are missing. Top predators have been shown to have a profound effect on community structure in these systems (Layer *et al.* 2011), and so their exclusion may omit an important source of variation in this data. However this was unavoidable since the presence of these species has not been systematically recoded at these sites. Additionally, the fish assemblage of the lake sites were sampled from the lake outflows, which likely contain a different fish community to that in the main lake. For instance, Pike (*Esox lucius)* are usually associated with slow moving or standing water bodies, and so would be unlikely to be sampled in the lake outflows, even if they were present in the main lake (although Pike was found to be present at lake sites on fourteen sampling occasions). Of the 434 sampling occasions on which fish were present at a site, Brown Trout was found 434 times reflecting its dominance in these systems. The next most common species was the European Eel (*Anguilla anguilla*), which was found on 136 sampling occasions. All other species (*Esox lucius, Gasterosteus aculeatus, Lampetra sp., Phoxinus*

*phoxinus and Salmo salar*) were found on less than 60 sampling occasions. The use of presence/absence data for this analysis (rather than counts of each species at each site) should help to overcome some of these limitations.

### 3.5.4 Conclusion

This is the first example, of which we are aware, of a large collection of replicated food webs distributed over both time and space. Our study is one of the first to address macroecological questions relating to the structure of food webs across time and a broad environmental gradient. Our analysis reveals fundamental structural changes occurring in the food webs as they respond to changes in acidity, these structural changes could have profound implications for the stability of the system, and may be limiting biological recovery. It would be instructive to further investigate the stability of these food webs, in order to more fully explore if there is an intrinsic food web inertia limiting the rate of recovery (Appendix A).

# 4 | Food web topological plasticity disrupts the provisioning of ecosystem services

## 4.1 Summary

There have been calls recently to direct the management of ecosystem services towards the conservation of the structure of ecological networks. However ecological networks are made up of many interactions all with the potential to influence one another. Thus, when directing management towards a particular desired interaction it is important to consider it in the context of the whole network. Carabid consumers in arable systems consume both weed seeds and gastropod resources, and can provide a measurable pest control benefit to farmers. However it is unknown to what extent the network in which these consumers are embedded rewires when alternative resources, such as gastropods, are available.

Here we use an exceptionally large dataset of 374 half-fields distributed across the UK and taken from the Farm Scale Evaluation (FSE) of genetically modified herbicide tolerant crop. We use these food webs to test if the presence of gastropods disrupts the ecosystem service of weed seed regulation. We found that increasing numbers of gastropod species are associated with a decline in the number of herbivores in each food web. There was a strong negative relationship between the herbivore and predator interaction frequency in each food web. The number of herbivores, and the herbivore interaction frequency was found to be related to the strength of weed regulation found in each half-field.

These results suggest that if management were directed toward manipulating network structure and reducing the carabid-gastropod interactions (i.e. by removing gastropods from the system) then this might result in a stronger weed regulation effect.

## 4.2 Introduction

The sustainable provision of ecosystems services has become a cornerstone of environmental research, management and policymaking (Royal Society 2009; Redford *et al.* 2012). Many of the services that humanity requires are driven by interactions between species (Montoya, Rogers & Memmott 2012), such as trophic interactions between consumers and their resources for the delivery of biological control (e.g. Macfadyen *et al.* 2011). Increasingly, we are learning that these interactions are embedded in a network of other links which reflect the structure and dynamics of the community present. Thus, any one interaction and the ecosystem service it supports can be influenced both positively and negatively by the composition of interactions making up the network in which it is embedded.

Changes in the composition of resource species within a food web, whether that be presence and absence, or changes to their relative abundances, will necessarily cause concomitant changes to their interactions with consumer species. Some interactions may disappear altogether, others will be reformed while many may change their relative strength. All these changes could potentially occur without marked changes to the consumer assemblage. Alternatively, a change in prey composition could lead to a rearrangement of all possible links and potentially complete turnover of the consumer assemblage. These two component parts of link turnover (Poisot *et al.* 2012) could go on to cause changes in the functioning and finally service provisioning of a particular network. Such turnover of links is essentially a rewiring of the existing network according to the composition and abundances of the consumer and resource community.

We know that environmental change causes species turnover (e.g. Benedick *et al.* 2006; Clough *et al.* 2007; Novotny *et al.* 2007), indeed species turnover is commonly evaluated as β-diversity between habitats and much of our attempts to manage ecosystem services is predicated upon conserving species against turnover (e.g. Benedick *et al.* 2006). More recently, an argument has been made that managing and

conserving links in networks is an important mechanism for assuring the stable delivery of ecosystem services (McCann 2007; Tylianakis *et al.* 2010; Gray *et al.* 2014). In part this is because changes in the composition of interactions can lead to changes in ecological functions despite no change in species richness (Tylianakis, Tscharntke & Lewis 2007). What is unknown, however, is the importance and contribution of link turnover to the support and delivery of ecosystem services in replicate networks of real-world ecosystems.

Here we investigate the relative importance of link and species turnover in a highly replicated network of two ecosystem services, weed seed and slug control, delivered by a common community of carabid beetle species in agricultural fields distributed across the UK. Carabid beetles are polyphagous predators and have been the subject of much research as they regulate weed seeds in agricultural systems (Bohan *et al.* 2011a) and consume gastropods pests (Bohan *et al.* 2000), both of which contribute to reduced crop yields. Thus they are potentially important contributors to the ecosystem service of pest control in agricultural systems. The diets of carabid beetles have been studied extensively (e.g. Larochelle 1990; Mundy *et al.* 2000; Saska 2008), some taxa are considered to be generalist omnivores (e.g. *Pterostichus sp*.) whilst others are specialists (e.g. *Harpalus sp*.). However the extent to which each species contributes to the ecosystem service of pest control is unknown. If seed specialists alone are enough to provide effective control of weed seeds, then the presence of gastropod resources shouldn't interfere with this service. However if omnivores are required to control weed seeds, which would also feed upon gastropods when they are present then the presence of gastropods in agricultural fields might disrupt the ecosystem service of weed seed control.

The nodes of our networks were formed from the abundances of species of weeds, slugs and carabids present in agricultural sample data, with observed trophic links gathered from the literature. Changes in species and link turnover are thus inferred by comparison between networks in different replicate fields, each with a

distinct community of weed seeds, slugs and carabids. Taking weed seed regulation as our standard ecosystem service, we ask: i) how this service is affected by the potentially disruptive and competing function of slug predation; ii) do both link and species turnover contribute towards this disruption; and, iii) should we manage species or links in real-world networks of ecosystem services?

## 4.3 Methods

### 4.3.1 Experimental design and data collection

The Farm Scale Evaluation (FSE) experiment extensively sampled the biodiversity in and around crop fields across the UK (Figure 15). Previously power analyses have shown that the nodes in the FSE dataset are fully sampled (Perry *et al.* 2003). More details of the experimental design and protocols for data collection can be found in (Champion *et al.* (2003) and Bohan *et al.* (2005), but briefly they are as follows:

The count data for the weed seedbanks, seed rain, carabids and gastropods comes from 66 spring-sown beet, 55 spring maize and 66 spring oilseed rape fields. The fields were distributed across the UK (Figure 15) and each field was sampled for one cropping year (Firbank *et al.* 2003) between 2000 and 2004. Each field was divided in two so that one half was sown with the conventional crop and the other the Genetically Modified Herbicide Tolerant (GMHT) variety. Data from both treatments were used for the analyses presented in this study, hence a total of 374 half-fields.

The pitfall-trapping of soil-surface-active invertebrates employed the method described by Brooks *et al.* (2003). Pitfall traps were distributed along transects which ran from the crop edge into the centre of each field in the spring (April /May) and summer (June/July), and in late summer (August). Viable seed available to the carabids for consumption via the return of weed seed to the seedbank (seed rain) was quantified using seed rain traps along the same transects within each field (Heard *et al.* 2003). The traps were emptied every 2 weeks throughout the growing season.

Gastropods were sampled as in Brooks *et al.* (2003), using baited refuge traps at the same positions used for the pitfall trapping in late April and in early August for spring oilseed rape, and in May and August for maize and beet. All invertebrates and non-crop seeds were identified to species, and counted. Counts were then pooled, by summation, to give a year-total estimate for each species in each half field, and from this the relative abundance of each species was calculated.

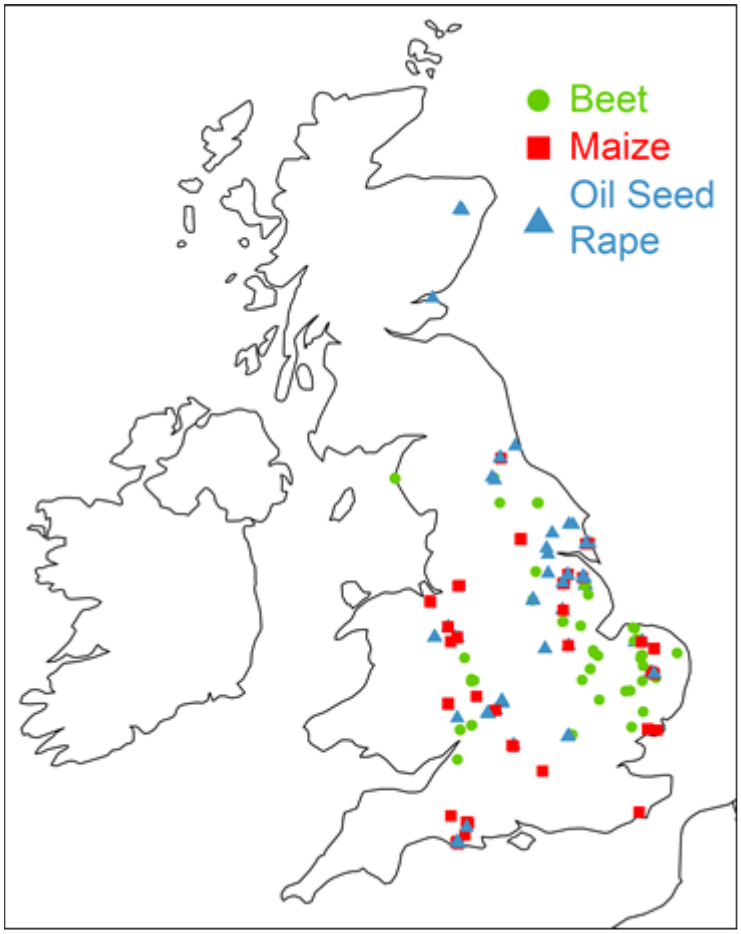



**Figure 15**. Map of the 187 sites from the FSE dataset used in this study.

### 4.3.2 Weed Regulation

To assess the regulation of weed seeds, seedbank samples were taken just prior to sowing in the experimental cropping year ($t_0$) and just prior to sowing in the

following cropping year ($t_1$). Seedbank abundance was estimated by taking soil cores along four transects running into each half field. Germinated seeds in the seedbank sample were counted and identified to species (Squire, Rodger & Wright 2000; Heard *et al.* 2003). The seedbank counts were then pooled, by summation, to give an estimate of the seedbank in each half field (total weeds) as well as dicotyledon, monocotyledon and individual species counts. Regulation was calculated from the change in seed bank counts between $t_0$ and $t_1$, using the following formula:

$$regulation = \ln\left(\frac{t_1 + 0.5}{t_0 + 0.5}\right) \quad (1)$$

so that for each half field a measure total, dicotyledon, monocotyledon and individual weed species regulation was calculated.

### 4.3.3 Food web construction

The species sample data were supplemented with carabid dietary information harvested from the literature. We assumed that where a carabid species, A, was observed to consume a resource species, B, in the literature and both these species were present in the sample data from one half field, then this interaction was realised (as in Goldwasser & Roughgarden 1993; Havens 1993; Layer *et al.* 2010a; Pocock, Evans & Memmott 2012). To standardise the (trophic interaction) sampling effort across all carabid species, and account for poorly studied carabid species which did not appear in the literature, it was assumed that each carabid would consume the same resources as other carabids within the same genus (see Goldwasser *et al.* 1993; Layer *et al.* 2010). A similar generalisation was made at the resource level. Where particular carabids was recorded to feed upon one species of gastropod or weed in the literature, we assumed that this carabid would also consume other resource species of the same genus (Gray *et al.* 2015b). This generalisation was done to reduce the numbers of isolated species within each network, and to avoid the bias towards more studied species (Ings *et al.* 2009; Woodward *et al.* 2010b).

Interaction frequency between each consumer and resource was calculated as the product of consumer relative abundance and resource relative abundance. Species abundance is known to be a major predictor of the strength of its interactions with other species (e.g. Reuman & Cohen 2005), hence weighting the links in this manner incorporates an estimate of the variation in interaction strength with species abundance, within each food web. Incorporating weighted links in this way builds upon the simple binary food web structure built from presence/absence data. Since species abundances vary wildly in response to its local environment the resulting network structure is also be more sensitive to change.

Following network construction, each carabid species was assigned to a trophic group based upon their role in each replicate network in which they are found. Carabid nodes linked only to gastropods were assigned to the 'predator' grouping, while those consuming only weeds were 'herbivores', and 'omnivores' were species linked to both gastropods and weeds. Thus, a particular carabid species might be a predator in one food web, a herbivore in another and an omnivore in yet another.

### 4.3.4 Statistical analysis

All analysis was done in R (R Core Team 2013) using the cheddar (Hudson *et al.* 2013), bipartite (Dormann, Gruber & Fruend 2008) and vegan (Oksanen *et al.* 2015) packages. Food web plots were created with the HiveR package (Krzywinski *et al.* 2012). Herbivore or predator interaction frequency for each food web was calculated as the sum of all those interactions belonging to these carabids feeding only on weeds or gastropods respectively (i.e. those feeding on both resource types were classified as omnivores and excluded).

Species and link turnover across the collection of food webs were measured using Bray-Curtis dissimilarity in the vegan package (Oksanen *et al.* 2015). Each food web is a realisation of interactions drawn from the metaweb (Dunne 2006), contingent on local species composition and abundances. While the dissimilarity of species between two sites is straightforward to quantify, link dissimilarity must be

decomposed into two parts; differences in interactions between networks originate from differences in species composition, and because shared species between the two realisations may interact differently (Poisot *et al.* 2012). The link dissimilarity presented here is that driven solely by changes in the underlying species composition. In order to assess how species and link turnover changes across the herbivore/predator gradient, we used the number of herbivores and predators within each network as factor levels with which to categorise the food webs (i.e. food webs with 1 herbivore, 2 herbivores, 3 herbivores etc). We ensured that no single food web appeared in more than one group by randomly assigning food webs to either their herbivore or predator group, and calculated the Bray-Curtis dissimilarity between the herbivore and predator groups.

In all models each site was treated as a replicate, as there was no repeat sampling from any one site. Linear regressions were used to test for a relationship between the number of gastropod species and the number of weed species and specialist herbivores within each food web. The count of weed and gastropod species, and herbivores was log(x+0.5) transformed to obtain normality. The relationship between number of herbivores and predators in each network was assessed using a Generalised Linear Model and a Quasiposisson error distribution to account for overdispersion, the number of predators in each network was log(x+0.5) transformed and used as the predictor variable. The relationship between the number of herbivore links and predator links for omnivore nodes only was assessed with linear regression using log(x) transformed predictor and explanatory variables to obtain normality. Due to the extreme distribution produced, the relationship between predatory interaction frequency and herbivory interaction frequency was fitted using LOWESS smoothing. The relationship between weed regulation and the number of herbivores or herbivory interaction frequency was assed using linear regressions where the number of herbivores had been log(x+0.5) transformed and the herbivory interaction frequency was log(x) transformed after removal of zeros by addition of the minimum value, to obtain normality. To directly test if the presence of gastropods interfered

with any weed seed regulation, count of carabid species within each food web was included as an interaction term.

## 4.4 Results

In total 811 unique trophic interactions were found between 41 carabid, 96 weed and 9 gastropod species (Figure 16). In the 'master' amalgamated food web (Figure 16) there were 17 herbivore carabid species, 6 predatory species and 18 omnivore species.

### 4.4.1 Service disruption

No pattern was found in the number of gastropod and weed species within each food web (Figure 17a), however as the number of gastropod species increased, the number of specialist herbivores within each food web decreased ($F_{1,372}$ = 11.8, p=<0.0001, Figure 17b).

There was an inverse relationship between the number of herbivores and predators within each food web ($F_{1,372}$ = 339.5, p=<0.0001, Figure A Appendix F) and between specialist predatory and herbivory interaction frequency (Figure 18). As the predatory interaction frequency increased across the networks, the herbivory interaction frequency reduced dramatically.

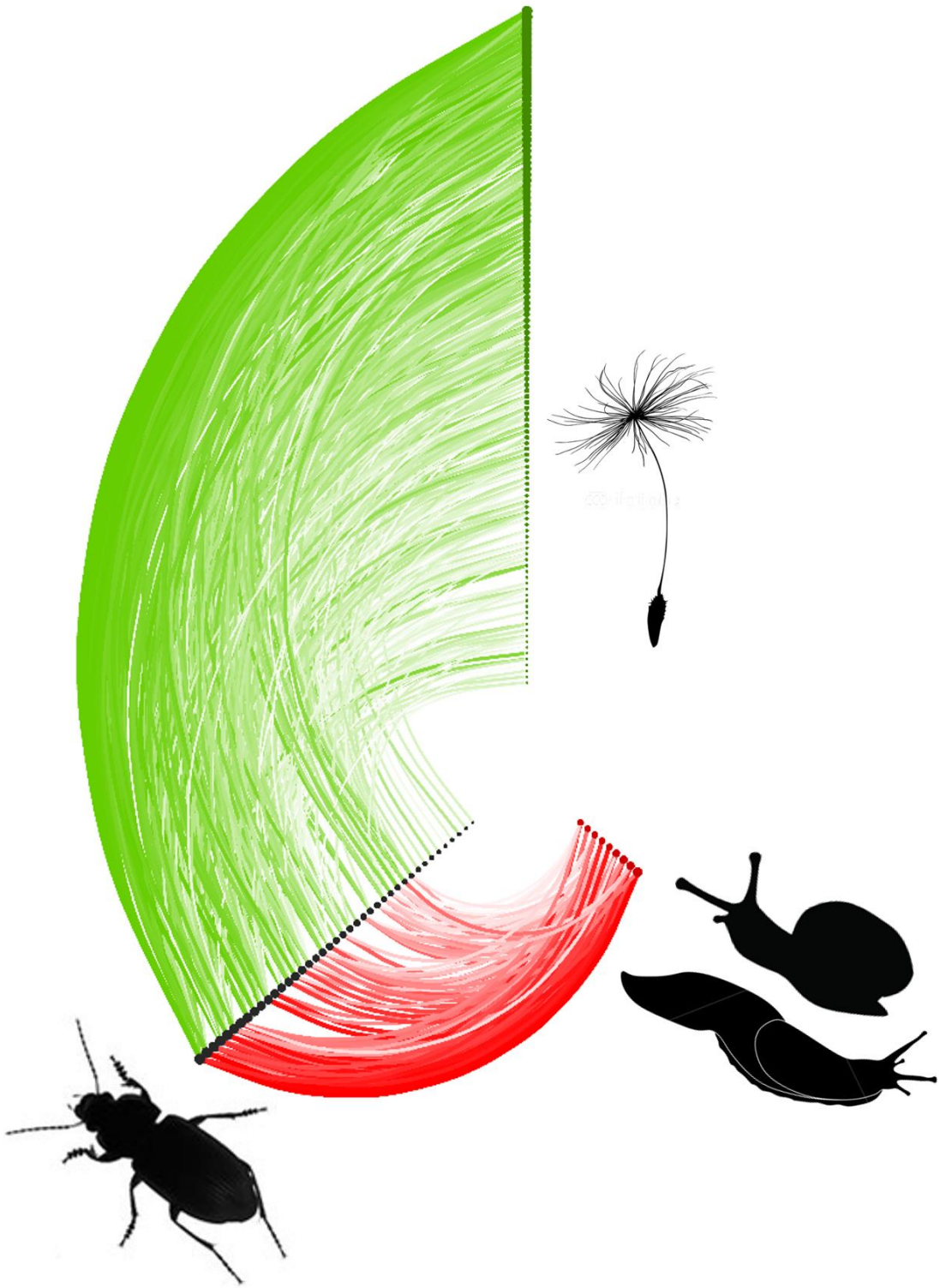


**Figure 16.** The 'master' food web. Carabid (black circles), weed (green circles) and gastropod (red circles) species nodes are sized proportionally to their ranked relative to how often they were found across all food webs. Link colour intensity and thickness is proportional to the strength of the interaction across all food webs.

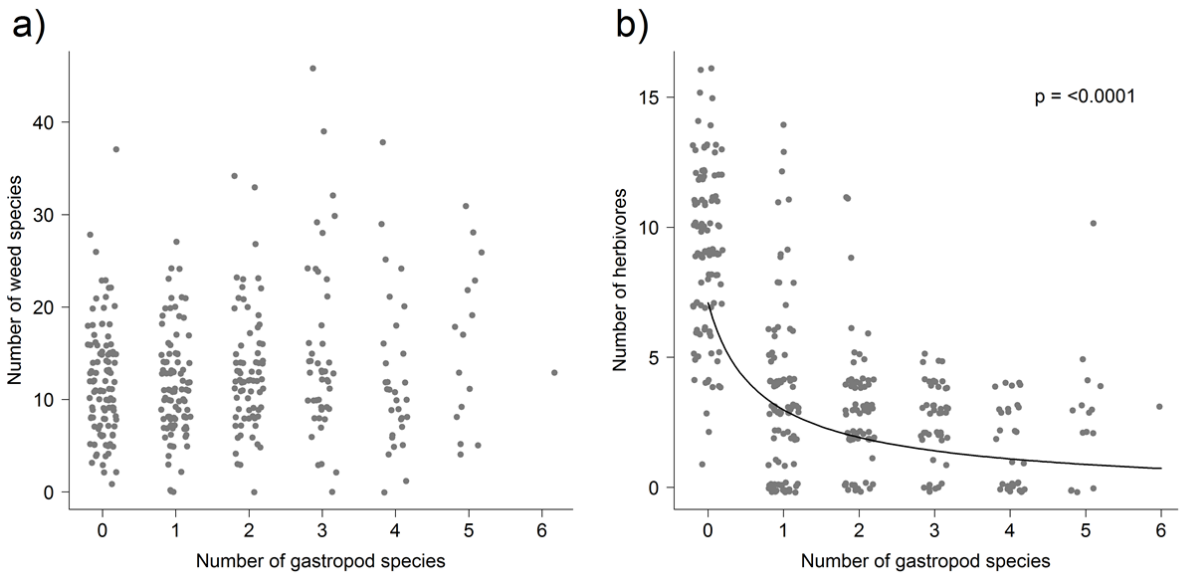



**Figure 17**. The relationships between the number of gastropod and weed species in each food web (a), and between the number of gastropod species and the number of herbivores within each food web (b).

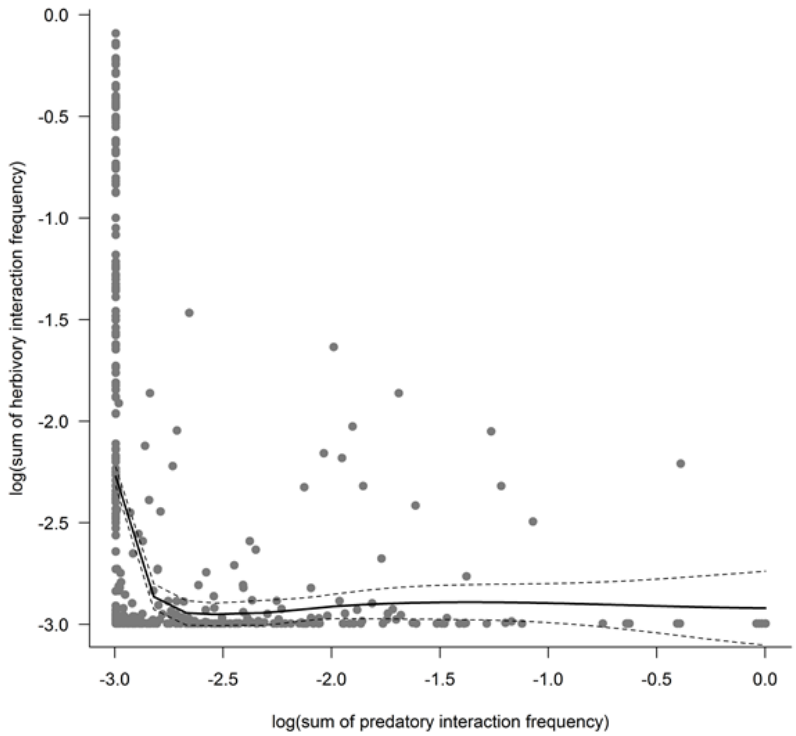



**Figure 18.** The sum of specialist herbivory interaction frequency and sum of specialist predatory interaction frequency for each food web. Lines show a LOWESS smoother with standard error.

Food webs with more specialist herbivores had stronger weed regulation ($F_{1,333}$ = 3.98, p=0.047, Figure B Appendix F), this trend was also found for monocot ($F_{1,333}$ = 6.42, p=0.01) and dicot weed regulation ($F_{1,333}$ = 4.57, p=0.03). This trend was also evident for total weed regulation by food webs with larger specialist herbivore interaction frequencies ($F_{1,333}$ = 5.16, p=0.02, Figure 19a), and also for monocot regulation ($F_{1,333}$ = 3.89, p=0.05) and dicot regulation ($F_{1,333}$ = 5.76, p=0.02). Those food webs which were more dominated by specialist herbivorous interactions more strongly down regulated weed seeds. These relationships were not found when considering the total herbivore interaction frequency (i.e. specialist herbivores plus weed feeding omnivore links) for each food web (Table A Appendix F), suggesting that it is the specialist herbivore interactions which are more strongly related to the level of weed regulation.

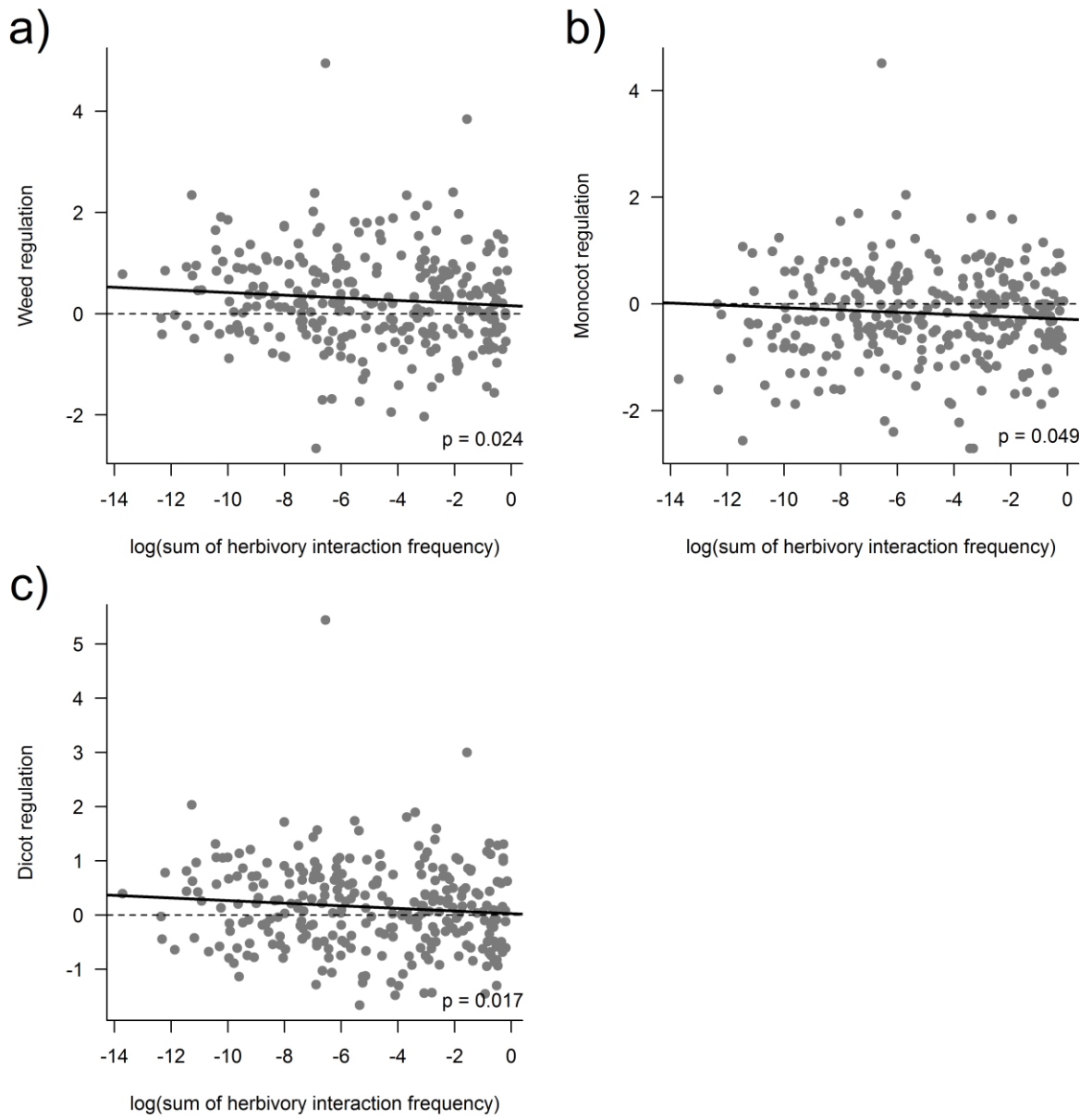



**Figure 19**. The level of total weed regulation (a), monocot regulation (b) and dicot regulation (c) related to the sum herbivory interaction frequency for each network.

This indicates that the presence of gastropods in fields interferes with the ecosystem service of weed regulation, their presence in a food web decreases the number of specialist herbivore carabids (Figure 17b), which in turn weakens the interaction frequency between specialist herbivore carabids and weeds (Figure 18) which is related to a decreases in weed seed regulation (Figure 19).

### 4.4.2 Link turnover

The variation in network structure was driven more strongly by changes in link composition than by changes in species composition (Figure 20). Although changes in carabid species composition were low (mean dissimilarity 0.44 ± 0.17), and most carabid species were found in most food webs, carabids were able to alter their diets across the gradient of available resources such that changes in link composition were higher (mean dissimilarity 0.65 ± 0.25).

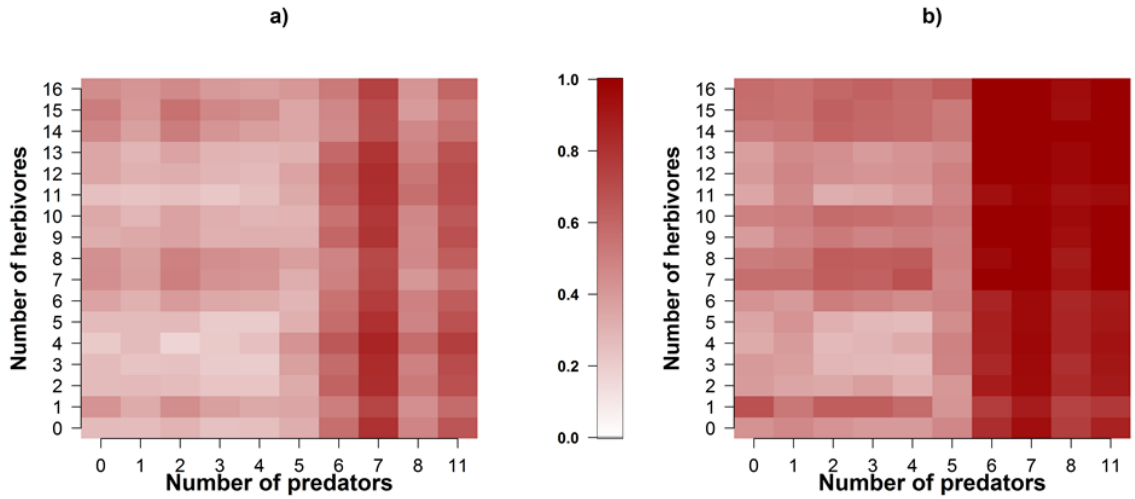



**Figure 20.** The turnover in carabid species (a) and links (b) between the food webs across the herbivore/predator gradient.

### 4.4.3 Management of ecosystem services

For each carabid, in each food web, there was an inverse relationship between the number of links to weed resources and number of links to gastropod resources (Figure 21), even amongst the omnivores, no omnivore node was found to have high numbers of links to both weeds and gastropods (t=-7.08, p=<0.0001, Table A, Figure C Appendix F). The occupancy of the potential link space by the most common carabid species *Pterostichus melanarius* is shown in Figure 21, this distribution was typical for the most abundant carabid species, many of the species occupied most but not all of this link space (Figure D in Appendix F), suggesting that most carabid species can

perform most roles in these networks.

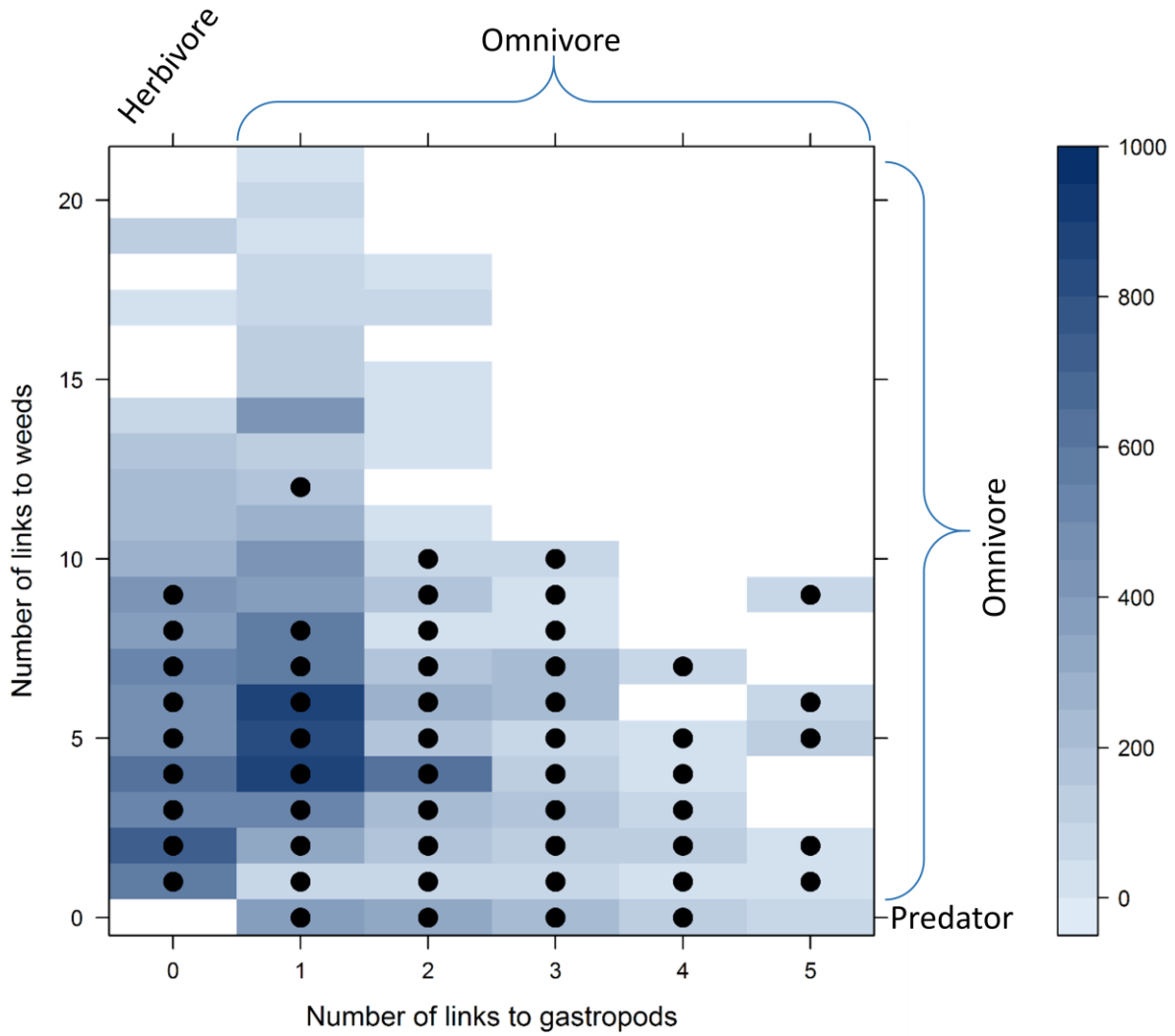



**Figure 21**. A density plot showing the inverse relationship between herbivore links and predator links. Within each food web, for each carabid species, the number of links to weed and gastropod resources are plotted. Some carabids were pure herbivores or predators, but most were omnivores. Colour indicates the count of each particular weed-gastropod link combination. The occupation of this space of potential feeding interactions for the most common carabid species *Pterostichus melanarius* is shown in black.

## 4.5 Discussion

This is the first study of which we are aware that explicitly links the replicated structure of ecological networks to the provisioning of an ecosystem service. We have presented evidence that the presence of alternative resources (gastropods) causes re-wiring within the food web and could disrupt the delivery of effective weed seed control; the presence of gastropods in the food webs reduced the strength of specialist herbivore interactions, which in turn was related to reduced weed seed regulation. Thus when targeting ecosystem management towards the delivery of desired services, it is important to consider the interactions underpinning those services and the local realisation of the regional metaweb, rather than the taxonomically defined trophic designation of a species. For example *P. meinarius* is a highly abundant generalist carabid species consuming both weed seeds and gastropods, but directing conservation towards this species will not deliver stable service provision (Bohan *et al.* 2011a). Directing conservation efforts towards removing carabid-gastropod links from the system (i.e. through the use of molluscicides) might increase the specialist carabid-weed seed interactions and ensure more stable service delivery. However this mechanism would be best tested with experimental evidence, and direct examination of the diet of carabids at the local scale, perhaps through molecular techniques (e.g. Eskelson *et al.* 2011; Lundgren, Saska & Honěk 2013).

The versatility and diet breadth of these species may be an important mechanism maintaining network structure over time, eliminating alternative resources to direct predation pressure on to weed seeds might cause carabid predators to be without resources at certain times of the year (i.e. spring, or before seed fall) and so might be lost from the system. This itself clearly poses a problem for stable service delivery, and how best to manage this system over time is a matter for further research.

Our results demonstrate plasticity in species roles within these networks, sometimes performing as specialist herbivores, sometimes specialist predators and sometimes generalist omnivores, suggesting that the global niche of these species is

modulated at the field scale. Carabid species are traditionally categorised into strict feeding guilds (e.g. Lövei & Sunderland 1996), although increasingly it is being recognised that these species are more versatile in their diet than previously thought (e.g. Lundgren, Saska & Honěk 2013). Our results demonstrate that the role a species plays within a complex food web is defined by its synecology, and thus the contribution each species makes toward a desired level of ecosystem functioning can only be assessed in the context of the food web. Modern conservation literature favours the maintenance of ecosystem functioning and the ecosystem services that result. A food web approach towards service management is advantageous as it explicitly considers the interactions between species which are the drivers of many ecosystem functions (Thompson *et al.* 2012; Gray *et al.* 2014). Our work here has identified a possible mechanism through which weed seed regulation could be enhanced through management targeted at manipulating food web structure, rather than the species composition *per se*.

The networks presented here are 'summary' food webs for the species found in each half field, they are local realisations of the regional metaweb, and hence they are not sensitive to behavioural differences between species across environmental gradients. They are reliant on the trophic information data harvested from the literature and as such there were many species for which no trophic information could be found and so were excluded from the final networks, these were mostly weed species (109), but also some gastropod species (3). Although links established from forced feeding trials were excluded there may be biases in the food webs due to the often small choice range available to carabids in laboratory trials. The sensitivity of these food webs could be improved by gathering more information on the trophic interactions found in these natural systems, for instance screening the guts of carabids for molecular markers (e.g. Lundgren, Saska & Honěk 2013) or resource DNA (e.g. Eskelson *et al.* 2011) would facilitate the identification of species specific trophic interactions. Given these limitations, *a priori* one might expect these food webs to be relatively structurally invariant, thus the trends presented here may in reality be even

more prominent in nature.

Weeds and gastropod pests exert a significant impact on agriculture, and considerable resources are diverted towards controlling these sources of reduced productivity. The negative effects of pesticides on the natural world are well documented (Royal Society 2009). Our results indicate that by harnessing the natural link plasticity found within these food webs, effective weed seed control through targeted management of food web structure could be achieved, potentially reducing the need for some pesticides (e.g. herbicides). Adopting a food web approach links pest populations to food web dynamics and ecosystem service provisioning, which can then be more accurately predicted and managed.

# 5 | The recovery of a freshwater food web from a catastrophic pesticide spill

## 5.1 Summary

Pesticides have strong negative direct effects in fresh waters, but understanding how these effects propagate through natural ecosystems is limited because research that considers the whole ecological community in a natural setting is rare.

Here, we investigate how an accidental spill of the insecticide Chlorpyrifos affects the structure and functioning of a natural river community. We quantify the direct impacts on pesticide sensitive arthropods, and the indirect effects mediated through the food web. We quantify the effect that this food web re-structuring had on a key ecosystem function, leaf litter decomposition. We use data collected regularly for up to 18 months after the spill to investigate the trajectory of recovery.

We find that the biomasses of pesticide sensitive species are reduced, while the biomasses of their competitors and resources increased. Major restructuring occurred within the food web such that the trophic transfer efficiency through the pesticide sensitive nodes was reduced. Invertebrate mediated leaf litter decomposition was reduced while microbial leaf litter decomposition was unchanged. Constrained correspondence analysis showed that community structure recovered by one year on. Ecosystem functioning (leaf litter decomposition) recovered more quickly than the structural aspects of the community, perhaps due to the high redundancy within the assemblage involved in leaf litter decomposition. This work demonstrates the resilience of natural freshwater systems to pesticide spills, this deeper and more holistic understanding will facilitate more effective mitigation and restoration efforts.

## 5.2 Introduction

The global human population is growing rapidly, this coupled with trends towards a more Western diet pattern has created an urgent need to assure future food security (Godfray *et al.* 2010). Meeting this demand for food will lead to ever more increased pesticide use, as the need to control agricultural pests becomes more pressing. While technological advances and innovations such as 'sustainable intensification' (Royal Society 2009) may help us to meet this challenge in the future, in the meantime the use of pesticides in agriculture continues to have wide ranging negative impacts upon biodiversity and ecosystem functioning (Rockström *et al.* 2009). A deeper understanding of the impacts of pesticides at the ecosystem scale is vital if we are to effectively mitigate against these negative impacts.

Most ecotoxicological work to date has necessarily been laboratory or mesocosm based, experiments in the laboratory have revealed precise toxicity and effect levels for individual pesticides on the survival and life-history of target species (e.g. *Gammarus pulex;* Xuereb *et al.* 2007). More complex, community level responses to pesticides have been revealed in micro- and mesocosm experiments (van den Brink *et al.* 1996; Van Wijngaarden *et al.* 1996; Traas *et al.* 2004) and field surveys (Chung, Wallace & Grubaugh 1993; Triebskorn *et al.* 2003; Malaj *et al.* 2014). Little is known about how pesticides affect natural whole communities, or how the ecosystem processes of natural communities are affected, due to the lack of replicated, controlled experiments in a natural setting. Additionally, little work has been done tracking the trajectory of recovery for communities after exposure to pesticides, especially in a natural setting (but see Raven & George 1989; Chung, Wallace & Grubaugh 1993). Predicting how pesticides will influence natural communities and ecosystem properties poses a far larger challenge than predicting responses of individual taxa in isolation (Relyea 2009; Altenburger *et al.* 2013), due to the interplay between direct and indirect effects (Brock, van Wijngaarden & van Geest 2000) and non-additive effects such as synergisms and antagonisms (Relyea & Hoverman 2006).

Chlorpyrifos is a widely used broad-spectrum organophosphate pesticide (insecticide and acaricide) which attacks insect (and arachnid) nervous systems. It is relatively non-persistent, measured natural water column half-lives for chlorpyrifos typically range from < 1 to 4.8 days, (Racke 1993; Barron & Woodburn 1995), and its principal degradation products are less toxic than the parent chemical leading its direct (toxic) effects to be relatively short-lived (Kramer *et al.* 1997). Specifically, chlorpyrifos can be toxic to most invertebrate and fish species, but often only at high concentrations, (Barron & Woodburn 1995). Within the invertebrates, crustaceans and invertebrate larvae are among the most sensitive species for instance the lethal concentration ($LC_{50}$) for the freshwater shrimp *G. pulex* is 0.07µg/L (Barron & Woodburn 1995). Molluscs are among the most resistant species, with many taxa having an $LC_{50}$ of >100µg/L (Barron & Woodburn 1995). Most fish species are also sensitive to chlorpyrifos toxicity given high enough concentrations, for instance, the $LC_{50}$ for Rainbow Trout (*Oncorhynchus mykiss*) is 7µg/L (Barron & Woodburn 1995). Compartments of the community not directly affected by chlorpyrifos toxicity are affected indirectly (see reviews by Barron & Woodburn 1995; Brock, van Wijngaarden & van Geest 2000; Giddings *et al.* 2014). For instance, Chlorpyrifos has been found to alter food-web structure in microcosms (Traas *et al.* 2004). Little work has been done to investigate the effects of Chlorpyrifos on the microbial community, which is an important component of the food web and contribute towards many ecosystem functions. Little work has been done linking chlorpyrifos exposure to ecosystem functioning, however mesocosm studies have found macroinvertebrate-mediated litter breakdown to be depressed (Brock *et al.* 1993; Cuppen *et al.* 1995).

Ecological networks such as food webs are a useful tool for studying the effects of stressors on communities (Ings *et al.* 2009; Thompson *et al.* 2012; Gray *et al.* 2014). A food web based approach to the study of pesticide exposure in natural ecosystems is particularly useful as it implicitly encompasses the range of indirect effects, allowing the full impacts of the pesticide on the whole community to be measured. The use of

trivariate food webs (Cohen, Jonsson & Carpenter 2003), whereby the mass and abundance of each node (taxa) in the network is incorporated, can reveal changes to the energy transfer efficiency through the food web. Study of the structure of the food web has provided explanations for unexpectedly slow recovery from perturbations in other systems (Scheffer & Carpenter 2003; Layer *et al.* 2011), and it might provide an explanation for the often slow recovery of natural communities following chlorpyrifos exposure (Raven & George 1989). Food web models have been used to successfully predict the indirect effects of mixtures of pesticides on experimental communities (Halstead *et al.* 2014), but this is yet to be applied to natural communities. Much of the research examining the indirect effects of pesticides on communities and ecosystem functioning has been descriptive rather than mechanistic (but see Traas *et al.* 2004; Halstead *et al.* 2014). As of yet the food web of natural communities exposed to chlorpyrifos as not been investigated in a controlled and repeated manner.

There is a need for controlled, replicated investigations into the direct and indirect effects of pesticides in natural ecosystems, as well as the trajectory of recovery that these communities follow. A spill of the insecticide chlorpyrifos in the River Kennet, Wiltshire UK, on 1st July 2013 provided an opportunity to address this gap. The River Kennet is a lowland chalk tributary of the River Thames in southern England, designated as a UK Site of Special Scientific Interest (SSSI). Its diverse macroinvertebrate fauna is dominated by Gammaridae, Baetidae, Ephemerellidae, Simuliidae and Chironomidae, which support an economically important salmonid game fishery (Wright *et al.* 2002, 2004). The spill of chlorpyrifos was likely to be a 'down-the-drain' incident, and entered the river through a water treatment works. Concentrations of 0.52–0.82µg/L were recorded coming from the main tertiary sewage treatment works in Marlborough, Wiltshire, on 2 and 5 July, respectively (Appendix C; Thompson *et al.*). The peak concentration was most likely missed, but even the recorded concentrations are sufficient to be acutely toxic to arthropods (Barron & Woodburn 1995; Giddings *et al.* 2014), particularly over extended periods

(i.e. >24h; Rubach *et al.* 2011).

Previous work has demonstrated the immediate effects of this spill on both community structure and measures of ecosystem functioning (Appendix C; Thompson *et al.*). Crucially, Thompson *et al.* (2015) used before-after-control-impact (BACI) data, allowing them to detect causal relationships between the pesticide spill and reduced invertebrate abundances. Thompson *et al.* (2015) found that the biomass of pesticide sensitive species, in particular the keystone detritivore *G. pulex*, was reduced in sites impacted by the pesticide, relative to control sites upstream of the spill. Indirect effects were detected as the biomass of non-pesticide sensitive taxa such as oligochaete worms, increased. Chlorophyll-α concentration, which is a proxy for algal biomass, was also increased. When the food web was plotted in trivariate spece, whereby the nodes of the network are plotted by their mass and abundance, the slope of the interactions through the pesticide sensitive nodes steepened in response to the pesticide. The slope of these interactions is a proxy for energy transfer through the food web, steeper slopes indicating reduced energy transfer efficiency. These changes to the structure of the food web was found to alter ecosystem functioning rates; invertebrate mediated leaf litter decomposition was reduced in impacted sites, while microbe driven decomposition was increased.

We build upon the work done by Thompson *et al.* (2015), and demonstrate the recovery of the river ecosystem one year on. We quantify the initial impact and recovery of the biomass of sensitive and non sensitive invertebrate species, use detailed mass and abundance data to construct quantified food webs to make specific predictions about the direct and indirect effects mediated through the food web, and investigate the effects on and recovery of leaf litter decomposition.

Specifically we will address the following hypotheses:

1. Community structure will be significantly impacted by the pesticide spill at impacted sites, and recover towards that of the control sites by one year on.
2. The biomass of pesticide sensitive taxa which have adult aerial life stage, such as the mayfly *Baetis sp*., will recover more quickly than those who do not, such as *G. pulex*.
3. The biomass of oligochaete worms and chlorophyll-α concentration, which were higher in impacted sites immediately after the spill, will recover to become indistinguishable from that of the control by one year on.
4. As the biomass of pesticide sensitive species recovers, so too will the energy transfer efficency through those nodes; the slope of the feeding interactions between pesticide sensitive taxa and their consumers and resources will become shallower over time.
5. The rate of leaf-litter decomposition at impacted sites will recover to be indistinguishable from that at the control sites by one year on. Microbial decomposition rates will recover more quickly than invertebrate mediated decomposition.
6. Structural changes to the community will take longer to recover than invertebrate mediated decomposition rates due to high redundancy and a mix of sensitive and non-sensitive species involved in decomposition.

## 5.3 Methods

Sampling began in July 2013, approximately 2 weeks after the pesticide spill. Seven sites were selected with similar channel forms and riparian surroundings, located approximately 1km apart, except for the most downstream site which is located approximately 5km from the second most downstream site (Figure 22). Each sampling round was repeated in full every two months (with the exception of electrofishing which was not done in every sampling round) up until the 9th sampling round in March 2015, when there was a six month gap before the final (10th) sampling round in September 2015. Due to limited resources not all samples from each site and date could be processed, details of which samples, from which site at which time points are presented below (Table 6).

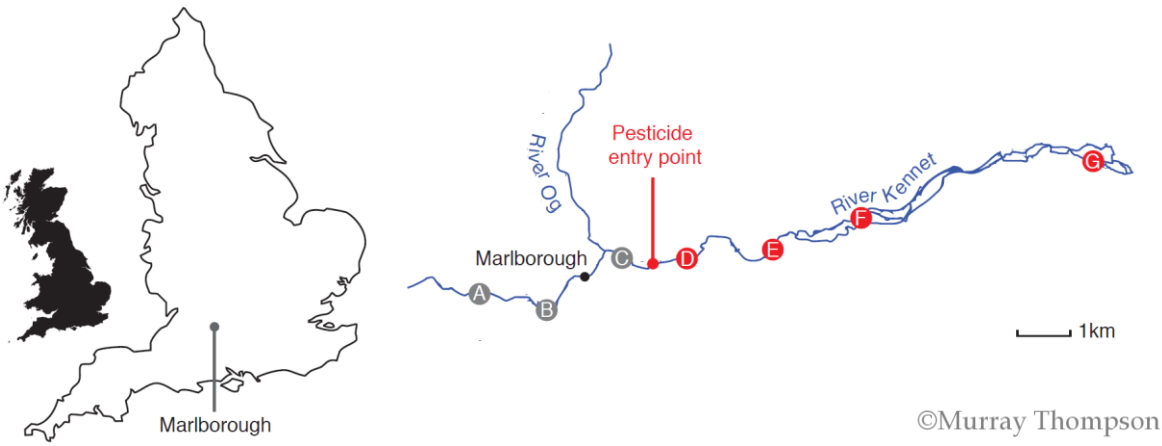



Figure 22. Map of the River Kennet with study sites marked; three upstream of the pesticide entry point in grey, and four below in red.

### 5.3.1 Community structure

To quantify the chlorophyll-α concentration, a proxy for algal biomass, ten stones were randomly selected from each reach, and using a quadrat of known area (17.28cm$^2$) of the upper surface scraped and washed into a sample bottle. Samples

were stored in the dark and frozen at -20°C until analysis. Data presented here were from six samples (Table 6), from sites A-F, and four of the total ten time points. The laboratory procedure was as follows: each sample was filtered through a Whatman GF/C glass fiber filter placed in a Büchner funnel by applying a vacuum. The filter was then quickly placed into 20ml of ice-cold 96 % ethanol, mashed up slightly with a glass rod, and placed on ice covered with a dark plastic bag to prevent degradation of chlorophyll molecules, which become highly sensitive to UV light during extraction. Samples were kept at 4°C in the dark overnight. 200µl of sample was transferred to a Nunc™ MicroWell™ 96-well optical bottom plate and chlorophyll-α content was then measured spectrophotometrically using a Biotek HT absorbance reader (Biotek, Swindon, U.K.). Absorption of the extraction was measured at 664 (chlorophyll-α) and 750 nm (turbidity), with and without the addition of 50µl of 0.1M HCL. The addition of HCL degraded the chlorophyll-α into its degradation products, pheophytins and pheophorbides (Steinman, Lamberti & Leavitt 2006), hence the skew in absorption values created as a result of degradation products and not chlorophyll-α was removed (Lorenzen 1967). The absorbance of the sample in a microplate was converted into a 1-cm path length corrected absorbance using the measured path length (Warren 2008):

$$A_{1cm} = \frac{A_{microplate} - B}{\pi r^2} \quad (1)$$

Where $A_{microplate}$ is the absorbance reading taken using the 96-well plate reader, B is the mean absorbance of 10 control wells containing only 200µl of ethanol, and $r$ is the radius of the well (0.325cm).

After path length correction, the chlorophyll-α and pheophytin concentration was calculated using the following equation (modified from Steinman, Lamberti & Leavitt 2006):

$$\text{Chlorophyll-}\alpha\ \&\ \text{pheophytin (mm/m}^2) = \frac{(A_{664} - A_{750}) \cdot E \cdot 10^4}{83.4 \cdot T} \quad (2)$$

$$\text{Pheophytin (mm/m}^2) = \frac{(A_{HCL\,664} - A_{HCL\,750}) \cdot E \cdot 10^4}{83.4 \cdot T} \quad (3)$$

Where $A_{664}$ and $A_{750}$ is the absorbance at 664nm and 750nm respectively, $A_{HCL\,664}$ and $A_{HCL\,750}$ is the absorbance after the addition of HCL at 664nm and 750nm respectively. $E$ is the volume of ethanol used (20ml), 83.4 is the absorption coefficient for chlorophyll-α in ethanol ($1g^{-1}cm^{-1}$), and $T$ is the scraped area of stone ($17.28cm^2$). Finally the concentration of chlorophyll-α was calculated by subtracting equation (3) from equation (2).

The diatom assemblage was characterised by selecting ten permanently submerged stones at each study site from unshaded areas. The biofilm was scrubbed and washed from the upper stone surface. Samples were immediately preserved by addition of Lugol's iodine and stored until further processing. Data presented here were from one sample processed (Table 6) from four sites (A, C, D and F) at 2 time points (September 2013 and September 2014). Slide preparation followed Battarbee *et al.* (2001), a minimum of 300 diatom valves were identified to species per sample using standard keys (see Appendix G) and abundances per unit area were determined as in Battarbee (1973). Linear dimensions were measured to the nearest 1µm to estimate diatom biovolume (Hillebrand *et al.* 1999). The first 30 specimens of all common (n > 30) species were measured, and where species were encountered less frequently, all specimens in the count were measured. Carbon content of the diatoms was estimated (Rocha & Duncan 1985) and then converted to dry mass (Sicko-Goad, Schelske & Stoermer 1984).

To determine the abundance of the invertebrates, ten Surber samples (area $0.0625m^2$; mesh 330µm) were taken from randomly chosen stony riffles at each site within a ~$50m^2$ stretch. Samples were immediately preserved in 96% ethanol until

further processing. In the laboratory, samples from four sites (A, C, D and F) and four time points (September 2013, March 2014, September 2014 and March 2015) were prioritised (Table 6). For each site, in each time point three samples were processed; invertebrates were sorted from debris, identified to species where possible (i.e. all except Diptera [identified to Family] and Annelidae [identified to Subclass]), and counted. Individuals were identified using a combination of published identification keys (Appendix G). Body size measurements were taken in the form of linear dimensions (head-capsule width or body length) for up to 60 specimens of each species in each site at each time point using a calibrated ocular micrometer, and individual dry mass determined from published length-dry mass regression equations (Table A Appendix G).

Quantitative depletion electrofishing was undertaken at 6 time points (September 2013, March 2014, September 2014, November 2014, March 2015 and September 2015; some time points were omitted to minimise stress on the fish assembledge) to assess fish abundance (Table 6). At each site a 50m stretch of the river was electrofished (after Carle & Strub 1978). Stop-nets were installed at both ends of the stretch, and three runs were completed, moving upstream and sweeping from one side of the river to the other. All fishes were counted and measured (fork length and body mass) before being released back into the stream alive. Population densities were estimated using the R package FSA (Ogle 2012) and iterative Maximum Weighted Likelihood statistics (Carle & Strub 1978). For each species, individual dry mass was calculated using length–mass regression equations (Table A Appendix G).

**Table 6.** The numbers of samples processed for each data type. Numbers in black indicate data presented here, in grey are data processed but not presented here. In all instances (except fish) gaps represent samples collected (see main text for numbers) but not yet processed. Fish were only surveyed on the dates indicated, only those indicated with a black Y are presented here.

| | Site | Jul-13 | Sep-13 | Nov-13 | Mar-14 | May-14 | Jul-14 | Sep-14 | Nov-14 | Mar-15 | Sep-15 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Chlorophyll | A | | **6** | | **6** | 6 | 6 | **6** | 6 | **6** | |
| | B | | **6** | | **6** | 6 | 6 | **6** | 6 | **6** | |
| | C | | **6** | 6 | **6** | 6 | 6 | **6** | 6 | **6** | |
| | D | | **6** | 6 | **6** | 6 | 6 | **6** | 6 | **6** | |
| | E | | **6** | 6 | **6** | 6 | 6 | **6** | 6 | **6** | |
| | F | | **6** | 6 | **6** | 6 | 6 | **6** | 6 | **6** | |
| | G | 6 | 5 | | 6 | 6 | 6 | | 6 | 6 | |
| Diatoms | A | | **1** | | 3 | | | **1** | | | |
| | B | | | | 3 | | | | | | |
| | C | | **1** | | 3 | | | **1** | | | |
| | D | | **1** | | 3 | | | **1** | | | |
| | E | | | | 3 | | | | | | |
| | F | | **1** | | 3 | | | **1** | | | |
| | G | | | | 3 | | | | | | |
| Invertebrates | A | | **4** | | **3** | | | **3** | | **3** | |
| | B | | 1 | | | | | | | | |
| | C | | **3** | | **3** | | | **3** | | **3** | |
| | D | | **3** | | **3** | | | **3** | | **3** | |
| | E | | 1 | | | | | | | | |
| | F | | **3** | | **3** | | | **3** | | **3** | |
| | G | | 1 | | | | | | | | |
| Fish | A | | **Y** | | Y | | | **Y** | Y | Y | Y |
| | B | | Y | | Y | | | Y | Y | Y | Y |
| | C | | **Y** | | Y | | | **Y** | Y | Y | Y |
| | D | | **Y** | | Y | | | **Y** | Y | Y | Y |
| | E | | Y | | Y | | | Y | Y | Y | Y |
| | F | | **Y** | | Y | | | **Y** | Y | Y | Y |
| | G | | Y | | Y | | | Y | Y | Y | Y |
| Leaf litter decomposition | A | | **10** | | **10** | | | **10** | | **10** | |
| | B | | 10 | | 10 | | | | | | |
| | C | | **10** | | **10** | | | **10** | | **10** | |
| | D | | **10** | | **10** | | | **10** | | **10** | |
| | E | | 10 | | 10 | | | | | | |
| | F | | **10** | | **10** | | | **10** | | **10** | |
| | G | | 10 | | 10 | | | | | | |

### 5.3.2 Food web construction and analysis

Quantitative trivariate food webs were constructed, where nodes (species) are plotted by their mass and abundance along with their interactions. This was done for four sites (A, C, D & F) and two time points (September 2013 & September 2014), which produced in total eight food webs. Feeding links were established by inferring links from the literature, and filled in for each network using the WebBuilder function (Chapter 2; Gray *et al.* 2015) in R (R Core Team 2013). This method is based on the assumption that a described feeding link would be realized between two species at a given study site if the same link has been described in another system where those species co-exist (e.g. Hall & Raffaelli 1991; Martinez 1991; Layer *et al.* 2010b; Pocock, Evans & Memmott 2012). In some instances, due to a paucity of trophic interaction data, feeding links were assigned on the basis of taxonomic similarity, (Table B Appendix G).

The slope of every trophic interaction in trivariate space was calculated using the method of Cohen *et al.* (2009) in the R package Cheddar (Hudson *et al.* 2013). We used link slopes to estimate changes in potential biomass flux between a resource and its consumer (Thompson *et al.*). A trophic interaction can be viewed as a vector from a resource to its consumer in mass-abundance space, a steepening of this slope indicates less efficient energy transfer and reduced biomass flux (Cohen *et al.* 2009).

### 5.3.3 Leaf litter decomposition

Breakdown rates of black alder (*Alnus glutinosa*) leaf litter was determined for each site at each time point. Ten replicates of bags with each of two mesh sizes were deployed in each site at each time point, each bag containing 3.00±0.3 g of air-dried leaf litter. Bags with fine mesh size (500µm) were used to exclude invertebrate detritivores, while coarse mesh size (10mm) allowed invertebrates access to the leaf litter. The bags were left to incubate in the river for 9 days, when collected they were frozen at -20°C. In the laboratory samples from four sites (A, C, D and F) and four time points (September 2013, March 2014, September 2014 and March 2015) were

prioritised for processing (Table 6). All ten replicate samples were processed. The leaf-litter was extracted and oven-dried at 80°C and re-weighed to determine the proportion remaining.

Leaf breakdown rates were expressed as the exponential decay rate coefficient, $k$ (after Woodward *et al.* 2012b):

$$k = -ln\left(\frac{m_1}{m_0 \cdot c}\right)/dd \quad (5)$$

Where $m_0$ is the initial leaf litter weight, and $m_1$ is the final leaf litter weight, $c$ is an air-dry to oven-dry conversion factor (calculated separately; 0.968) and $dd$ is the number of degree-days (the temperature multiplied by the number of incubation days). Total ($k_{total}$) and microbially mediated breakdown rates ($k_{microbe}$) were determined from the coarse-mesh and fine-mesh bags, respectively. Rates of invertebrate-mediated breakdown were derived by calculating the percent of litter mass remaining in coarse-mesh and fine-mesh bags in each bag pair and then calculating a new $k$ value ($k_{invertebrate}$) based on this difference:

$$k_{invertebrate} = 1 - \left((1 - P_{coarse}) - \left(1 - P_{fine}\right)\right) \quad (6)$$

Where $P_{coarse}$ is the proportional weight of leaf litter remaining in a coarse bag, and $P_{fine}$ is the proportional weight of leaf litter remaining in a fine bag.

To control for seasonal temperature differences across the sampling time points, temperature data were obtained for Environment Agency monitoring stations located within the sampling area. One temperature reading was selected to coincide with each sampling time point, where possible temperature data from the same site were used.

### 5.3.4 Statistical analysis

All statistical analysis was done in R (R Core Team 2013). Constrained Correspondence Analysis (CCA), with sampling time point (September 2013 or September 2014) fitted as the sole constrained axis, was done to assess the impact of the pesticide spill on, and recovery of, community structure. The explanatory power of time, treatment and an interaction between the two was assessed using Permutational multivariate analysis of variance (PERMANOVA) with 9999 permutations.

Generalised linear mixed effects models (GLMM) were used to test for a significant effect of time and treatment on a range of response variables; chlorophyll-α concentration, the biomass of some pesticide sensitive taxa (*Baetis*, *G. pulex*), non-pesticide sensitive taxa (Oligochaeta), link slopes and decomposition rates. In all cases treatment was nested within time in order to assess any differences between control and impact at each time point. Site and sample month was treated as having a random effect on the intercept of the linear relationship. Chlorophyll-α concentration and biomasses were $\log_{10}$ transformed to meet test assumptions. All GLMMs were performed using the nlme package in R (Pinheiro *et al.* 2014).

## 5.4 Results

### 5.4.1 Community structure

The CCA model revealed that community structure was significantly affected by the pesticide spill (**Figure 23**), and that the community had recovered one year on from the spill (indicated by the reduction in distance separating the two treatments over time). PERMANOVA revealed that the interaction between time and treatment had an effect on community structure (Table 7). The treatment/time component of the CCA model explained 30% of the variation. This confirms our first hypothesis.

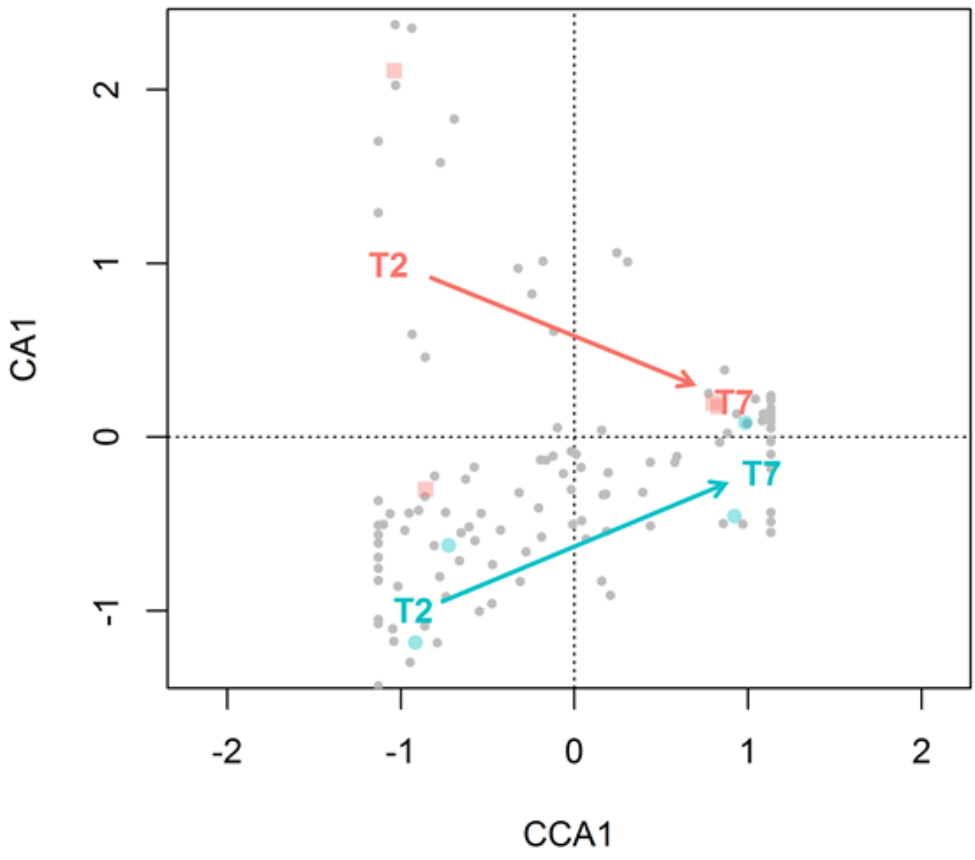



**Figure 23**. CCA plot with time fitted as the sole constrained variable. Site scores are shown as either red squares (impacted sites), or blue circles (control sites). Species scores are shown in grey. Time points are separated; T2 = September 2013, T7 = September 2014.

**Table 7**. Effect of treatment and time on community structure (PERMANOVA).

| Source | d.f. | SS | Pseudo-F | P-value |
|---|---|---|---|---|
| Treatment/time | 3 | 1.11906 | 1.0881 | 0.0446 |
| treatment | 1 | 0.55254 | 1.0746 | 0.1092 |
| Residual | 4 | 2.05681 | | |

Both time and treatment had an effect on chlorophyll-α concentrations, (Table 8), chlorophyll-α concentrations were higher in impacted sites after the pesticide spill but became indistinguishable from that of the control sites by one year on (September 2014) indicating that the algal assemblage was indirectly affected by the pesticide spill (Figure 24), although the chlorophyll-α concentration at impacted and control sites did diverge again in March 2015. This is in accordance with anecdotal evidence, there was visible evidence of a large algal bloom in the months after the spill.

Invertebrate biomass was also affected by both treatment and time (Table 8); the biomass of *Baetis sp.* was lower at impacted sites immediately after the pesticide spill (Figure 25) and recovered to become indistinguishable from that of the control sites by March 2014. *G. pulex* in contrast had a lower biomass in impacted sites through almost the entire sampling period (Figure 25) only returning to become indistinguishable from that of the control sites in March 2015. This confirms our second hypothesis, *Baetis sp*. recovered more quickly than *G. pulex* following the pesticide spill. Oligochaete worms had a higher biomass in impacted sites after the spill (Figure 25). Although the oligochaete biomass did not fall, it became indistinguishable from that of the control sites over time, confirming our third hypothesis.

Fish biomass was unaffected by treatment and time (Table 8), suggesting that the fish community was unaffected by the pesticide spill.

**Table 8**. Statistics of fit for the multiple mixed effects models. All models include a main effect of treatment, time nested within treatment, and a random effect of site on the intercept of the linear relationship.

| Response variable | Predictor variable & interactions | d.f. | F-value | P-value |
|---|---|---|---|---|
| Log(Chlorophyll-α concentration) | Time/treatment | 4 | 7.80 | 0.0001 |
| | Time | 3 | 28.64 | 0.0001 |
| Log(*Baetis sp.* biomass) | Time/treatment | 4 | 3.67 | 0.0144 |
| | Time | 3 | 2.12 | 0.1168 |
| Log(*G. pulex* biomass) | Time/treatment | 4 | 16.17 | 0.0001 |
| | Time | 3 | 5.99 | 0.0021 |
| Log(*Oligochaeta* biomass) | Time/treatment | 4 | 7.06 | 0.0003 |
| | Time | 3 | 4.92 | 0.0062 |
| Log(Fish biomass) | Time/treatment | 3 | 0.307 | 0.8199 |
| | Time | 2 | 0.094 | 0.9099 |
| Algae – arthropod link slopes | Time/treatment | 2 | 107.9 | 0.0001 |
| | Time | 1 | 0.06 | 0.805 |
| Algae – non-arthropod link slopes | Time/treatment | 2 | 7.45 | 0.0006 |
| | Time | 1 | 0.0001 | 0.999 |
| Arthropod – fish link slopes | Time/treatment | 2 | 0.879 | 0.416 |
| | Time | 1 | 0.513 | 0.474 |
| Non-arthropod – fish link slopes | Time/treatment | 2 | 0.501 | 0.607 |
| | Time | 1 | 0.0951 | 0.758 |
| Log(invertebrate mediated decomposition) | Time/treatment | 4 | 26.82 | 0.0001 |
| | Time | 3 | 15.17 | 0.0001 |
| Log(Microbial mediated decomposition) | Time/treatment | 4 | 1.97 | 0.1028 |
| | Time | 3 | 12.78 | 0.0001 |
| Log(Total decomposition) | Time/treatment | 4 | 43.54 | 0.0001 |
| | Time | 3 | 57.36 | 0.0001 |

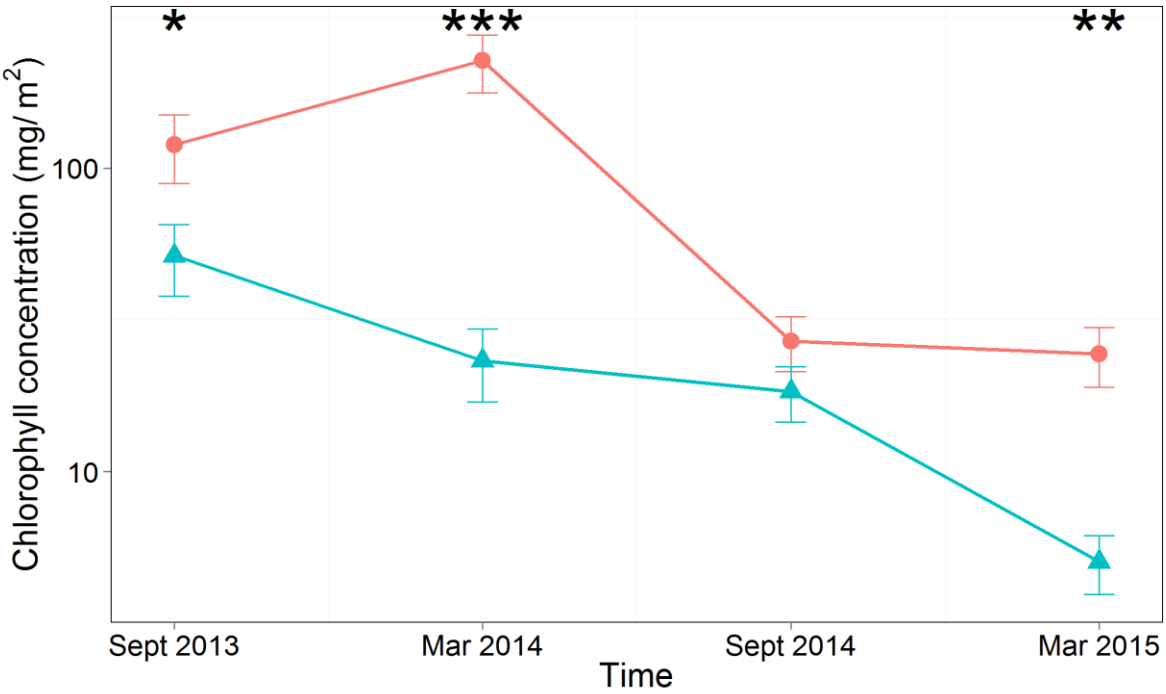



**Figure 24**. The concentration of Chlorophyll-α over time in impacted (red, circles) and control (blue, triangles) sites. Points show mean values, ± s.e.m. Those time points where impact and control concentrations are different are indicated by significance stars; * = p = 0.05 to 0.01, ** = p = 0.01 to 0.001, *** = p <= 0.001 = ***.

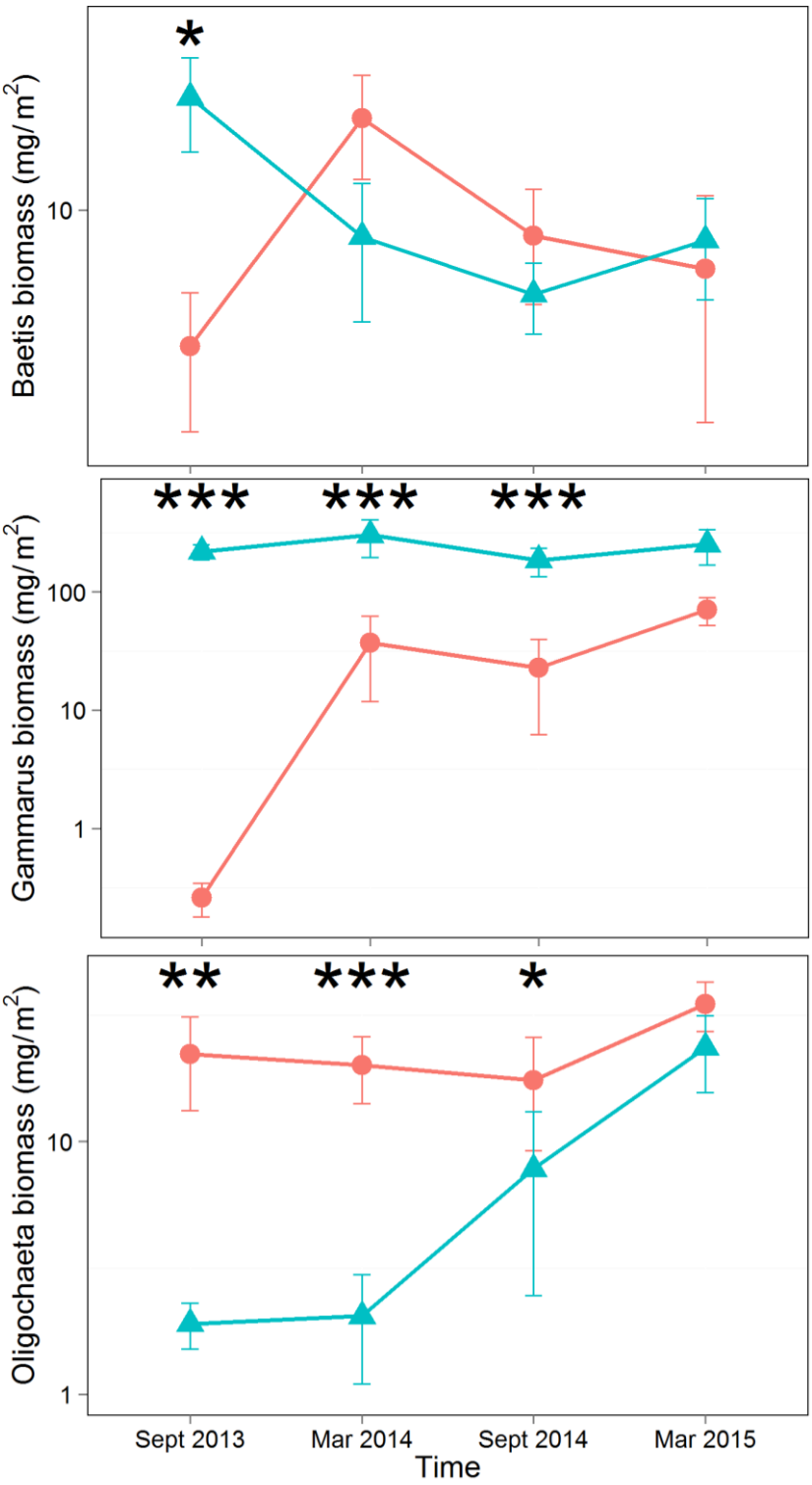



**Figure 25**. The biomass of *Baetis sp*. (top), *G. pulex* (middle) and Oligochaeta (bottom) over time in impacted (red, circles) and control sites (blue, triangles). Points show mean values, ± s.e.m. Those time points where impact and control biomasses are different are indicated by significance stars; * = p = 0.05 to 0.01, ** = p = 0.01 to 0.001, *** = p <= 0.001 = ***.

### 5.4.2 Food web

The food web experienced significant re-structuring as a result for the pesticide (Figure 26), there was a thinning of the middle of the food web as pesticide-sensitive species were lost. In particular, *G. pulex*, which is an important detritivore and a keystone species (i.e. a species more abundant than expected for its size) was much reduced in both size and abundance and so lost its keystone position within the food web (black triangles in Figure 26). By September 2014 it had recovered to its normal positon within the food web. *Baetis sp.*, another pesticide sensitive taxon was also reduced in both body mass and abundance in impacted sites. It also recovered by September 2014 (black upside down triangles in Figure 26). In contrast, Oligochaeta increased in both body size and abundance in impacted sites to take a dominant position within the food web, it still occupied this position in trivariate space in September 2014 (black squares in Figure 26). There was an appearance of larger diatoms, and larger diatom species immediately after the spill, these were lost by September 2014 (Figure 26), which is in accordance with our third hypothesis.

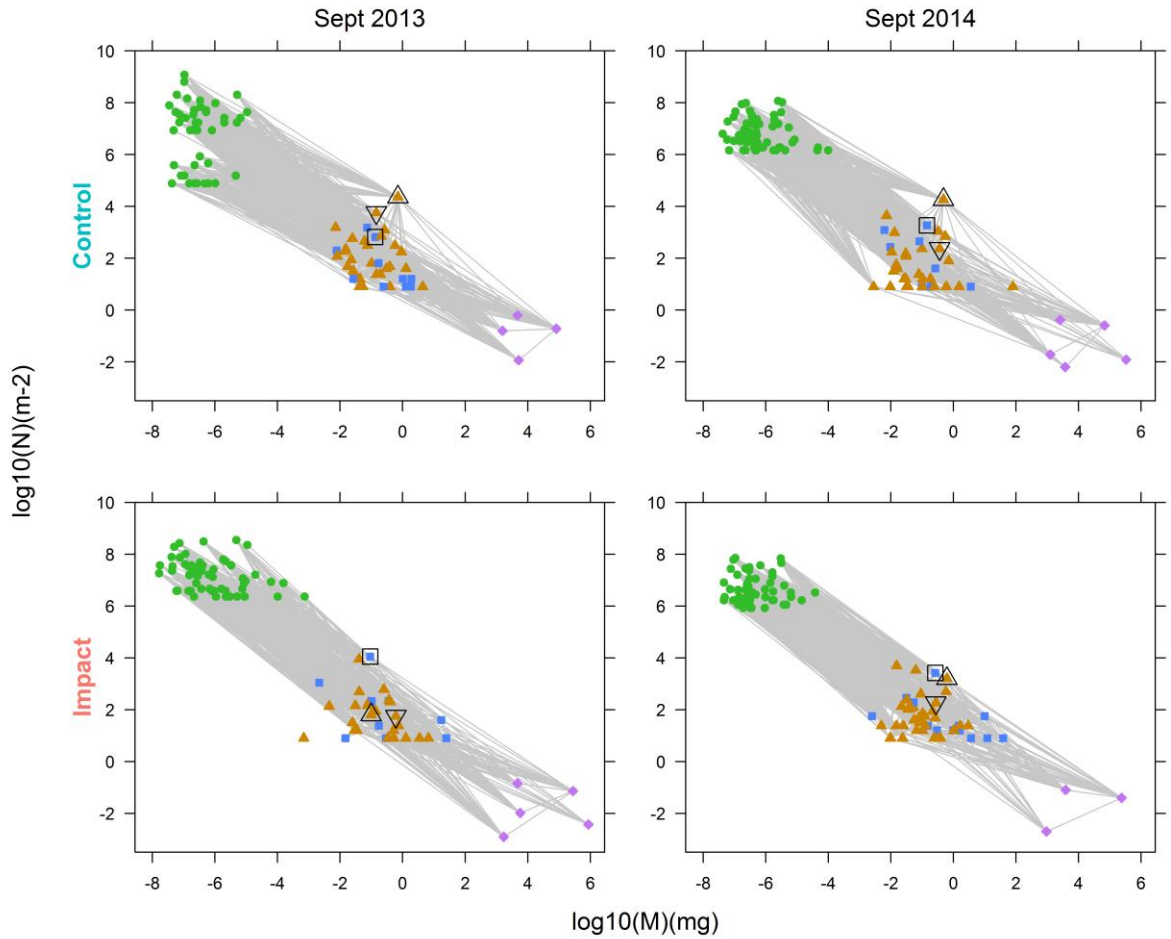



**Figure 26.** Trivariate food webs where species are plotted by their average mass on the x-axis and abundance on the y-axis. Trophic interactions between species are shown in grey. Green circles = algae, orange triangles = arthropods, blue squares = non-arthropod invertebrates, pink diamonds = fish. The locations of some key taxa have been highlighted; black upward pointing triangles = *G. pulex*, black downward pointing triangles = *Baetis sp.*, black squares = Oligochaetes.

This re-structuring of the food web is reflected in the changes in link slopes between resources and consumers within the food web (Figure 27; Table 8), in support of our fourth hypothesis. Link slopes between arthropods and their algae resources were more negative (i.e. steeper) in impacted sites than control sites in September 2013 (top left Figure 27). This indicates altered mass-abundance scaling within the food web and reduced energy transfer efficiency through the algae-arthropod pathway

within impacted communities. Link slopes between algae and non-arthropod invertebrates were also affected by the pesticide (Table 8). Although pairwise mean comparisons revealed no within time point differences between impacted and control food webs there was a trend towards progressively less negative (i.e. shallower) primary link slopes for non-arthropod invertebrates in impacted sites over time (top right Figure 27). This indicates that over time the energy efficiency transfer from algae through non-arthropod invertebrates within impacted food webs increased, and decreased in control sites.

There was no effect of treatment and time of the slope of the links between arthropod and non-arthropod resources and their fish consumers (Table 8, Figure 27).

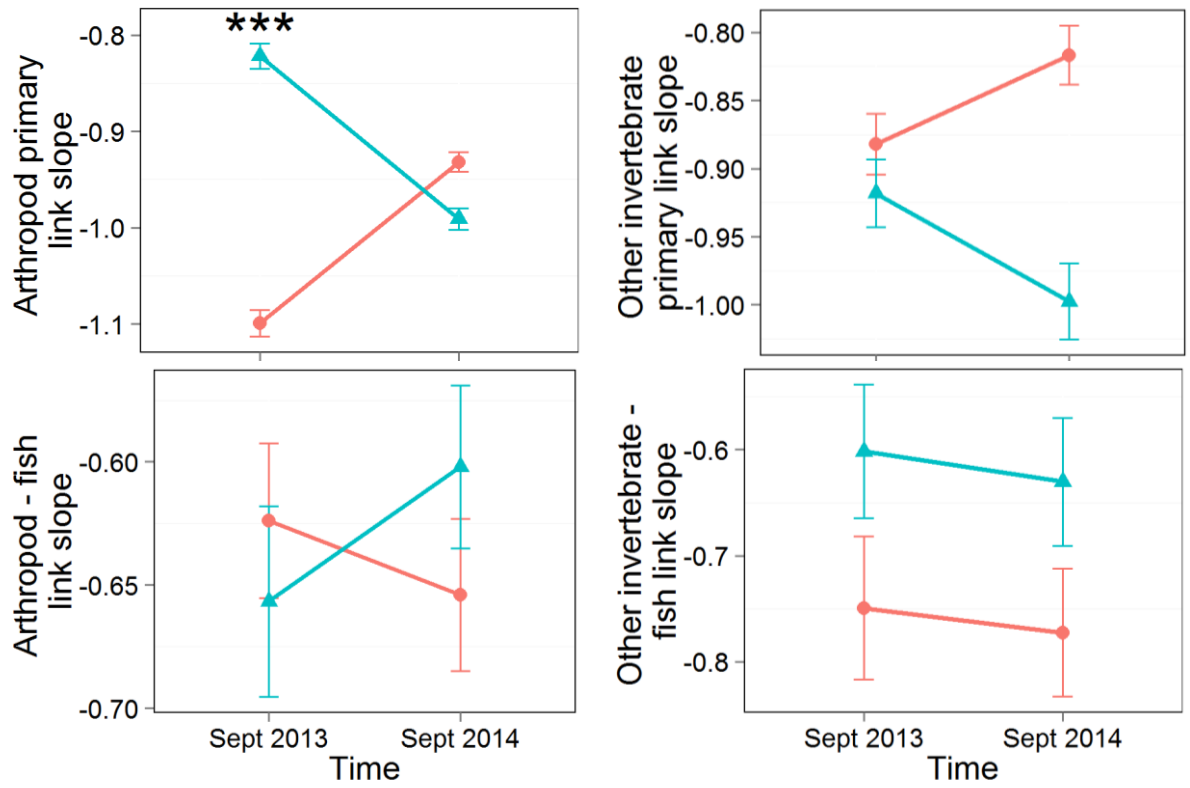



**Figure 27**. Link angles between invertebrates (arthropods and non-arthropods) and their algal resources (primary links) and fish consumers, over time in impacted (red, circles) and control sites (blue, triangles). Points show mean values, ± s.e.m. Those time points where impact and control slopes are different are indicated by significance stars; * = p = 0.05 to 0.01, ** = p = 0.01 to 0.001, *** = p <= 0.001 = ***.

### 5.4.3 Ecosystem processing

Total decomposition of leaf litter was lower in impacted sites up to 6 months after the pesticide spill (top Figure 28; Table 8). This was driven by the decline in invertebrate mediated decomposition (middle Figure 28; Table 8). No effect of the pesticide was seen on the rate of microbial decomposition (bottom Figure 28; Table 8), despite a trend towards higher rates immediately after the spill. These results support our fifth and sixth hypotheses; here the rate of microbial decomposition was found to be unaffected by the pesticide (in contrast to Thompson *et al.*), while invertebrate mediated decomposition took up to a year to recover. The biomass of *G. pulex*, a keystone detritivore, took longer to recover after the pesticide spill than the rate of invertebrate leaf litter decomposition.

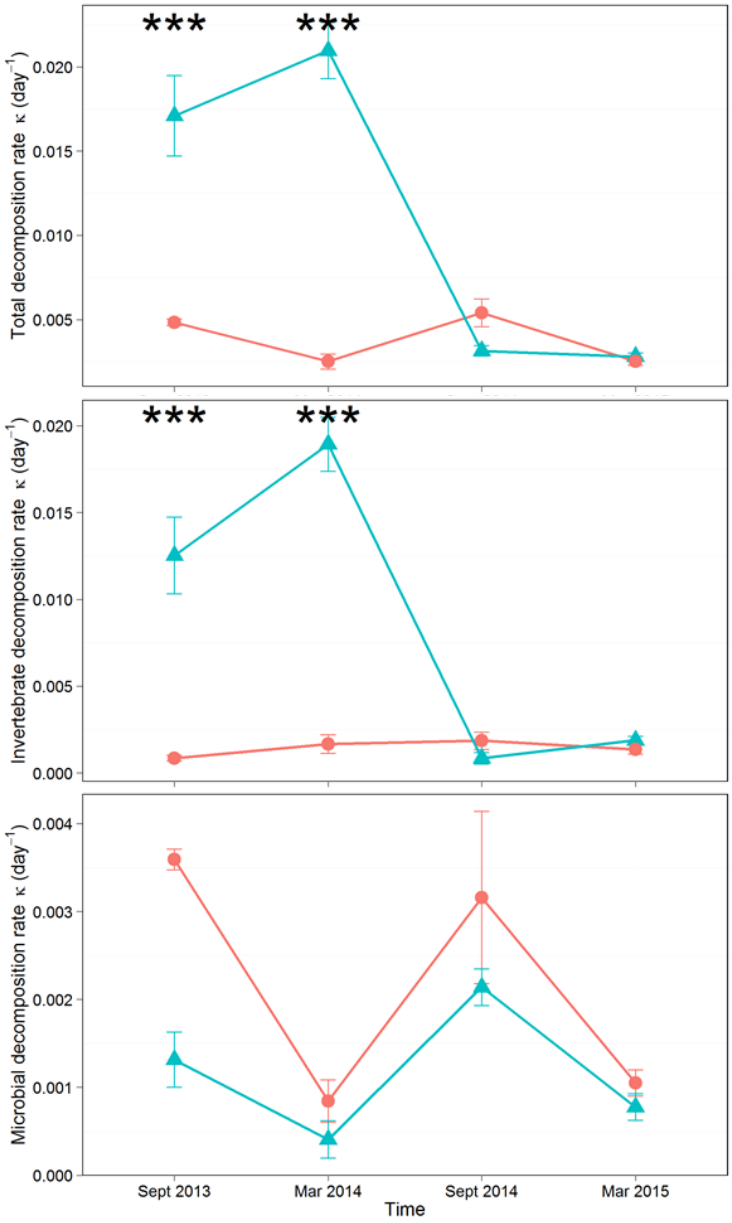



**Figure 28.** The decomposition rate expressed as temperature corrected rate of decomposition, over time in impacted (red, circles) and control sites (blue, triangles). Points show mean values, ± s.e.m. Those time points where impact and control means are different are indicated by significance stars; * = p = 0.05 to 0.01, ** = p = 0.01 to 0.001, *** = p <= 0.001 = ***.

## 5.5 Discussion

The spill of the insecticide chlorpyrifos on the River Kennet in 2013 affected almost all aspects of the ecological community. The loss of sensitive arthropod species caused significant reductions in the biomass of key invertebrate groups. These direct effects propagated through the food web, causing non-pesticide sensitive taxa to increase in biomass (increased biomass of Oligochaeta at impacted sites), this could be because those taxa were released from competitive pressure. Algae increased their biomass (increased chlorophyll-α concentration at impacted sites, larger diatoms in impacted sites), this could be because they were released from predation pressure from the pesticide sensitive taxa. Mass-abundance scaling between arthropods and their resources and consumers was altered such that there was a potential reduction in the flow of biomass through arthropod mediated feeding pathways (steeper feeding link slopes). There were knock on effects for ecosystem processing as invertebrate mediated decomposition rates were reduced. In general recovery time was faster for ecosystem processing (decomposition) than for structural aspects of the community, the exception being *Baetis,* whose biomass recovered by March 2014, before decomposition rates had. The CCA showed clear recovery of community structure over time.

The response and trajectory of recovery of Oligochaeta and *G. pulex* is interesting as it suggests that the increased biomass of Oligochaeta may have been preventing the full recovery of *G. pulex*, perhaps through competitive exclusion. A mechanism like this has been suggested to be responsible for the slow biological recovery of freshwater food webs affected by acidification (Layer, Hildrew & Woodward 2013), whereby under acidified conditions generalist herbivore-detritivores dominate and prevent the recolonization of specialist herbivores as pH increases. *G. pulex* is a highly generalist collector-gatherer species, feeding on both detritus and the algal biofilm, Oligochaeta is a broad taxonomic unit containing many taxa which also feed on both detritus and the algal biofilm, hence there is likely a high

overlap in the resources of these two nodes (Tachet *et al.* 2002). Alternatively *G. pulex* may have been slow to recover due to dispersal limitations, taxa such as *Baetis sp.* which have a winged adult stage which would have aided recovery, and indeed insect species did recover relatively quickly. *G. pulex* is fully aquatic as so would have likely recolonised from above and below the affected stretch of the river. If dispersal limitations were the cause of the slow recovery of *G. pulex*, then there would likely be a difference in the recovery rates of this species at the impacted sites, with the most upstream site, which is closest to a source population unaffected by the pesticide, recovering quickest. This trend was not apparent in the data but may emerge as more samples are processed from the remaining time points.

No effect of the pesticide was seen on the fish assemblage. This is unsurprising that there was no discernible effect on the fish biomass given how mobile they are and that their diet is known to be subsidised along the stretches of the river that commercial angling is found, which includes all of the impacted sites used here.

Previous work has revealed that these effects extend into the microbial world, Thompson *et al.* (2015), found that the functional potential of the microbial assemblage was higher in the impacted sites in September 2013, as was the abundance of genes associated with organophosphate use and ammonia oxidation which would likely be a response to the widespread arthropod death. Microbes account for most of the world's biodiversity, they drive key ecosystem processes and biogeochemical cycles (e.g. nitrogen cycle) and interact with higher trophic levels. It is likely therefor that the direct effects on the microbial assemblage (through pesticide molecules available to metabolise), as well as indirect effects (through the glut of arthropod carcasses available for decomposition) in turn caused knock on effects for those species interacting with the microbial assemblage (such as the meiofauna). Thompson *et al.* (2015) found that microbe mediated leaf litter decomposition was higher at impacted sites immediately after the pesticide spill, and indeed this trend was seen here too, although not statistically significant. This suggests a compensatory mechanism such

that microbial decomposers were able to increase their decomposition rates, perhaps through increased biomass due to the glut of additional resources (in the form of arthropod carcasses) which partially compensated for the reduced invertebrate decomposition.

Here we have demonstrated the use of food webs for better understanding the effects of perturbations as they propagate through the ecological community. Food webs have also been used to assess the effects of other stressors, such as acidification and eutrophication, where interactions within food webs can shape both the ecosystem impact and the rate and trajectory of recovery (Ledger & Hildrew 2005; Layer *et al.* 2010b; Rawcliffe *et al.* 2010). Thus, such an approach allows us to move beyond partial taxonomic or trait-based views to bioassessment, to one that explicitly incorporates species interactions in ecological networks and the ecosystem processes that result.

This study demonstrates the resilience of freshwater communities to perturbations, whilst the structure and functioning of the food web was altered by the pesticide spill, there were many alternative pathways through which biomass could flow allowing the overall network structure to be resistant to change and dampening impacts on the top predators facilitating a short recovery time. As more samples from each time point are processed we will gain a better understanding of the resilience and resistance of the different components of the community, as well as their trajectory of recovery over time.

There were instances in this data where impacted variable means became indistinguishable from control variable means, not through change over time in the impacted data, but due to change over time of control data, or change in both (i.e. in *Baetis sp.* and Oligochaeta biomass). Whilst counterintuitive, this can still be taken as evidence of recovery but highlights the need to better understand the baseline variability from season to season and year to year in these data. As more samples from each time point are processed a better understanding of this variability and more

robust trends will emerge.

Studies of the effects of pesticides at the ecosystem level are rare in natural settings (Köhler & Triebskorn 2013); this study contributes to filling this knowledge gap. The projected increase in the worldwide use of pesticides (Tilman *et al.* 2002) has the potential to cause substantial negative impacts to our waterways. For instance, Malaj *et al.* (2014) estimate there to be acute lethal effects of organic chemicals in 14%, and chronic long-term effects in 42% of European waterways. A deeper, more holistic understanding of the effects of pesticides which enter our waterways, as well as an understanding of how those effects propagate through the food web will facilitate more effective mitigation and restoration efforts.

# 6 | General Discussion

The main aim of this project was to examine the suitability of food webs as a tool for monitoring the impacts of anthropogenic stressors on the environment. The evidence is clear that in some situations consideration of the structure of ecological networks is vital to fully understanding the response of an ecosystem to environmental change (e.g. Scheffer & Carpenter 2003; Tylianakis, Tscharntke & Lewis 2007; Henson, Craze & Memmott 2009), but these examples used detailed food webs with information about the weight of the nodes and links, information not usually available from biomonitoring schemes. Here I have demonstrated that food webs built from routine biomonitoring data, (i.e. species lists) and trophic information harvested from the literature can provide a deeper understanding of ecological communities than species lists alone (Chapters 3 & 4). Chapter 3 demonstrates that freshwater food webs built from routine biomonitoring data can reveal insights into how the structure of those food webs is affected by acidification, with implications for the ability of those food webs to recover from acidification. The results presented in Chapter 4 demonstrate that classifying carabid consumer species in the context of the food web in which they appear, and considering the structure of the resultant food web, leads to a more powerful prediction about the level of weed seed regulation. However I have also found instances where these coarse, binary food webs are not sensitive enough to be a useful biomonitoring tool; the freshwater food webs built in Chapter 3 were not sensitive enough to reveal consistent changes in community structure at each site over time, although we know from the species data that significant changes in the species assemblages, especially at the lower trophic levels, are apparent at many sites (Murphy *et al.* 2014). However when these food webs are augmented with information about the predicted biomass flows between nodes they reveal new insights into the apparent lag in biological recovery after the amelioration of acidification (Appendix A).

The food webs in Chapter 5 were built using mass and abundance data for each of the nodes. This additional information provided a deeper understanding of the effects of the pesticide on the substructure of the networks, the efficiency of energy transfer across the trophic levels, and the effects these changes had on ecosystem processing. Collecting data of this sort as part of routine biomonitoring would necessitate a greater workload, but this could be offset by the added value the data brings in terms of a deeper understanding of the dynamics of the community in question (Chapter 1; Gray et al. 2014).

The key to fully realising the potential for ecological networks in biomonitoring science is to build networks which reflect the underlying changes in community dynamics which respond to environmental change. Chapter 2 (Gray *et al.* 2015b) clearly demonstrates that as the quality and quantity of the collection of trophic interactions increases, so too does the quality of the constructed food web. Hosting this dataset on an open access website (https://sites.google.com/site/foodwebsdatabase/), with an established mechanism for researchers to donate data should allow the quality and breadth of it to grow over time.

An important step toward improving the quality of interaction datasets could be to assess the number of times a particular interaction appears in a particular dataset, as well as the number of times an interaction could have occurred but did not (i.e. species found at the same site but not found to interact). If a particular interaction has been observed many times across many systems, it is probably reasonable to assume it also occurs at other sites where those species co-exist. However, if it has only been observed rarely, or at a site with very different characteristics than the one in question (for instance contrasting environmental conditions, or significantly different community assemblages) this assumption might not be so reasonable. Alongside this, the functional response of consumers to the abundance of their resources (Holling 1966), as well as the dependency of interaction strength on the

abundance of the consumer could be incorporated. As the volume of trophic information data continues to grow, the evaluation of the realism of predicted links will improve over time.

With each passing year, methodological advancements are increasing the ease with which interactions between species can be characterised. A great number of published papers now use these methods to construct ecological networks or characterise interactions (e.g. Harper *et al.* 2005; Foltan *et al.* 2005; Navarro *et al.* 2010; Wilson *et al.* 2010; Clare *et al.* 2011; Newmaster *et al.* 2013). These advances also mean that biomonitoring data can be collected more quickly and cheaply (Gibson *et al.* 2015).

Novel approaches to determining network structure are being developed all the time, for instance a machine leaning algorithm based on basic prior knowledge and some logical rules has been used to construct an agricultural invertebrate food web (Bohan *et al.* 2011b). With some alterations to the basic structure, it is possible that this algorithm could be applied to other suitable datasets. Bayesian approaches to predicting network structure has been used outside of ecology for many years (e.g. Heckerman, Geiger & Chickering 1995), researchers are now beginning to apply these methods to ecological datasets (e.g. de Sassi, Staniczenko & Tylianakis 2012). However, as is common in ecology, the widespread adoption of these new techniques is hindered by the quantity of suitable data available to develop these methods on. For instance, the method of Sassi *et al.* (2012) could be adapted to predict the strength of interactions in freshwater systems, but a detailed dataset of the biomass flow between freshwater consumers and resources, complete with body size and abundance information, is rare.

Long term monitoring data could be very valuable in furthering our understanding of the realisation of and variability in strength of interactions between species. Just as species abundances can vary wildly from one year to the next, so too will the interactions between those species. Long term monitoring could help us to capture some of that variability, as well as understand the influences on that variation.

Long-term studies are rare, other than the Upland Waters Monitoring Network (UWMN; Kernan *et al.* 2010), classic examples include the work of Likens *et al.* (1977) at the Hubbard Brook Experimental Forest, and Slavik *et al.* (2004) at the Kuparuk River station of the Long-Term Ecological Research (LTER) network. Within the UK there is also the Environmental Change Network which has been running for more than 23 years. The UWMN recently changed its protocols such that biological samples are now stored in such a way that future molecular work, and in particular molecular identification of gut contents, is possible. The value of these datasets will increase progressively over time, as it is only with multiple decades of data that more subtle long-term trends, such as responses to climate change, are able to be detected. Despite the obvious value of long term monitoring data, it is often difficult to extract research funding for such strategic research, which often appears to fail to meet the "novelty" criteria of many research councils' remits (Appendix D; Gray *et al.* 2015a).

Large datasets are still relatively rare in ecological research, although this is beginning to change as more 'Big Data' approaches become available. Once ecological advances were limited by the labour-intensive methods by which empirical data was collected, and the resultant paucity of good quality datasets available for research. Now however, due to new technologies and advanced computing power the challenge for many researchers has shifted towards the ability to process the vast quantities of data that are being produced, and to interpret their ecological significance (Woodward *et al.* 2014). Furthermore, with large ecological datasets come additional challenges, such as sampling consistency or meta-data collection and curation (Raffaelli *et al.* 2014). In particular, given a sufficiently large sample, a statistical test will almost always exhibit a significant difference, unless the effect size is exactly zero. Very small differences between samples, even if significant, may not be meaningful (Sullivan & Feinn 2012). Trends found in observation data may be slight, and accompanied by large variation, such as those found in Chapter 4, requiring a researcher to consider the effect size of the trend in question, and it's

biological significance. Unlike significance tests, effect size is independent of and therefore not confounded by, sample size (Sullivan & Feinn 2012). One explanation for the small effect sizes seen in Chapter 4, is that 80% of the variation in this seed regulation data has already been ascribed to changes in the cropping regimes in the preceding years (Bohan *et al.* 2011c), although this is also likely to have an effect on the carabid assemblage and therefore the counts of herbivores at each site. As discussed in Chapter 4, often observational data is often best suited for the identification of trends in order to formulate formal hypotheses and design the accompanying experiments.

Chapter 5 provides the opposite problem, that of small sample sizes. Here there was no true replication as all the experimental sites were along one river, and thus external forces affecting the whole river could not be controlled. Recovery of the impacted sites couldn't be defined as change in the desired variable over time, but rather that the difference between control and impacted sites should eventually decrease to zero. This is related to a common problem in restoration ecology, that of 'shifting baselines' (Pauly 1995), long term reference conditions may themselves be changing due to anthropogenic effects such as climate change (such as within the UWMN). In the short term, impacts such as other pollution incidents upstream of the control sites or changing weather will have affected the algal biomass or rate of decomposition along the river. Thus although decomposition rates at the impacted sites did not change over time (Figure 28), they can be considered recovered from the pesticide spill because the decomposition rates at the control sites did change over time (perhaps due to colder weather during September 2014 and March 2015), and became indistinguishable from the impacted sites. This perhaps counterintuitive result can be understood and interpreted correctly because multiple control and impacted sites were studied repeatedly over time. The 'one reach at a time' (Bernhardt *et al.* 2007) approach to river restoration, where restorations tend to be ad-hoc and a combination of techniques unique to each site with little control or replication (Friberg

*et al.* 2011), inevitably suffers from this problem of small sample sizes, and leads to weak hypothesis testing and reduced predictive power.

In addition to integrating ecological networks into biomonitoring approaches, incorporating a fully ecological and evolutionary perspective could also bring much added value (Appendix D; Gray *et al.* 2015a). Currently, the predictive capacity of traditional biomonitoring approaches is restricted and will have a limited ability to adapt in the face of rapid and global habitat modification and climate change. This approach alongside the incorporation of measures of ecosystem functioning and aided by new technologies such as novel molecular techniques, may facilitate the future development of a more comprehensive and effective biomonitoring framework.

# 7 | References

Aizen, M.A., Sabatino, M. & Tylianakis, J.M. (2012) Specialization and Rarity Predict Nonrandom Loss of Interactions from Mutualist Networks. *Science*, **335**, 1486–1489.

Allesina, S. (2011) Predicting trophic relations in ecological networks: A test of the Allometric Diet Breadth Model. *Journal of Theoretical Biology*, **279**, 161–168.

Allesina, S. & Pascual, M. (2009) Food web models: A plea for groups. *Ecology Letters*, **12**, 652–662.

Allouche, O., Tsoar, A. & Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223–1232.

Altenburger, R., Backhaus, T., Boedeker, W., Faust, M. & Scholze, M. (2013) Simplifying complexity: Mixture toxicity assessment in the last 20 years. *Environmental Toxicology and Chemistry*, **32**, 1685–1687.

Arseneau, K.M., Driscoll, C.T., Brager, L.M., Ross, K.A. & Cumming, B.F. (2011) Recent evidence of biological recovery from acidification in the Adirondacks (New York, USA): a multiproxy paleolimnological investigation of Big Moose Lake. *Canadian Journal of Fisheries and Aquatic Sciences*, **68**, 575–592.

Baird, D.J., Brock, T.C.M., de Ruiter, P.C., Boxall, A.B.A., Culp, J.M., Eldridge, P., Hammen, U., Jak, R.G., Kidd, K.A. & DeWitt, T. (2001) The food web approach in the environmental management of toxic substances. *Ecological Variability: Separating Anthropogenic from Natural Causes of Ecosystem Impairment* (eds D.J. Baird), & G.A. Burton), pp. 83–122. SETAC Press, Pensacola, USA.

Balloch, D., Davies, C.E. & Jones., F.H. (1976) Biological assessment of water quality in three British rivers: the North Esk (Scotland), the Ivel (England) and the Taf (Wales). *Water Pollution Control (UK)*.

Barnes, C. (2008) Predator and prey body sizes in marine food webs. *Ecology*, **89**, 881.

Barron, M.G. & Woodburn, K.B. (1995) Ecotoxicology of chlorpyrifos. *Reviews of Environmental Contamination and Toxicology*, **144**, 1–93.

Bascompte, J., Jordano, P., Melian, C.J. & Olesen, J.M. (2003) The nested assembly of plant-animal mutualistic networks. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 9383–9387.

Battarbee, R.W. (1973) A new method for the estimation of absolute microfossil numbers, with reference especially to diatoms. *Limnology and Oceanography*, **18**, 647–653.

Battarbee, R.W., Jones, V.J., Flower, R.J., Cameron, N.G., Bennion, H., Carvalho, L. & Juggins, S. (2001) Diatoms. *Tracking Environmental Change Using Lake Sediments. Volume 3: Terrestrial, Algal and Siliceous Indicators* pp. 151–202. Kluwer Academic

Publishers, Dordrecht, The Netherlands.

Battarbee, R.W., Monteith, D.T., Juggins, S., Evans, C.D., Jenkins, A. & Simpson, G.L. (2005) Reconstructing pre-acidification pH for an acidified Scottish loch: A comparison of palaeolimnological and modelling approaches. *Environmental Pollution*, **137**, 135–149.

Benedick, S., Hill, J.K., Mustaffa, N., Chey, V.K., Maryati, M., Searle, J.B., Schilthuizen, M. & Hamer, K.C. (2006) Impacts of rain forest fragmentation on butterflies in northern Borneo: Species richness, turnover and the value of small fragments. *Journal of Applied Ecology*, **43**, 967–977.

Bennion, H., Simpson, G.L., Anderson, N.J., Clarke, G., Dong, X.H., Hobaek, A., Guilizzoni, P., Marchetto, A., Sayer, C.D., Thies, H. & Tolotti, M. (2011) Defining ecological and chemical reference conditions and restoration targets for nine European lakes. *Journal of Paleolimnology*, **45**, 415–431.

Berlow, E.L., Dunne, J.A., Martinez, N.D., Stark, P.B., Williams, R.J. & Brose, U. (2009) Simple prediction of interaction strengths in complex food webs. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 187–191.

Berlow, E.L., Neutel, A.-M., Cohen, J.E., De Ruiter, P.C., Ebenman, B., Emmerson, M., Fox, J.W., Jansen, V.A.A., Jones, I.J., Kokkoris, G.D., Logofet, D.O., McKane, A.J., Montoya, J.M. & Petchey, O. (2004) Interaction strengths in food webs: issues and opportunities. *Journal of Animal Ecology*, **73**, 585–598.

Bernhardt, E.S., Sudduth, E.B., Palmer, M.A., Allan, J.D., Meyer, J.L., Alexander, G., Follastad-Shah, J., Hassett, B., Jenkinson, R., Lave, R., Rumps, J. & Pagano, L. (2007) Restoring rivers one reach at a time. *Restoration Ecology*, **15**, 482–493.

Birk, S., Willby, N.J., Kelly, M.G., Bonne, W., Borja, A., Poikane, S. & van de Bund, W. (2013) Intercalibrating classifications of ecological status: Europe's quest for common management objectives for aquatic ecosystems. *Science of the Total Environment*, **454-455**, 490–499.

Bohan, D.A., Boffey, C.W.H., Brooks, D.R., Clark, S.J., Dewar, A.M., Firbank, L.G., Haughton, A.J., Hawes, C., Heard, M.S., May, M.J., Osborne, J.L., Perry, J.N., Rothery, P., Roy, D.B., Scott, R.J., Squire, G.R., Woiwod, I.P. & Champion, G.T. (2005) Effects on weed and invertebrate abundance and diversity of herbicide management in genetically modified herbicide-tolerant winter-sown oilseed rape. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 463–474.

Bohan, D.A., Bohan, A.C., Glen, D.M., Symondson, W.O.C., Wiltshire, C.W. & Hughes, L. (2000) Spatial dynamics of predation by carabid beetles on slugs. *Journal of Animal Ecology*, **69**, 367–379.

Bohan, D.A., Boursault, A., Brooks, D.R. & Petit, S. (2011a) National-scale regulation of the weed seedbank by carabid predators. *Journal of Applied Ecology*, **48**, 888–898.

Bohan, D.A., Caron-Lormier, G., Muggleton, S., Raybould, A. & Tamaddoni-Nezhad, A. (2011b) Automated Discovery of Food Webs from Ecological Data Using Logic-Based Machine Learning. *Plos One*, **6**.

Bohan, D.A., Powers, S.J., Champion, G., Haughton, A.J., Hawes, C., Squire, G.,

Cussans, J. & Mertens, S.K. (2011c) Modelling rotations: Can crop sequences explain arable weed seedbank abundance? *Weed Research*, **51**, 422–432.

Bonada, N., Prat, N., Resh, V.H. & Statzner, B. (2006) Developments in aquatic insect biomonitoring: a comparative analysis of recent approaches. *Annual Review of Entomology*, **51**, 495–523.

Bray, R.J. & Curtis, J.T. (1957) An ordination of the upland forest communities of southern Winsconin. *Ecological Monographs*, **27**, 325–349.

Briand, F. (1983) Environmental Control of Food Web Structure. *Ecology*, **64**, 253–263.

Briand, F. & Cohen, J.E. (1984) Community food webs have scale-invariant structure. *Nature*, **307**, 264–267.

van den Brink, P.J., van Wijngaarden, R.P.A., Lucassen, W.G.H., Brock, T.C.M. & Leeuwangh, P. (1996) Effects of the insecticide Dursban 4e (active ingredient chlorpyrifos) in outdoor experimental ditches: II. Invertebrate community responses and recovery. *Environmental Toxicology and Chemistry*, **15**, 1143–1153.

Brock, T.C.M., Vet, J.J.R.M., Kerkhofs, M.J.J., Lijzen, J., Zuilekom, W.J. Van & Gijlstra, R. (1993) Fate and Effects of the Insecticide Dursban® 4E in Indoor Elodea-Dominated and Macrophyte-Free Freshwater Model Ecosystems : III. Aspects of Ecosystem Functioning. *Archives of Environmental Contamination and Toxicology*, **169**, 160–169.

Brock, T.C.M., van Wijngaarden, R.P.A. & van Geest, G.J. (2000) Ecological risks of pesticides in freshwater ecosystems Part 2 : Insecticides. *Alterra-rapport 089*, 128.

Brooks, D.R., Bohan, D.A., Champion, G.T., Haughton, A.J., Hawes, C., Heard, M.S., Clark, S.J., Dewar, A.M., Firbank, L.G., Perry, J.N., Rothery, P., Scott, R.J., Woiwod, I.P., Birchall, C., Skellern, M.P., H.Walker, J., Baker, P., Bell, D., Browne, E.L., Dewar, A.J.G., Fairfax, C.M., Garner, B.H., Haylock, L.A., Horne, S.L., Hulmes, S.E., Mason, N.S., Norton, L.R., Nuttall, P., Randle, Z., Rossall, M.J., Sands, R.J.N., Singer, E.J. & Walker, M.J. (2003) Invertebrate responses to the management of genetically modified herbicide-tolerant and conventional spring crops. I. Soil-surface-active invertebrates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **358**, 1847–1862.

Brose, U., Cushing, L., Berlow, E.L., Jonsson, T., Banašek-Richter, C., Bersier, L.-F., Blanchard, J.L., Brey, T., Carpenter, S.R., Blandenier, M.-F.C., Cohen, J.E., Dawah, H.A., Dell, T., Edwards, F., Harper-Smith, S., Jacob, U., Knapp, R.A., Ledger, M.E., Memmott, J., Mintenbeck, K., Pinnegar, J.K., Rall, B.C., Rayner, T., Ruess, L., Ulrich, W., Warren, P., Williams, R.J., Woodward, G., Yodzis, P. & Martinez, N.D. (2005) Body Sizes of Consumers and Their Resources. *Ecology*, **86**, 2545–2545.

Butts, C. (2013) sna: Tools for Social Network Analysis. R package version 2.3-2. http://CRAN.R-project.org/package=sna.

Carle, F.L. & Strub, M.R. (1978) A new method for estimating population size from removal data. *Biometrics*, **34**, 621–630.

Champion, G.T., May, M.J., Bennett, S., Brooks, D.R., Clark, S.J., Daniels, R.E., Firbank, L.G., Haughton, A.J., Hawes, C., Heard, M.S., Perry, J.N., Randle, Z., Rossall, M.J.,

Rothery, P., Skellern, M.P., Scott, R.J., Squire, G.R. & Thomas, M.R. (2003) Crop management and agronomic context of the Farm Scale Evaluations of genetically modified herbicide-tolerant crops. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **358**, 1801–1818.

Chung, K., Wallace, J. & Grubaugh, J. (1993) The impact of insecticide treatment on abundance, biomass and production of litterbag fauna in a headwater stream: a study of pretreatment, treatment and recovery. *Limnologica*, **28**, 93–106.

Clare, E.L. (2014) Molecular detection of trophic interactions: emerging trends, distinct advantages, significant considerations and conservation applications. *Evolutionary Applications*, **7**, 1144–1157.

Clare, E.L., Barber, B.R., Sweeney, B.W., Hebert, P.D.N. & Fenton, M.B. (2011) Eating local: Influences of habitat on the diet of little brown bats (*Myotis lucifugus*). *Molecular Ecology*, **20**, 1772–1780.

Clark, J.M., Van Der Heijden, G.M.F., Palmer, S.M., Chapman, P.J. & Bottrell, S.H. (2011) Variation in the sensitivity of DOC release between different organic soils following $H_2SO_4$ and sea-salt additions. *European Journal of Soil Science*, **62**, 267–284.

Clark, C.M. & Tilman, D. (2008) Loss of plant species after chronic low-level nitrogen deposition to prairie grasslands. *Nature*, **451**, 712–5.

Cleveland, W., Grosse, E., Shyu, W., Chambers, J. & Hastie, T. (1992) Local regression models. *Statistical Models in S* (eds J.M. Chambers), & T.J. Hastie), pp. 309–376. Wadsworth & Brooks/Cole.

Clough, Y., Holzschuh, A., Gabriel, D., Purtauf, T., Kleijn, D., Kruess, A., Steffan-Dewenter, I. & Tscharntke, T. (2007) Alpha and beta diversity of arthropods and plants in organically and conventionally managed wheat fields. *Journal of Applied Ecology*, **44**, 804–812.

Cohen, J.E., Jonsson, T. & Carpenter, S.R. (2003) Ecological community description using the food web, species abundance, and body size. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 1781–1786.

Cohen, J.E. & Mulder, C. (2014) Soil invertebrates, chemistry, weather, human management, and edaphic food webs at 135 sites in The Netherlands: SIZEWEB. *Ecology*, **95**, 578.

Cohen, J.E., Newman, C.M. & Briand, F. (1985) A Stochastic Theory of Community Food Webs: II. Individual Webs. *Proceedings of the Royal Society B: Biological Sciences*, **224**, 449–461.

Cohen, J.E., Schittler, D.N., Raffaelli, D.G. & Reuman, D.C. (2009) Food webs are more than the sum of their tritrophic parts. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 22335–40.

Cuppen, J.G.M., Gylstra, R., Sv, B., Budde, B.J. & Brock, T.C.M. (1995) Effects of nutrient loading and insecticide application on the ecology of *Elodea* -dominated freshwater microcosms: responses of macroinvertebrate detritivores, breakdown of plant litter, and final conclusions. *Archiv für Hydrobiologie*, **134**, 157–177.

Dormann, C.F., Gruber, B. & Fruend, J. (2008) Introducing the bipartite package: Analysing ecological networks. *R news*, **8**, 8–11.

Driscoll, C.T., Lawrence, G.B., Bulger, A.J., Butler, T.J., Cronan, C.S., Eagar, C., Lambert, K.F., Likens, G.E., Stoddard, J.L. & Weathers, K.C. (2001) Acidic Deposition in the Northeastern United States: Sources and Inputs, Ecosystem Effects, and Management Strategies. *BioScience*, **51**, 180.

Dunne, J.A. (2006) The network structure of food webs. *Ecological Networks: Linking Structure to Dynamics in Food Webs* pp. 27–86.

Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002a) 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*, **99**, 12917–12922.

Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002b) Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, **5**, 558–567.

Edwards, F.K., Lauridsen, R.B., Fernandes, W.P.A., Beaumont, W.R.C., Ibbotson, A.T., Scott, L., Davies, C.E. & Jones, J.I. (2009) Re-introduction of Atlantic salmon, *Salmo salar* L., to the Tadnoll Brook, Dorset. *Proceedings of the Dorset Natural History and Archaeological Society*, **130**, 9–16.

Eklof, A., Helmus, M.R., Moore, M. & Allesina, S. (2012) Relevance of evolutionary history for food web structure. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 1588–1596.

Elton, C.S. (1927) *Animal Ecology*. Sidgwick & Jackson, London, UK.

Eskelson, M.J., Chapman, E.G., Archbold, D.D., Obrycki, J.J. & Harwood, J.D. (2011) Molecular identification of predation by carabid beetles on exotic and native slugs in a strawberry agroecosystem. *Biological Control*, **56**, 245–253.

European Commission. (2000) Establishing a framework for community action in the field of water policy. Directive 2000/60/EC of the European Parliament and of the Council. , L 327, 1–72.

Evans, C.D., Cullen, J.M., Alewell, C., Kopácek, J., Marchetto, A., Moldan, F., Prechtel, A., Rogora, M., Veselý, J. & Wright, R. (2001) Recovery from acidification in European surface waters. *Hydrology and Earth System Sciences*, **5**, 283–298.

Evans, C.D., Jones, T.G., Burden, A., Ostle, N., Zieliński, P., Cooper, M.D. a, Peacock, M., Clark, J.M., Oulehle, F., Cooper, D. & Freeman, C. (2012) Acidity controls on dissolved organic carbon mobility in organic soils. *Global Change Biology*, **18**, 3317–3331.

Evans, C.D., Monteith, D.T. & Harriman, R. (2001) Long-term variability in the deposition of marine ions at west coast sites in the UK Acid Waters Monitoring Network: Impacts on surface water chemistry and significance for trend determination. *Science of the Total Environment*, **265**, 115–129.

Evans, C.D., Monteith, D.T., Reynolds, B. & Clark, J.M. (2008) Buffering of recovery from acidification by organic acids. *Science of the Total Environment*, **404**, 316–325.

Feld, C.K., Birk, S., Bradley, D.C., Hering, D., Kail, J., Marzin, A., Melcher, A., Nemitz,

D., Pedersen, M.L., Pletterbauer, F., Pont, D., Verdonschot, P.F.M. & Friberg, N. (2011) From Natural to Degraded Rivers and Back Again: A Test of Restoration Ecology Theory and Practice. *Advances in Ecological Research*, **44**, 119–209.

Firbank, L.G., Heard, M.S., Woiwod, I.P., Hawes, C., Haughton, A.J., Champion, G.T., Scott, R.J., Hill, M.O., Dewar, A.M., Squire, G.R., May, M.J., Brooks, D.R., Bohan, D.A., Daniels, R.E., Osborne, J.L., Roy, D.B., Black, H.I.J., Rothery, P. & Perry, J.N. (2003) An introduction to the farm-scale evaluations of genetically modified herbicide-tolerant crops. *Journal of Applied Ecology*, **40**, 2–16.

Foltan, P., Sheppard, S., Konvicka, M. & Symondson, W.O.C. (2005) The significance of facultative scavenging in generalist predator nutrition: detecting decayed prey in the guts of predators using PCR. *Molecular Ecology*, **14**, 4147–58.

Friberg, N., Bonada, N., Bradley, D.C., Dunbar, M.J., Edwards, F.K., Grey, J., Hayes, R.B., Hildrew, A.G., Lamouroux, N., Trimmer, M. & Woodward, G. (2011) Biomonitoring of Human Impacts in Freshwater Ecosystems: The Good, the Bad and the Ugly. *Advances in Ecological Research*, **44**, 1–68.

Gibson, J.F., Shokralla, S., Curry, C., Baird, D.J., Monk, W.A., King, I. & Hajibabaei, M. (2015) Large-Scale Biomonitoring of Remote and Threatened Ecosystems via High-Throughput Sequencing. *PLoS ONE*, **10**, e0138432.

Giddings, J.M., Williams, W.M., Solomon, K.R. & Giesy, J.P. (2014) Ecological Risk Assessment for Chlorpyrifos in Terrestrial and Aquatic Systems in the United States. *Reviews of Environmental Contamination and Toxicology*, **231**, 119–162.

Gilbert, A.J. (2009) Connectance indicates the robustness of food webs when subjected to species loss. *Ecological Indicators*, **9**, 72–80.

Gilljam, D., Thierry, A., Edwards, F.K., Figueroa, D., Ibbotson, A.T., Jones, J.I., Lauridsen, R.B., Petchey, O.L., Woodward, G. & Ebenman, B. (2011) Seeing Double: Size-Based and Taxonomic Views of Food Web Structure. *Advances in Ecological Research*, **45**, 67–133.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M. & Toulmin, C. (2010) Food Security: The Challenge of Feeding 9 Billion people. *Science*, **327**, 812–818.

Goldwasser, L. & Roughgarden, J. (1993) Construction and analysis of a large Caribbean food web. *Ecology*, **74**, 1216–1233.

Goldwasser, L. & Roughgarden, J. (1997) Sampling effects and the estimation of food-web properties. *Ecology*, **78**, 41–54.

Gravel, D., Poisot, T., Albouy, C., Velez, L. & Mouillot, D. (2013) Inferring food web structure from predator–prey body size relationships. *Methods in Ecology and Evolution*, **4**, 1083–1090.

Gray, C., Baird, D.J., Baumgartner, S., Jacob, U., Jenkins, G.B., O'Gorman, E.J., Lu, X., Ma, A., Pocock, M.J.O., Schuwirth, N., Thompson, M. & Woodward, G. (2014) Ecological networks: the missing links in biomonitoring science. *Journal of Applied Ecology*, **51**, 1444–1449.

Gray, C., Bista, I., Creer, S., Demars, B.O.L., Falciani, F., Monteith, D.T., Sun, X. &

Woodward, G. (2015a) Freshwater conservation and biomonitoring of structure and function: genes to ecosystems. *Aquatic Functional Biodiversity: Ecological and Evolutionary Approaches* (eds A. Belgrano, G. Woodward & U. Jacob), Elsevier.

Gray, C., Figueroa, D.H., Hudson, L.N., Ma, A., Perkins, D. & Woodward, G. (2015b) Joining the dots: an automated method for constructing food webs from compendia of published interactions. *Food Webs*, **5**, 11–20.

Hall, S.J. & Raffaelli, D. (1991) Food-Web Patterns: Lessons from a Species-Rich Web. *Journal of Animal Ecology*, **60**, 823–841.

Hall, S.J. & Raffaelli, D.G. (1997) Food web patterns: What do we really know? *Multitrophic Interactions* (eds A.C. Gange & V.K. Brown), pp. 395–417. Blackwells Scientific Publications, Oxford, UK.

Halstead, N.T., McMahon, T.A., Johnson, S.A., Raffel, T.R., Romansic, J.M., Crumrine, P.W. & Rohr, J.R. (2014) Community ecology theory predicts the effects of agrochemical mixtures on aquatic biodiversity and ecosystem properties. *Ecology Letters*, **17**, 932–941.

Harper, G.L., King, R. a., Dodd, C.S., Harwood, J.D., Glen, D.M., Bruford, M.W. & Symondson, W.O.C. (2005) Rapid screening of invertebrate predators for multiple prey DNA targets. *Molecular Ecology*, **14**, 819–827.

Havens, K.E. (1993) Predator-Prey Relationships in Natural Community Food Webs. *Oikos*, **68**, 117.

Heard, M.S., Hawes, C., Champion, G.T., Clark, S.J., Firbank, L.G., Haughton, a J., Parish, a M., Perry, J.N., Rothery, P., Roy, D.B., Scott, R.J., Skellern, M.P., Squire, G.R. & Hill, M.O. (2003) Weeds in fields with contrasting conventional and genetically modified herbicide-tolerant crops. II. Effects on individual species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **358**, 1833–1846.

Heckerman, D., Geiger, D. & Chickering, D. (1995) Learning Bayesian networks: the combination of knowledge and statistical data. *Machine Learning*, **20**, 197–243.

Hegland, S.J., Dunne, J., Nielsen, A. & Memmott, J. (2010) How to monitor ecological communities cost-efficiently: The example of plant-pollinator networks. *Biological Conservation*, **143**, 2092–2101.

Heleno, R., Devoto, M. & Pocock, M. (2012) Connectance of species interaction networks and conservation value: Is it any good to be well connected? *Ecological Indicators*, **14**, 7–10.

Henriksen, A., Fjeld, E. & Hesthagen, T. (1999) Critical load exceedance and damage to fish populations. *Ambio*, **28**, 583–586.

Henson, K.S.E., Craze, P.G. & Memmott, J. (2009) The Restoration of Parasites , Parasitoids , and Pathogens to Heathland Communities. *Ecology*, **90**, 1840–1851.

Hildrew, A.G. (2009) Sustained Research on Stream Communities: A Model System and The Comparative Approach. *Advances in Ecological Research*, **41**, 175–312.

Hillebrand, H., Dürselen, C.-D., Kirschtel, D., Pollingher, U. & Zohary, T. (1999) Biovolume calculation for pelagic and benthic microalgae. *Journal of Phycology*, **424**, 403–424.

Holling, C.S. (1966) The Functional Response of Invertebrate Predators to Prey Density. *Memoirs of the Entomological Society of Canada*, **98**, 1–86.

Hudson, L.N., Emerson, R., Jenkins, G.B., Layer, K., Ledger, M.E., Pichler, D.E., Thompson, M.S.A., O'Gorman, E.J., Woodward, G. & Reuman, D.C. (2013) Cheddar: analysis and visualisation of ecological communities in R. *Methods in Ecology and Evolution*, **4**, 99–104.

Huxham, M., Raffaelli, D. & Pike, A. (1995) Parasites and food web patterns. *Journal of Animal Ecology*, 168–176.

Ings, T.C., Montoya, J.M., Bascompte, J., Bluthgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E., Lewis, H.M., Olesen, J.M., van Veen, F.J.F., Warren, P.H. & Woodward, G. (2009) Ecological networks - beyond food webs. *Journal of Animal Ecology*, **78**, 253–269.

Jordán, F. (2009) Keystone species and food webs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 1733–1741.

Kaiser-Bunbury, C.N., Muff, S., Memmott, J., Muller, C.B., Caflisch, A., Müller, C.B. & Caflisch, A. (2010) The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behaviour. *Ecology Letters*, **13**, 442–452.

Kernan, M., Battarbee, R.W., Curtis, C., Monteith, D.T. & Shillands, E.M. (2010) *Recovery of Lakes and Streams in the UK from Acid Rain. The United Kingdom Acid Waters Monitoring Network 20 Year Interpretative Report*. Report to the Department for Environment, Food and Rural Affairs.

Köhler, H.-R. & Triebskorn, R. (2013) Wildlife ecotoxicology of pesticides: can we track effects to the population level and beyond? *Science*, **341**, 759–65.

Kramer, P.R.G., Jonkers, D.A., van Liere, L., Brock, T.C.M., Brink, P.J. van den & Hartgers, E.M. (1997) Impact of community nutrient status on fate and effects of insecticides in shallow freshwater ecosystems. *Interactions of nutrients and toxicants in the food chain of aquatic ecosystems* (eds P.R.G. Kramer, D.A. Jonkers & L. van Liere), National Institute of Public Health and the Environment, RIVM report 703715001, Bilthoven, The Netherlands.

Kremen, C. (2005) Managing ecosystem services: what do we need to know about their ecology? *Ecology Letters*, **8**, 468–479.

Kremen, C., Ullman, K.S. & Thorp, R.W. (2011) Evaluating the Quality of Citizen-Scientist Data on Pollinator Communities. *Conservation Biology*, **25**, 607–617.

Krzywinski, M., Birol, I., Jones, S.J. & Marra, M. a. (2012) Hive plots-rational approach to visualizing networks. *Briefings in Bioinformatics*, **13**, 627–644.

Larochelle, A. (1990) *The Food of Carabid Beetles (Coleoptera: Carabidae, Including Cicindelinae)*. Fabreries, Supplément 5.

Latora, V. & Marchiori, M. (2001) Efficient Behavior of Small-World Networks. *Physical Review Letters*, **87**, 198701.

Lawton, J.H. (1989) Food webs. *Ecological concepts* (ed J.M. Cherrett), pp. 43–78. Blackwell Scientific, Oxford, UK.

Layer, K., Hildrew, A.G., Jenkins, G.B., Riede, J.O., Rossiter, S.J., Townsend, C.R. & Woodward, G. (2011) Long-Term Dynamics of a Well-Characterised Food Web: Four Decades of Acidification and Recovery in the Broadstone Stream Model System (ed G Woodward). *Advances in Ecological Research*, **44**, 69–117.

Layer, K., Hildrew, A., Monteith, D. & Woodward, G. (2010a) Long-term variation in the littoral food web of an acidified mountain lake. *Global Change Biology*, **16**, 3133–3143.

Layer, K., Hildrew, A.G. & Woodward, G. (2013) Grazing and detritivory in 20 stream food webs across a broad pH gradient. *Oecologia*, **171**, 459–471.

Layer, K., Riede, J.O., Hildrew, A.G. & Woodward, G. (2010b) Food Web Structure and Stability in 20 Streams Across a Wide pH Gradient. *Advances in Ecological Research*, **42**, 265–299.

Ledger, M.E., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N. & Woodward, G. (2013) Extreme Climatic Events Alter Aquatic Food Webs. A Synthesis of Evidence from a Mesocosm Drought Experiment. *Advances in Ecological Research*, **48**, 343–395.

Ledger, M.E., Brown, L.E., Edwards, F.K., Milner, A.M. & Woodward, G. (2012) Drought alters the structure and functioning of complex food webs. *Nature Climate Change*, **3**, 223–227.

Ledger, M.E. & Hildrew, A.G. (2005) The ecology of acidification and recovery: changes in herbivore-algal food web linkages across a stream pH gradient. *Environmental Pollution*, **137**, 103–118.

Levine, S. (1980) Several measures of trophic structure applicable to complex food webs. *Journal of Theoretical Biology*, **83**, 195–207.

Likens, G.E., Bormann, F.H., Pierce, R.S., Eaton, J.S. & Johnson, N.M. (1977) *Biogeo-Chemistry of a Forested Ecosystem*. Springer-Verlag, New York.

Lorenzen, C. (1967) Determination of chlorophyll and pheopigments: spectrophotometric equations. *Limnology and Oceanography*, **12**, 343–346.

Lövei, G.L. & Sunderland, K.D. (1996) Ecology and Behavior of Ground Beetles (Coleoptera: Carabidae). *Annual Review of Entomology*, **41**, 231–256.

Lundberg, P., Ranta, E. & Kaitala, V. (2000) Species loss leads to community closure. *Ecology Letters*, **3**, 465–468.

Lundgren, J.G., Saska, P. & Honěk, A. (2013) Molecular approach to describing a seed-based food web: The post-dispersal granivore community of an invasive plant. *Ecology and Evolution*, **3**, 1642–1652.

Macfadyen, S., Craze, P.G., Polaszek, A., van Achterberg, K. & Memmott, J. (2011) Parasitoid diversity reduces the variability in pest control services across time on farms. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 3387–3394.

Malaj, E., von der Ohe, P.C., Grote, M., Kühne, R., Mondy, C.P., Usseglio-Polatera, P., Brack, W. & Schäfer, R.B. (2014) Organic chemicals jeopardize the health of freshwater ecosystems on the continental scale. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 9549–54.

Marcogliese, D.J. & Cone, D.K. (1997) Food webs: a plea for parasites. *Trends in Ecology & Evolution*, **12**, 320–325.

Martinez, N.D. (1991) Artifacts or attributes? Effects of resolution on the Little Rock Lake food web. *Ecological Monographs*, **61**, 367–392.

Masters, Z.O.E., Peteresen, I., Hildrew, A.G. & Ormerod, S.J. (2007) Insect dispersal does not limit the biological recovery of streams from acidification. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **17**, 375–383.

May, R.M. (1972) Will a large complex system be stable? *Nature*, **238**, 413–414.

McCann, K.S. (2000) The diversity-stability debate. *Nature*, **405**, 228–233.

McCann, K. (2007) Protecting biostructure. *Nature*, **446**, 29.

Memmott, J., Martinez, N.D. & Cohen, J.E. (2000) Predators, Parasitoids and Pathogens: Species Richness, Trophic Generality and Body Sizes in a Natural Food Web. *Journal of Animal Ecology*, **69**, 1–15.

Metcalfe-Smith, J.L. (1996) Biological water-quality assessment of rivers: Use of macroinvertebrate communities. *River Restoration* (eds G. Petts & P. Calow), pp. 17–59. Blackwell Science Ltd, Oxford, UK.

Monteith, D.T. & Evans, C.D. (1998) *United Kingdom Acid Waters Monitoring Network 10 Year Report. Analysis and Interpretation of Results, April 1988 - March 1998. Report to the Department for Environment, Food and Rural Affairs (Contract EPG 1/3/160)*.

Monteith, D.T. & Evans, C.D. (2005) The United Kingdom Acid Waters Monitoring Network: a review of the first 15 years and introduction to the special issue. *Environmental Pollution*, **137**, 3–13.

Monteith, D.T., Evans, C.D., Henrys, P.A., Simpson, G.L. & Malcolm, I.A. (2014) Trends in the hydrochemistry of acid-sensitive surface waters in the UK 1988–2008. *Ecological Indicators*, **37**, 287–303.

Monteith, D.T., Hildrew, A.G., Flower, R.J., Raven, P.J., Beaumont, W.R.B., Collen, P., Kreiser, A.M., Shilland, E.M. & Winterbottom, J.H. (2005) Biological responses to the chemical recovery of acidified fresh waters in the UK. *Environmental Pollution*, **137**, 83–101.

Monteith, D.T., Stoddard, J.L., Evans, C.D., de Wit, H.A., Forsius, M., Høgåsen, T., Wilander, A., Skjelkvåle, B.L., Jeffries, D.S., Vuorenmaa, J., Keller, B., Kopácek, J. & Vesely, J. (2007) Dissolved organic carbon trends resulting from changes in atmospheric deposition chemistry. *Nature*, **450**, 537–540.

Montoya, J.M., Pimm, S.L. & Solé, R. V. (2006) Ecological networks and their fragility. *Nature*, **442**, 259–264.

Montoya, D., Rogers, L. & Memmott, J. (2012) Emerging perspectives in the restoration of biodiversity-based ecosystem services. *Trends in Ecology and Evolution*, **27**, 666–672.

Montoya, J.M. & Solé, R. V. (2002) Small world patterns in food webs. *Journal of Theoretical Biology*, **214**, 405–412.

Morris, R.J., Gripenberg, S., Lewis, O.T. & Roslin, T. (2014) Antagonistic interaction networks are structured independently of latitude and host guild. *Ecology Letters*,

**17**, 340–9.

Mulder, C. & Elser, J.J. (2009) Soil acidity, ecological stoichiometry and allometric scaling in grassland food webs. *Global Change Biology*, **15**, 2730–2738.

Mundy, C.A., Allen-Williams, L.J., Underwood, N. & Warrington, S. (2000) Prey selection and foraging behaviour by *Pterostichus cupreus* L. (Col., Carabidae) under laboratory conditions. *Journal of Applied Entomology*, **124**, 349–358.

Murphy, J.F., Davy-Bowker, J., McFarland, B. & Ormerod, S.J. (2013) A diagnostic biotic index for assessing acidity in sensitive streams in Britain. *Ecological Indicators*, **24**, 562–572.

Murphy, J.F., Winterbottom, J.H., Orton, S., Simpson, G.L., Shilland, E.M. & Hildrew, A.G. (2014) Evidence of recovery from acidification in the macroinvertebrate assemblages of UK fresh waters: A 20-year time series. *Ecological Indicators*, **37**, 330–340.

Naeem, S. (1998) Species redundancy and ecosystem reliability. *Conservation Biology*, **12**, 39–45.

Navarro, S.P., Jurado-Rivera, J.A., Gómez-Zurita, J., Lyal, C.H.C. & Vogler, A.P. (2010) DNA profiling of host-herbivore interactions in tropical forests. *Ecological Entomology*, **35**, 18–32.

Newmaster, S.G., Thompson, I.D., Steeves, R.A.D., Rodgers, A.R., Fazekas, A.J., Maloles, J.R., McMullin, R.T. & Fryxell, J.M. (2013) Examination of two new technologies to assess the diet of woodland caribou: video recorders attached to collars and DNA barcoding. *Canadian Journal of Forest Research*, **43**, 897–900.

Novotny, V., Miller, S.E., Hulcr, J., Drew, R.A.I., Basset, Y., Janda, M., Setliff, G.P., Darrow, K., Stewart, A.J.A., Auga, J., Isua, B., Molem, K., Manumbor, M., Tamtiai, E., Mogia, M. & Weiblen, G.D. (2007) Low beta diversity of herbivorous insects in tropical forests. *Nature*, **448**, 692–5.

Ogle, D.H. (2012) FSA: Fisheries stock analysis. R package version 0.8.3.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2015) vegan: Community Ecology Package. R package version 2.3-1. http://CRAN.R-project.org/package=vegan.

Olesen, J.M., Bascompte, J., Dupont, Y.L. & Jordano, P. (2007) The modularity of pollination networks. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19891–19896.

Olito, C. & Fox, J.W. (2014) Species traits and abundances predict metrics of plant – pollinator network structure , but not pairwise interactions. *Oikos*, 1–9.

Patrick, S., Monteith, D.T. & Jenkins, A. (Eds. ). (1995) *UK Acid Waters Monitoring Network: The First Five Years. Analysis and Interpretation of Results, April 1988 - March 1993*. London.

Patrick, S., Waters, D., Juggins, S. & Jenkins, A. (1991) *The United Kingdom Acid Waters Monitoring Network: Site Descriptions and Methodology Report. Report to the Department of the Environment and Department of the Environment Northern Ireland*. London.

Pauly, D. (1995) Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology & Evolution*, **10**, 430.

Peralta, G., Frost, C.M., Rand, T.A., Didham, R.K. & Tylianakis, J.M. (2014) Complementarity and redundancy of interactions enhance attack rates and spatial stability in host-parasitoid food webs. *Ecology*, **95**, 1888–1896.

Perry, J.N., Rothery, P., Clark, S.J., Heard, M.S. & Hawes, C. (2003) Design, analysis and statistical power of the Farm Scale Evaluations of genetically modified herbicide-tolerant crops. *Journal of Applied Ecology*, **40**, 17–31.

Petchey, O.L., Beckerman, A.P., Riede, J.O. & Warren, P.H. (2008) Size, foraging, and food web structure. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 4191–4196.

Petchey, O.L., Beckerman, A.P., Riede, J.O. & Warren, P.H. (2011) Fit, efficiency, and biology: some thoughts on judging food web models. *Journal of Theoretical Biology*, **279**, 169–71.

Piechnik, D.A., Lawler, S.P. & Martinez, N.D. (2008) Food-web assembly during a classic biogeographic study: Species' 'trophic breadth' corresponds to colonization order. *Oikos*, **117**, 665–674.

Pimm, S.L. (1980) Properties of Food Webs. *Ecology*, **61**, 219–255.

Pimm, S.L. (1984) The complexity and stability of ecosystems. *Nature*, **307**, 321–326.

Pimm, S.L., Russell, G.J., Gittleman, J.L. & Brooks, T.M. (1995) The Future of Biodiversity. *Science*, **269**, 347–350.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & Team, T.R.C. (2014) nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-117, URL: http://CRAN.R-project.org/package=nlme.

Pocock, M.J.O., Evans, D.M. & Memmott, J. (2012) The Robustness and Restoration of a Network of Ecological Networks. *Science*, **335**, 973–977.

Poelen, J.H., Simons, J.D. & Mungall, C.J. (2014) Global biotic interactions: An open infrastructure to share and analyze species-interaction datasets. *Ecological Informatics*, **24**, 148–159.

Poisot, T., Canard, E., Mouillot, D., Mouquet, N., Gravel, D. & Jordan, F. (2012) The dissimilarity of species interaction networks. *Ecology Letters*, **15**, 1353–61.

R Core Team. (2013) R: A language and environment for statistical computing.

Racke, K.D. (1993) Environmental fate of chlorpyrifos. *Reviews of Environmental Contamination and Toxicology*, **131**, 1–150.

Raffaelli, D. (2007) Food Webs, Body Size and the Curse of the Latin Binomial. *From Energetics to Ecosystems: The Dynamics and Structure of Ecological Systems* (eds N. Rooney, K.S. McCann & D.L.G. Noakes), pp. 53–64. Springer Netherlands.

Raffaelli, D., Bullock, J.M., Cinderby, S., Durance, I., Emmett, B., Harris, J., Hicks, K., Oliver, T.H., Paterson, D. & White, P.C.L. (2014) *Big Data and Ecosystem Research Programmes*, 1st ed. Elsevier Ltd.

Raffaelli, D.G. & Friedlander, A.M. (2012) Biodiversity and ecosystem functioning: an ecosystem-level approach. *Marine Biodiversity and Ecosystem Functioning:*

*Frameworks, methodologies, and integration* (eds M. Solan, R.J. Aspden & D.M. Paterson), pp. 149–163. Oxford University Press, Oxford, UK.

Rall, B.C., Brose, U., Hartvig, M., Kalinkat, G., Schwarzmüller, F., Vucic-Pestic, O. & Petchey, O.L. (2012) Universal temperature and body-mass scaling of feeding rates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **367**, 2923–34.

Raven, P.J. & George, J.J. (1989) Recovery by riffle macroinvertebrates in a river after a major accidental spillage of chlorpyrifos. *Environmental Pollution*, **59**, 55–70.

Rawcliffe, R., Sayer, C.D., Woodward, G., Grey, J., Davidson, T.A. & Jones, J.I. (2010) Back to the future: using palaeolimnology to infer long-term changes in shallow lake food webs. *Freshwater Biology*, **55**, 600–613.

Redford, K.H., Groves, C., Medellin, R.A. & Robinson, J.G. (2012) Conservation Stories, Conservation Science, and the Role of the Intergovernmental Platform on Biodiversity and Ecosystem Services. *Conservation Biology*, **26**, 757–759.

Relyea, R.A. (2009) A cocktail of contaminants: how mixtures of pesticides at low concentrations affect aquatic communities. *Oecologia*, **159**, 363–76.

Relyea, R. & Hoverman, J. (2006) Assessing the ecology in ecotoxicology: a review and synthesis in freshwater systems. *Ecology Letters*, **9**, 1157–1171.

Reuman, D.C. & Cohen, J.E. (2005) Estimating relative energy fluxes using the food web, species abundance, and body size. *Advances in Ecological Research*, **36**, 137–182.

Rocha, O. & Duncan, A. (1985) The relationship between cell carbon and cell volume in freshwater algal species used in zooplanktonic studies. *Journal of Plankton Research*, **7**, 279–294.

Rockström, A., Steffen, J.W., Noone, K., Persson, Å., Chapin III, F.S., Lambin, E., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H., Nykvist, B., De Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falke, M. & Fole, J. (2009) Planetary Boundaries: Exploring the Safe Operating Space for Humanity. *Ecology and Society*, **14**, 32.

Rohr, R.P., Scherer, H., Kehrli, P., Mazza, C. & Bersier, L.-F. (2010) Modeling food webs: exploring unexplained structure using latent traits. *The American Naturalist*, **176**, 170–177.

Rosemond, A.D., Reice, S.R., Elwood, J.W. & Mulholland, P.J. (1992) The effects of stream acidity on benthic invertebrate communities in the south-eastern United States. *Freshwater Biology*, **27**, 193–209.

Round, F.E. (1990) Diatom Communities their Response to Changes in Acidity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **327**, 243–249.

Royal Society. (2009) *Reaping the Benefits: Science and the Sustainable Intensification of Global Agriculture*. London, UK.

Rubach, M.N., Crum, S.J.H. & Van Den Brink, P.J. (2011) Variability in the dynamics of mortality and immobility responses of freshwater arthropods exposed to chlorpyrifos. *Archives of Environmental Contamination and Toxicology*, **60**, 708–721.

de Ruiter, P.C., Neutel, A.M. & Moore, J.C. (1995) Energetics, Patterns of Interaction Strengths, and Stability in Real Ecosystems. *Science*, **269**, 1257–1260.

SanClements, M.D., Oelsner, G.P., McKnight, D.M., Stoddard, J.L. & Nelson, S.J. (2012) New insights into the source of decadal increases of dissolved organic matter in acid-sensitive lakes of the northeastern United States. *Environmental Science & Technology*, **46**, 3212–9.

Saska, P. (2008) Effect of diet on the fecundity of three carabid beetles. *Physiological Entomology*, **33**, 188–192.

de Sassi, C., Staniczenko, P.P.A. & Tylianakis, J.M. (2012) Warming and nitrogen affect size structuring and density dependence in a host-parasitoid food web. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **367**, 3033–41.

Sáyago, R., Lopezaraiza-Mikel, M., Quesada, M., Álvarez-Añorve, M.Y., Cascante-Marín, A. & Bastida, J.M. (2013) Evaluating factors that predict the structure of a commensalistic epiphyte-phorophyte network. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20122821.

Sayer, M.D.J., Reader, J.P. & Dalziel, T.R.K. (1993) Freshwater acidification: effects on the early life stages of fish. *Reviews in Fish Biology and Fisheries*, **3**, 95–132.

Scheffer, M. & Carpenter, S.R. (2003) Catastrophic regime shifts in ecosystems: linking theory to observation. *Trends in Ecology and Evolution*, **18**, 648–656.

Sicko-Goad, L.M., Schelske, C.L. & Stoermer, E.F. (1984) Estimation of intracellular carbon and silica content of diatoms from natural assemblages using morphometric techniques. *Limnology and Oceanography*, **29**, 1170–1178.

Simpson, J.C. & Norris, R.H. (2000) *Biological Assessment of River Quality: Development of AUSRIVAS Models and Outputs* (eds JF Wright, DW Sutcliffe, and MT Furse). Freshwater Biological Association, The Ferry House, Far Sawrey, Ambleside, Cumbria, LA22 0LP, UK.

Simpson, G.L., Shilland, E.M., Winterbottom, J.M. & Keay, J. (2005) Defining reference conditions for acidified waters using a modern analogue approach. *Environmental Pollution*, **137**, 119–133.

Skjelkvåle, B.L., Stoddard, J.L., Jeffries, D.S., Tørseth, K., Høgåsen, T., Bowman, J., Mannio, J., Monteith, D.T., Mosello, R., Rogora, M., Rzychon, D., Vesely, J., Wieting, J., Wilander, a. & Worsztynowicz, a. (2005) Regional scale evidence for improvements in surface water chemistry 1990-2001. *Environmental Pollution*, **137**, 165–176.

Slavik, K., Peterson, B.J., Deegan, L.A., Bowden, W.B., Hershey, A.E. & Hobbie, J.E. (2004) Long-term responses of the Kuparuk River ecosystem to phosphorus fertilization. *Ecology*, **85**, 939–954.

Solé, R. V., Ferrer-Cancho, R., Montoya, J.M. & Valverde, S. (2003) Selection, tinkering, and emergence in complex networks. *Complexity*, **8**, 20–33.

Solé, R. V. & Montoya, J.M. (2001) Complexity and fragility in ecological networks. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 2039–2045.

Squire, G.R., Rodger, S. & Wright, G. (2000) Community-scale seedbank response to

less intense rotation and reduced herbicide input at three sites. *Annals of Applied Biology*, **136**, 47–57.

Steinman, A.D., Lamberti, G.A. & Leavitt, P.R. (2006) Biomass and Pigments of Benthic Algae. *Methods in Stream Ecology*, 2nd ed (eds R.F. Hauer & G.A. Lamberti), pp. 357–379. Elsevier Academic Press Inc.

Stoddard, J.L., Jeffries, D.S., Lukewille, A., Clair, T.A., Dillon, P.J., Driscoll, C.T., Forsius, M., Johannessen, M., Kahl, J.S., Kellogg, J.H., Kemp, A., Mannio, J., Monteith, D.T., Murdoch, P.S., Patrick, S., Rebsdorf, A., Skjelkvale, B.L., Stainton, M.P., Traaen, T., van Dam, H., Webster, K.E., Wieting, J. & Wilander, A. (1999) Regional trends in aquatic recovery from acidification in North America and Europe. *Nature*, **401**, 575–578.

Strong, J.S. & Leroux, S.J. (2014) Impact of Non-Native Terrestrial Mammals on the Structure of the Terrestrial Mammal Food Web of Newfoundland, Canada. *Plos One*, **9**, e106264.

Sullivan, G.M. & Feinn, R. (2012) Using Effect Size - or Why the P Value Is Not Enough. *Journal of graduate medical education*, **4**, 279–82.

Tachet, H., Richoux, P., Bournaud, M. & Usseglio-Polatera, P. (2002) *Invertébrés d'Eau Douce*, 2nd correc. CNRS, Paris.

Tang, S., Pawar, S. & Allesina, S. (2014) Correlation between interaction strengths drives stability in large ecological networks. *Ecology Letters*, **17**, 1094–1100.

Thébault, E. & Fontaine, C. (2010) Stability of Ecological Communities and the Architecture of Mutualistic and Trophic Networks. *Science*, **329**, 853–856.

Thompson, M.S.A., Bankier, C., Bell, T., Dumbrell, A.J., Gray, C., Ledger, M.E., Lehman, K., McKew, B.A., Sayer, C.D., Shelley, F., Trimmer, M., Warren, S.L. & Woodward, G. Gene-to-ecosystem impacts of a catastrophic pesticide spill: testing a multilevel bioassessment approach in a river ecosystem. *Freshwater Biology*.

Thompson, R.M., Brose, U., Dunne, J.A., Hall, R.O., Hladyz, S., Kitching, R.L., Martinez, N.D., Rantala, H., Romanuk, T.N., Stouffer, D.B. & Tylianakis, J.M. (2012) Food webs: reconciling the structure and function of biodiversity. *Trends in Ecology and Evolution*, **27**, 689–697.

Thompson, R.M., Dunne, J.A. & Woodward, G. (2012) Freshwater food webs: towards a more fundamental understanding of biodiversity and community dynamics. *Freshwater Biology*, **57**, 1329–1341.

Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R. & Polasky, S. (2002) Agricultural sustainability and intensive production practices. *Nature*, **418**, 671–7.

Townsend, C.R., Hildrew, A.G. & Schofield, K. (1987) Persistence of Stream Invertebrate Communities in Relation to Environmental Variability. *Journal of Animal Ecology*, **56**, 597–613.

Traas, T.P., Janse, J.H., Van den Brink, P.J., Brock, T.C.M. & Aldenberg, T. (2004) A freshwater food web model for the combined effects of nutrients and insecticide stress and subsequent recovery. *Environmental Toxicology and Chemistry*, **23**, 521–

529.
Triebskorn, R., Adam, S., Behrens, A., Beier, S., Bohmer, J., Braunbeck, T., Casper, H., Dietze, U., Gernhofer, M. & Honnen, W. (2003) Establishing Causality between Pollution and Effects at Different Levels of Biological Organization: The VALIMAR Project. *Human and Ecological Risk Assessment*, **9**, 171–194.
Tylianakis, J.M., Laliberté, E., Nielsen, A. & Bascompte, J. (2010) Conservation of species interaction networks. *Biological Conservation*, **143**, 2270–2279.
Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007) Habitat modification alters the structure of tropical host-parasitoid food webs. *Nature*, **445**, 202–205.
Vázquez, D.P., Chacoff, N.P. & Cagnolo, L. (2009) Evaluating multiple determinants of the structure of plant-animal mutualistic networks. *Ecology*, **90**, 2039–2046.
Vázquez, D.P., Melián, C.J., Williams, N.M., Blüthgen, N., Krasnov, B.R., Poulin, R., Vazquez, D.P., Melian, C.J., Williams, N.M., Bluthgen, N., Krasnov, B.R. & Poulin, R. (2007) Species abundance and asymmetric interaction strength in ecological networks. *Oikos*, **116**, 1120–1127.
Verdú, M. & Valiente-Banuet, A. (2011) The relative contribution of abundance and phylogeny to the structure of plant facilitation networks. *Oikos*, **120**, 1351–1356.
Vizentin-Bugoni, J., Maruyama, P.K. & Sazima, M. (2014) Processes entangling interactions in communities: forbidden links are more important than abundance in a hummingbird-plant network. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20132397.
Warren, C.R. (2008) Rapid Measurement of Chlorophylls with a Microplate Reader. *Journal of Plant Nutrition*, **31**, 1321–1332.
Watts, D.J. & Strogatz, S.H. (1998) Collective dynamics of 'small-world' networks. *Nature*, **393**, 440–2.
Van Wijngaarden, R.P.A., Van den Brink, P.J., Crum, S.J.H., Brock, T.C.M., Leeuwangh, P. & Voshaar, O.J.H. (1996) Effects of the insecticide Dursban 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: I . comparison of short-term toxicity between the laboratory and the field. *Environmental Toxicology and Chemistry*, **15**, 1133–1142.
Williams, R.J. & Martinez, N.D. (2000) Simple rules yield complex food webs. *Nature*, **404**, 180–183.
Williams, R.J. & Martinez, N.D. (2004) Limits to trophic levels and omnivory in complex food webs: theory and data. *The American Naturalist*, **163**, 458–68.
Wilson, E.E., Sidhu, C.S., Levan, K.E. & Holway, D.A. (2010) Pollen foraging behaviour of solitary Hawaiian bees revealed through molecular pollen analysis. *Molecular Ecology*, **19**, 4823–4829.
De Wit, H.A., Mulder, J., Hindar, A. & Hole, L. (2007) Long-term increase in dissolved organic carbon in streamwaters in Norway is response to reduced acid deposition. *Environmental Science & Technology*, **41**, 7706–7713.
Woodward, G., Benstead, J.P., Beveridge, O.S., Blanchard, J., Brey, T., Brown, L.E., Cross, W.F., Friberg, N., Ings, T.C., Jacob, U., Jennings, S., Ledger, M.E., Milner,

A.M., Montoya, J.M., O'Gorman, E.J., Olesen, J.M., Petchey, O.L., Pichler, D.E., Reuman, D.C., Thompson, M.S.A., Van Veen, F.J.F. & Yvon-Durocher, G. (2010a) Ecological Networks in a Changing Climate. *Advances in Ecological Research*, **42**, 71–138.

Woodward, G., Blanchard, J., Lauridsen, R.B., Edwards, F.K., Jones, J.I., Figueroa, D., Warren, P.H. & Petchey, O.L. (2010b) Individual-Based Food Webs: Species Identity, Body Size and Sampling Effects. *Advances in Ecological Research*, **43**, 211–266.

Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Reuman, D.C. & Ledger, M.E. (2012a) Climate change impacts in multispecies systems: drought alters food web size structure in a field experiment. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **367**, 2990–2997.

Woodward, G., Dumbrell, A.J., Baird, D.J. & Hajibabaei, M. (2014) Preface. *Advances in Ecological Research*, **51**, ix–xiii.

Woodward, G., Gessner, M.O., Giller, P.S., Gulis, V., Hladyz, S., Lecerf, A., Malmqvist, B., McKie, B.G., Tiegs, S.D., Cariss, H., Dobson, M., Elosegi, A., Ferreira, V., Graca, M.A.S., Fleituch, T., Lacoursiere, J.O., Nistorescu, M., Pozo, J., Risnoveanu, G., Schindler, M., Vadineanu, A., Vought, L.B.-M. & Chauvet, E. (2012b) Continental-Scale Effects of Nutrient Pollution on Stream Ecosystem Functioning. *Science*, **336**, 1438–1440.

Woodward, G. & Hildrew, A.G. (2001) Invasion of a stream food web by a new top predator. *Journal of Animal Ecology*, **70**, 273–288.

Woodward, G. & Hildrew, A.G. (2002) Food web structure in riverine landscapes. *Freshwater Biology*, **47**, 777–798.

Woodward, G., Speirs, D.C. & Hildrew, A.G. (2005) Quantification and resolution of a complex, size-structured food web. *Advances in Ecological Research*, **36**, 85–135.

Wright, J.F., Clarke, R.T., Gunn, R.J.M., Kneebone, N.T. & Davy-Bowker, J. (2004) Impact of major changes in flow regime on the macroinvertebrate assemblages of four chalk stream sites, 1997-2001. *River Research and Applications*, **20**, 775–794.

Wright, J.F., Furse, M.T. & Armitage, P.D. (1993) RIVPACS - a technique for evaluating the biological quality of rivers in the UK. *European Water Pollution Control*, **3**, 15–25.

Wright, J.F., Gunn, R.J.M., Winder, J.M., Wiggers, R., Vowles, K., Clarke, R.T. & Harris, I. (2002) A comparison of the macrophyte cover and macroinvertebrate fauna at three sites on the River Kennet in the mid 1970s and late 1990s. *Science of The Total Environment*, **282-283**, 121–142.

Xuereb, B., Noury, P., Felten, V., Garric, J. & Geffard, O. (2007) Cholinesterase activity in *Gammarus pulex* (Crustacea Amphipoda): characterization and effects of chlorpyrifos. *Toxicology*, **236**, 178–89.

Yan, N.D., Leung, B., Keller, W., Arnott, S.E., Gunn, J.M. & Raddum, G.G. (2003) Developing conceptual frameworks for the recovery of aquatic biota from acidification. *Ambio*, **32**, 165–169.

Yodzis, P. (1989) *Introduction to Theoretical Ecology*. Harper & Row, Cambridge, UK.
Young, R.G., Matthaei, C.D. & Townsend, C.R. (2008) Organic matter breakdown and ecosystem metabolism: functional indicators for assessing river ecosystem health. *Journal of the North American Benthological Society*, **27**, 605–625.

# Appendix A | The stability of the Upland Waters Monitoring Network food webs

## A.1 Introduction

The relationship between complexity and stability in natural ecosystems has long been a central focus of ecological research (MacArthur 1955; Elton 1958; May 1972; Yodzis 1981; Pimm 1984; McCann 2000). Initially it was suggested that diversity and complexity should stabilise food webs (MacArthur 1955; Elton 1958), as they increase the redundancy of nodes and links thus reducing the important of any one node or interaction. Early theoretical work by May (May 1972) contradicted this, and found that complexity increased instability. However, this work was done using randomly constructed food webs with interaction strengths sampled from a normal distribution, which is not what is observed in nature. In fact, Yodzis (1981) found that empirically constructed food webs were more stable than their random counterparts, suggesting that the distribution of interaction strengths could be crucial in determining the stability of natural systems.

The structure and stability of freshwater food webs as they recover from acidification is an important avenue of research, as it is the 'ecological inertial' of these food webs which is a cited mechanism to explain their slow recovery from acidification. It has been suggested that acidified food webs are more stable and resistant the re-invasion of acid sensitive species and thus resists change as acidity ameliorates (see Chapter 3; Lundberg *et al.* 2000; Ledger & Hildrew 2005; Kernan *et al.* 2010; Layer *et al.* 2010). In an attempt to address this hypothesis, I have assessed the structure of the UWMN food webs and variability of the communities across the pH gradient. To do this I first approximated the likely biomass flows across the food webs using allometric scaling relationships based on body mass (Tang *et al.* 2014). I then measured multiple network metrics and assess their relationship with hydrochemical stress (pH gradient).

Turnover of species and links is the reciprocal of persistence (the time a community remains unchanged after a perturbation; Figure 0.A; Pimm 1984), a more persistent community is a more stable community. Bray-Curtis link turnover is

measured much the same as species turnover. Similarly, variation in species abundances is a similar mechanism, and often linked to community stability (e.g. Pimm 1984; Mellin *et al.* 2010). A community whose species abundances were highly variable overtime would be expected to be more unstable.

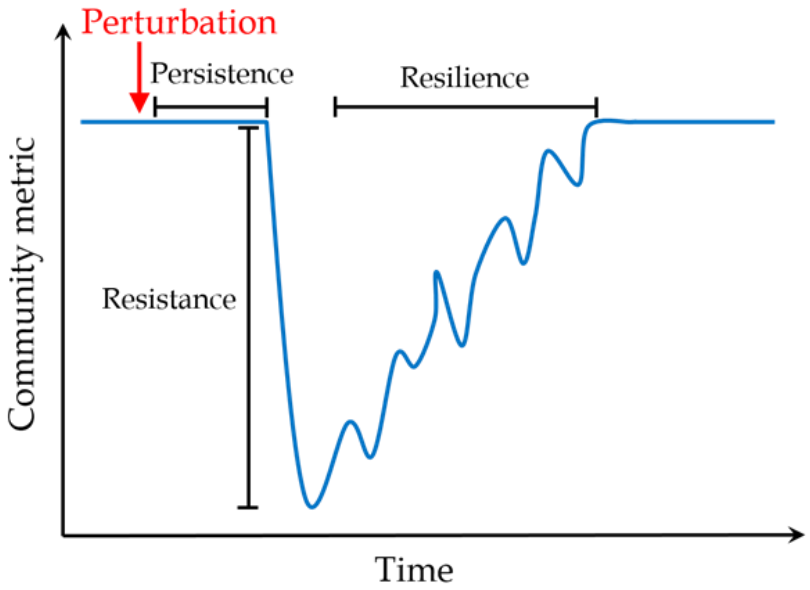



**Figure 0.A**. Some aspects of community stability. After a perturbation, persistence refers to the time a variable lasts before it is changed to a new variable, resilience refers to the time it takes for a variable to return to equilibrium, and resistance refers to the degree to which a variable is changed.

Recent advances allow the examination of network substructure. The substructural scale lies between that of an individual node and the whole network. The concept of a core/periphery structure (Figure 0.B) in social networks, whereby the structure of a network is governed by a highly interconnected core surrounded by a more loosely connected periphery, has been a major avenue of investigation in complex network research (Borgatti & Everett 2000; Csermely *et al.* 2013). Examination of the relative core size yields an insight into the flexibility and controllability (Csete & Doyle 2004; Liu, Slotine & Barabási 2011) of a variety of networks. Within a

biological context, a large core might indicate greater redundancy within the flows of a network, and therefore greater robustness to perturbations such as species loss (Appendix B). Thus, given that a greater core size might provide greater redundancy, it might also provide greater persistence and resistance (the degree to which a community is changed following a perturbation; Figure 0.A) to perturbations. The rich club coefficient measures the connectivity between nodes within the core (Zhou & Mondragon 2004), highly connected nodes within the core of networks heavily influence the functioning of that network, as has been demonstrated in the flow of rumours in social networks (Masuda & Konno 2006) or the transfer of information in the Internet (Zhou & Mondragón 2004). The high connectivity of the core of freshwater food webs have been found to buffer the food web from the effects of drought (Appendix B; Lu *et al.* in review). Experimental food webs exposed to drought conditions lose many species to extinction, they undergo major re-wiring within the core which maintains the overall core/periphery structure which in turn maintains the networks robustness to simulated species removal.

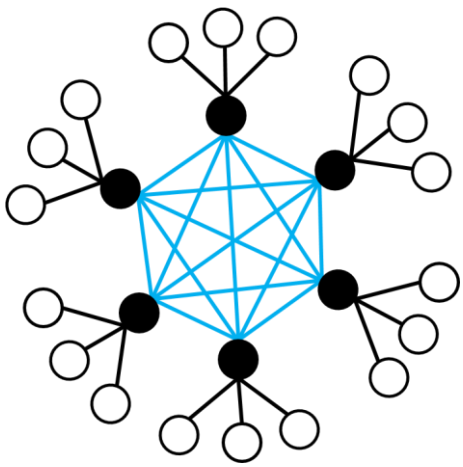


**Figure 0.B.** An example network with a strong core/periphery structure. Core nodes (solid black) are both highly connected, and highly connected to one another (blue links). Peripheral nodes (empty black) are both weakly connected, and weakly connected to core nodes.

The robustness of a food web to species extinctions gives another measure of stability. If a system is able to withstand many species extinctions before cascading secondary extinctions occur then that system is deemed to be more robust and hence more stable. Species extinctions from a food web can be simulated by sequential targeted removal of species, the systems robustness is determined from the point at which the food web collapses (Dunne *et al.* 2002).

The distribution of flows across a food web is theoretically linked to the stability of that system, and the presence of few strong and many weak interactions within a system has been suggested to have a stabilising effect (Kokkoris *et al.* 2002; Neutel *et al.* 2007), thus unstable systems might be expected to have a more even distribution of biomass flows. The distribution of flows through a network can be measures using Ulanomicz's (2004) Mutual Information metric.

Borrelli (2015) analysed the substructure of food webs, and found that some three-node motifs were more dynamically stable than others (tri-trophic chain, apparent competition and direct competition), and also found that these same motifs appeared in ecological networks more often than would be expected by chance. Thus, the occurrence of these three-node motifs might have a stabilising effect on food web structure.

Here I investigate the hypothesis that, concurrent with previous theory and findings (Lundberg *et al.* 2000; Ledger & Hildrew 2005; Kernan *et al.* 2010; Layer *et al.* 2010), acidified food webs will be more stable than non-acidified food webs.

## A.2 Methods

### A.2.1 Community matrix

The Upland Waters Monitoring Network (UWMN) food webs from chapter 3 were used in this analysis. The community matrix for each food webs was augmented

with estimates of biomass flow using metabolic scaling theory using the method of Tang *et al.* (2014). The off-diagonal entries in the community matrix M was parameterised using Holling's type 1 functional response. The type I functional response is given by the classic Lotka-Volterra equations:

$$\frac{dx_i}{dt} = x_i \left[ g_i(x_i) - \sum_{j \in pred(i)} a_{ij} x_{ij} + \sum_{k \in prey(i)} e_{ij} a_{ij} x_k \right] \quad (1)$$

Where $x_i$ is the biomass of species $i$, $g_i$ is a function depending on $x_i$ only, usually representing the growth of species $i$, determined by its reproduction and death rates. For any consumer-resource pair, $a_{ij}$ is the search rate of consumer $i$ for its resource $j$. $e_{ij}$ is the conversion efficiency of resource into consumer biomass.

Next, metabolic scaling theory (Peters 1986; Yodzis & Innes 1992; Brown *et al.* 2004; Reuman *et al.* 2008, 2009; Pawar *et al.* 2012; Rall *et al.* 2012) was used to find estimates for these parameters.

A. Body size scaling for biomass $x_i^*$:

Body size information harvested from the literature (Brose *et al.* 2006; Gilljam *et al.* 2011; Pawar *et al.* 2012; Ledger *et al.* 2012) as well as data collected as part of the River Kennet study (see Chapter 5) was used to ascribe an average body mass to each UWMN species. In some cases it was necessary to use a genus averaged value as no data at the species level could be found. Equation (2) was used to ascribe a biomass to each species:

$$x_i^* = 10^{x_0 + 3\gamma + \varepsilon_i} \cdot m_i^{1+\gamma} \quad (2)$$

Where $m_i$ is the body mass of species $i$ in kg. The parameter $\gamma$ is the scaling exponent

for numerical abundance, typically taking values between −1.25 and −0.1, and taken here to be −0.675 (as in Tang *et al.* (2014)). The intercept parameter $x_0$ is negative and was set to −1.16 for all webs (Cyr *et al.* 1977; Leaper & Raffaelli 1999). $\varepsilon_i$'s denote the residuals of the regression line, and were sampled from a Gaussian distribution with mean zero and standard deviation of 0.1 (as in Tang *et al.* (2014)).

B. Body size scaling for mass-specific search rate, $a_{ij}$:

To find a mass specific search rate for each consumer $i$ and resource $j$ the results of (Pawar *et al.* (2012), as in Tang *et al.* (2014) were used:

$$a_{ij} = 10^{a_0} m_i^{\beta_i} f(k_{ij}) \quad (3)$$

where $a_0$ = −3.50 is a constant, $m_i$ is the body size of consumer $i$ (in kg), and $k_{ij} = \frac{m_j}{m_i}$ is the resource to consumer body size ratio. $f(k_{ij}) = \frac{k_{ij}^{0.46}}{1+k_{ij}^{K}}$ is a function that quantifies the well-documented unimodal relationship between search and consumption rates and size ratios (see Tang *et al.* 2014). As in Tang *et al.* (2014) and based on previous results $K = 2$. To account for uncertainty in the scaling relationship, the exponent $\beta_i$ was sampled independently from a normal distribution with mean -0.15 and standard deviation 0.052 for each consumer $i$.

C. Conversion efficiency $e_{ij}$

Conversion efficiency was assumed to be a uniformly distributed random variable within empirically feasible ranges (0.2 ± 0.1 for herbivores, and 0.5 ± 0.1 for carnivores). Tang *et al.* (2014) found their main results were insensitive to choice of $e_{ij}$.

### A.2.2 Measure of hydrochamical stress

Principal Component Analysis (PCA) was performed on the water chemistry data of each site. The water chemistry variables used were yearly average pH (also yearly minimum), Acid Neutralising Capacity (also yearly minimum), alkalinity, $H^+$, conductivity, $NO_3$, Soluble Monomeric Aluminium, Soluble Non-Labile Monomeric Aluminium, Soluble Labile Monomeric Aluminium, Dissolved Organic Carbon, Na, Cl, $SO_4$, $PO_4$. Yearly mean (or minimum) values for these variables were centred to zero and scaled by their standard deviations, and sample scores on the first PC axis (PC1) extracted for use as a proxy for water chemical stress.

The Euclidean distance in multivariate space between the first and last sampling year for each site was recoded as a proxy for the degree of change in hydrochemical variables over the course of monitoring (Figure 0.C).

Each network metric was regressed against PC1, or the Euclidean distance for that site, and any trend assessed with Generalised Linear Mixed Effects models. For each model, site and year were used as random effects, but a range of random effects structures were investigated for each response variable, the best model was selected on the basis of AIC.

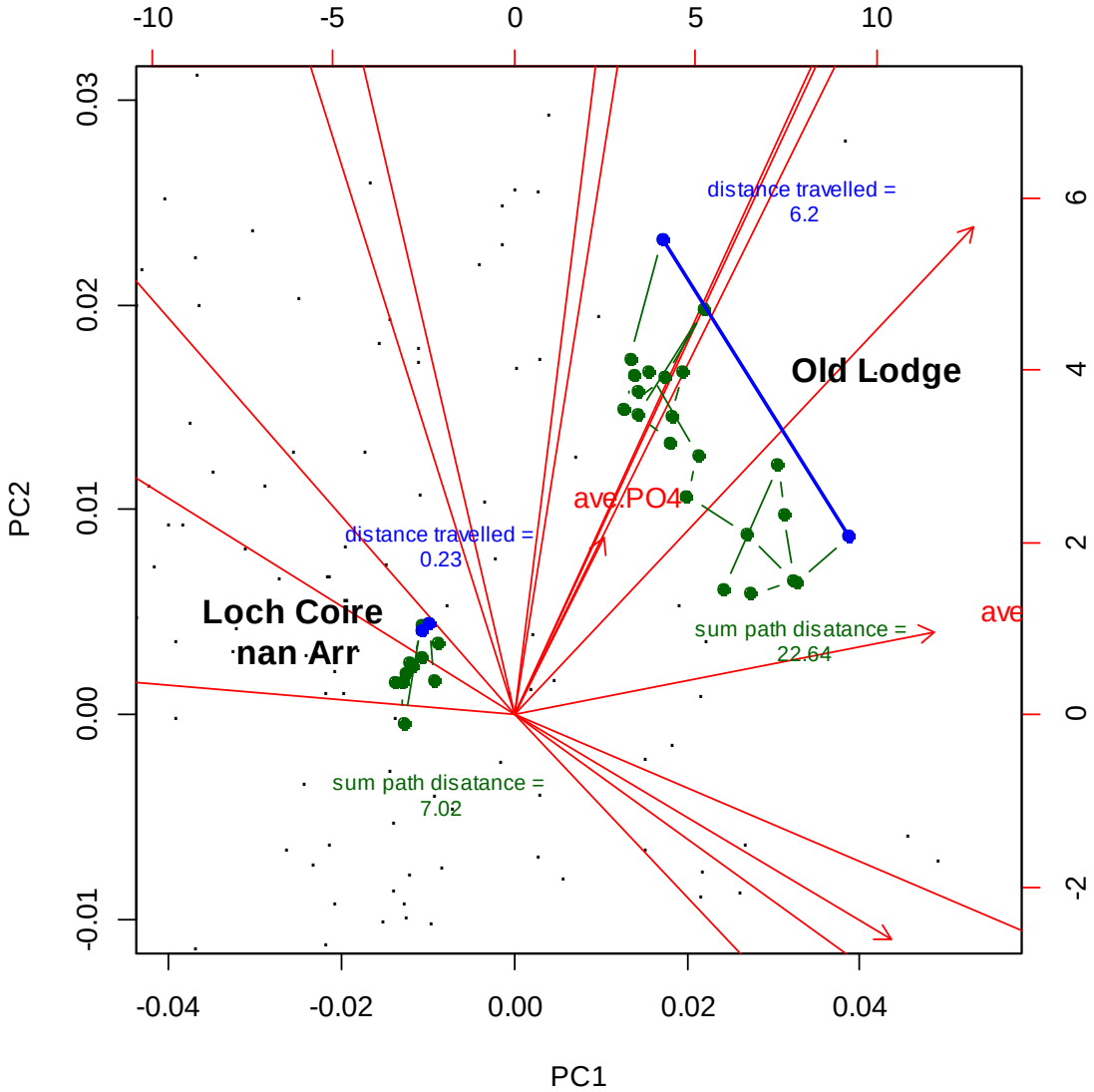



**Figure 0.C**. An example of two sites plotted in multivariate hydrochemical space. Old Lodge is located in the south of England, it was highly acidified in the 1980's and has changed markedly in its acidity over the course of monitoring. Loch Coir nan Arr is located in the far north west of Scotland, was not considered acidified at the onset of monitoring and it's pH has changed very little since then.

### A.2.3 Measures of stability

A range of metrics were used to quantify the stability of these food webs over time and in relation to the degree of hydrochemical stress they were under. Link turnover across the collection of food webs was measured using Bray-Curtis dissimilarity in the vegan package (Oksanen *et al.* 2012).

The coefficient of variation in the relative abundance of each species over time at each site was calculated, and a mean value across all species recorded for each site.

The core/periphery structure of each food web was measured, the core of a food web is defined as a cohesive compartment of the network where species are both highly connected and highly interconnected. This structure is then surrounded by a more loosely connected periphery. To find the core/periphery boundary nodes were ranked by their degree (number of links). A node with a rank $r$ has degree $k_r$. The number of links that this node shares with nodes of a higher rank is $k_r^+$. The core is defined by detecting a change of the behaviour of $k_r^+$ as a function of $r$, and the boundary of the core is defined by the node with rank $r^*$ where $k_{r*}^+ > k_r^+$ for $r > r^*$ (Borgatti & Everett 2000). The size of each food webs core, proportional to its total size was recoded.

The density of connections within the core of each food web was measured using the Rich Club score (Ma & Mondragón 2015):

$$\emptyset_r = \frac{2}{r(r-1)} \sum_{i=1}^{r} k_i^+ = \frac{2E_r}{r(r-1)} \quad (4)$$

where $E_r$ is the number of links shared by the highest ranked r nodes and $r(r-1)/2$ is the maximum number of possible links among these nodes. The connectivity of a core is given by $\emptyset_{r*}$ whereby a fully connected core has a value of $\emptyset_{r*} = 1$ and a fully disconnected core gives $\emptyset_{r*} = 0$.

The robustness of each network to simulated species loss was recoded using the method of Dunne *et al.* (2002), where by species were ordered by their degree, and sequentially removed from the network. Secondary extinctions occurred where consumers were left with no resources. The total number of primary extinctions required to cause network collapse (the loss of 50% of the food web), proportional to

total network size was recoded.

In order to examine the distribution of flows, the Mutual Information (MI) of each food web was measured. MI measures the evenness of flows across a network, such that high values indicate a more uneven distribution of flows (Ulanowicz 2004):

$$MI = k \sum_{j,i} \left(\frac{T_{ji}}{T..}\right) log\left(\frac{T_{ji}T..}{T_{j.}T_{.i}}\right) \quad (5)$$

where $T_{ij}$ is the rate of the internal transfer from resource species $j$ to consumer species $i$, and $k$ is a scalar constant.

The occurrence of certain motifs within each food web was examined. For each food web, 30 null networks was created using the Curveball algorithm (Strona *et al.* 2014), which maintains the number of consumers and resources each node has, but randomises who those connections are with. The frequency of each of teh three motifs was measured in each of the 30 randomisations of each food web, and a z-score computed as follows:

$$z_i = \frac{X_i - \overline{X_l}}{\sigma_l} \quad (6)$$

Where $X_i$ is the frequency of the $i$th motif in each empirical food web, $\overline{X_l}$ the mean frequency of the $i$th motif in the randomised networks, and $\sigma_l$ the standard deviation. A z-score greater than or less than 0 indicates that the occurrence of that particular motif in the food web is greater than or less than what you would expect by chance.

## A.3 Results

The first PCA axis (PC1) corresponded closely with the acidity gradient across the UWMN sites, where high PC1 values refer to low pH and high aluminium concentrations (Figure 0.D).

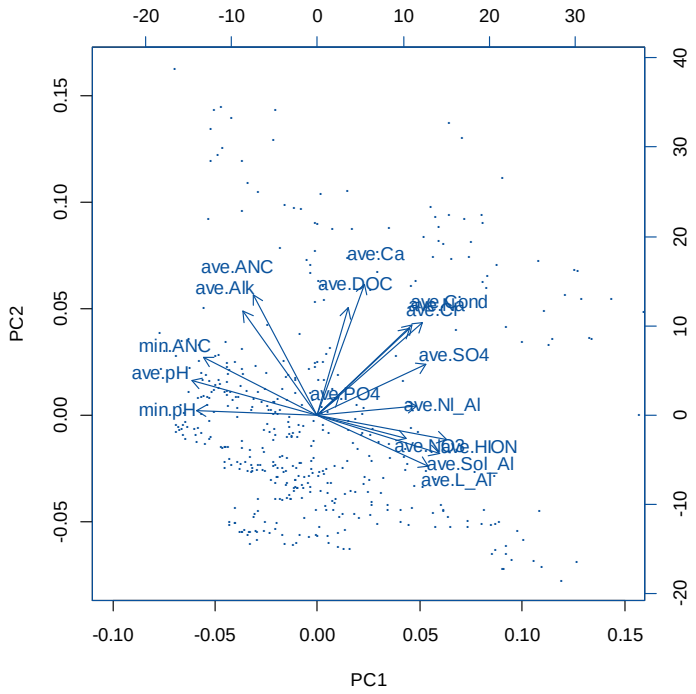



**Figure 0.D.** Ordination of hydochemical data. The first axis corresponds strongly to an acidity gradient, while the second axis is more related to $Ca^{2+}$ and DOC concentrations. min.pH = minimum yearly pH, ave.pH = mean yearly pH, min.ANC = minimum annual ANC, ave.Alk = mean yearly alkalinity, ave.ANC = mean yearly ANC, ave.DOC = mean yearly DOC, ave.Ca = mean yearly $Ca^{2+}$, ave.Na = mean yearly $Na^{+}$, ave.Cl = mean annual $Cl^{-}$, ave.Cond = mean annual conductivity, ave.PO4 = mean annual $PO_4$, ave.SO4 = mean annual $SO_4$, ave.Nl_Al = mean annual non-labile aluminium, ave.HION = mean annual $H^{+}$, ave.NO3 = mean annual $NO_3$, ave.Sol_Al = mean annual soluable aluminium, ave.L_Al = mean annual labile aluminium.

Link turnover between food webs of consecutive years was greater for stream food webs under greater hydrochemical stress, and was unaffected by hydrochemical stress at lake sites (Table 0.1, Figure 0.E). The link composition of stream food webs experienced greater turnover from one year to the next under greater hydrochemical stress.

**Table 0.1**. Statistics of fit for the multiple linear models. All mixed effects models include site and year as random effects (see main text for details). PC1 and PC2 refers to site scores taken from the first and second axis of the PCA performed on hydrochemical data (**Figure 0.D**).

| Response variable | Predictor variable & interactions | d.f. | F-value | P-value |
|---|---|---|---|---|
| Mixed effects models | | | | |
| Link turnover | PC1 | 1 | 29.669 | <0.0001 |
| | PC2 | 1 | 24.551 | <0.0001 |
| | type | 1 | 127.765 | <0.0001 |
| | PC2 * type | 1 | 19.691 | <0.0001 |
| Core size | PC1 | 1 | 19.67 | <0.0001 |
| | type | 1 | 37.94 | <0.0001 |
| Rich club score | PC1 | 1 | 11.44 | 0.0039 |
| | PC2 | 1 | 11.53 | 0.0029 |
| Robustness | PC1 | 1 | 22.31 | <0.0001 |
| | type | 1 | 38.26 | <0.0001 |
| Log(Mutual Information) | PC1 | 1 | 72.07 | <0.0001 |
| Linear models | | | | |
| Coefficient of variation (species abundance) | Euclidean distance | 1 | 7.38 | 0.0129 |

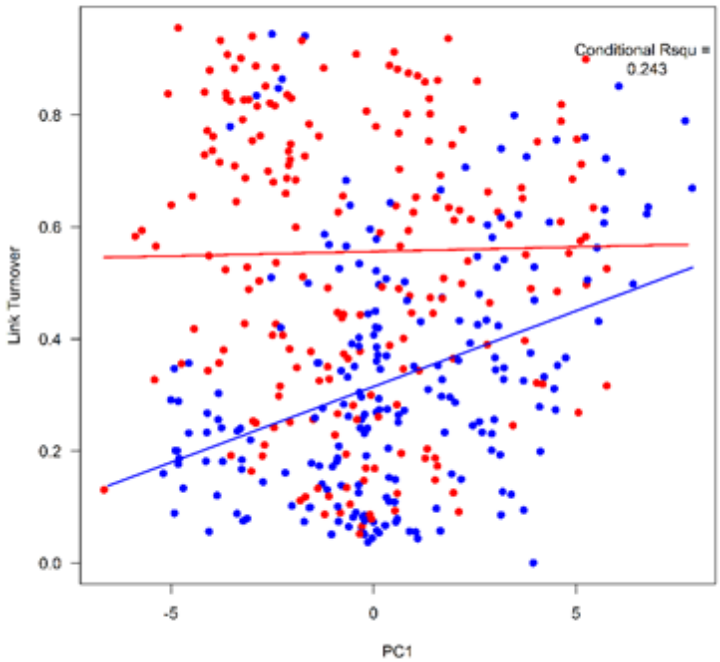



**Figure 0.E.** Link turnover increases with hydrochemical stress at stream sites (blue), and is unchanged at lake sites (red). Lines give the fit of the mixed effects model (see Table 0.1).

Species relative abundances was more variable over time at those sites which had experienced the most change in their hydrochemistry (**Figure 0.F**). One site stood apart from the general relationship, Loch Coire Fionnaraich had the lowest variability in species abundances, but had also been monitored for fewest years, only 10 whilst for the other sites the mean ± s.e.m. was 19.64 ± 0.80 years.

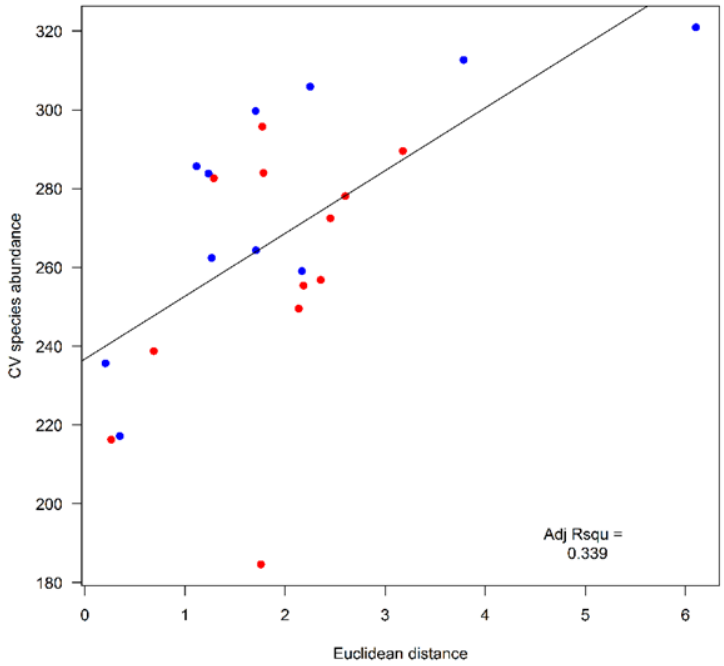



**Figure 0.F**. The mean coeficient of variation (CV) in individual species relative abundances for each site is greater at those sites who's hydrochemistry had changed the most over the course of monitoring (the euclidean distance in multivariatye space between the first and years of monitoring for each site). The line gives the fit of the linear model (see Table 0.1). Blue = stream sites, red = lake sites.

The relative core size was greater in food webs under less hydrochemical stress (Figure 0.G). Acidified food webs had smaller cores, this relationship was the same for both streams and lakes, although stream food webs had larger cores than lake food webs.

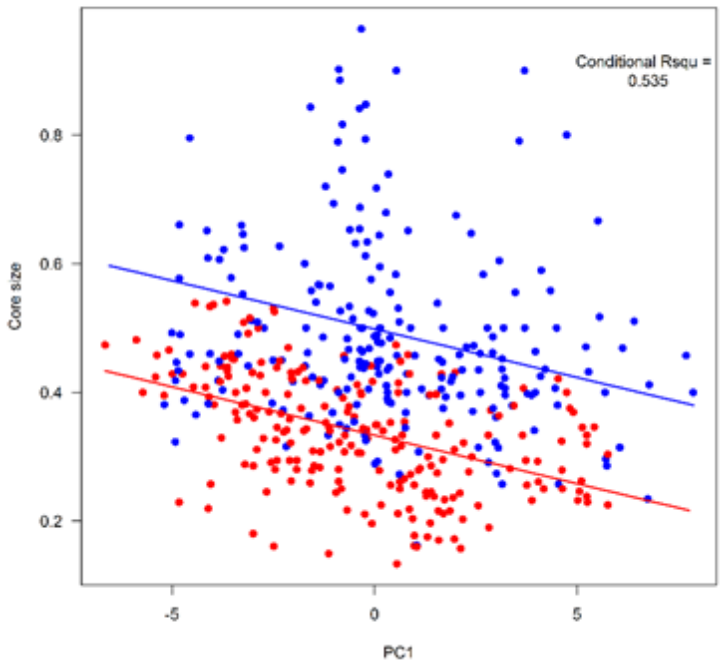



**Figure 0.G**. The size of the food web cores (relative to the whole network size) varies with hydrochemical stress. Those food webs under more hydrochemical stress have smaller cores. Lines give the fit of the mixed effects model (see Table 0.1). Blue = stream sites, red = lake sites.

The density of connections within the core of each food web, as measured by the rich club coefficient, was higher in those food webs which were under greater hydrochemical stress (Figure 0.H). As acidity increased, so too did the density of connections within the cores of the food webs, this relationship was found to be the same for streams and lakes.

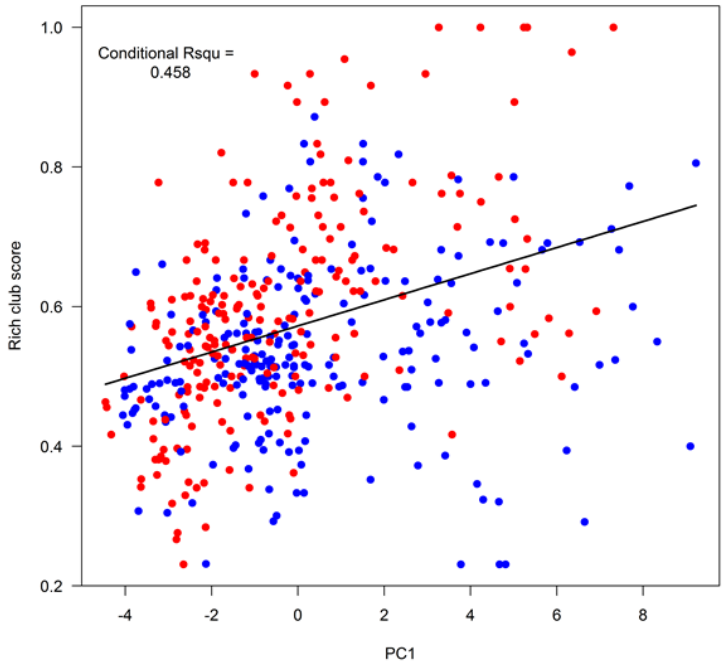



**Figure 0.H**. The density of connections within the core of each food web was higher for those food webs under greater hydrochemical stress. Line give the fit of the mixed effects model (see Table 0.1). Blue = stream sites, red = lake sites.

The robustness of food webs to the simulated removal of high degree nodes was lower for those which were under more hydrochemical stress (Figure 0.I). This relationship was found to be the same for lakes and streams, although lake food webs were less robust than stream food webs.

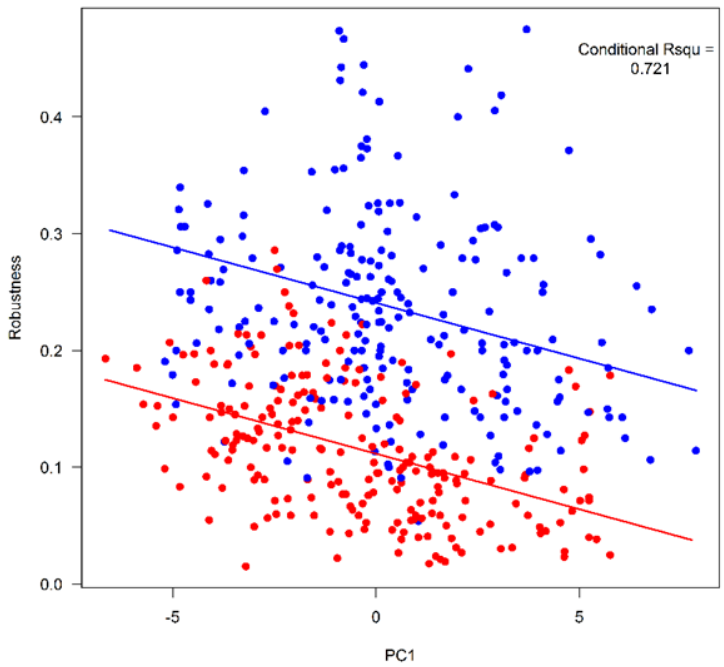



Figure 0.I. The robustness of each food web to simulated species removal was lower for those food webs under higher hydochemical stress. Lines give the fit of the mixed effects model (see Table 0.1). Blue = stream sites, red = lake sites.

The distribution of flows across the food webs was affected by the hydrochemistry at that site, biomass flows were more evenly distributed in those food webs which were under greater hydrochemical stress (Figure 0.J). As acidity decreased the distribution of flows became more uneven, as reflected in a higher mutual information score. This was found to be the same for lake and stream sites.

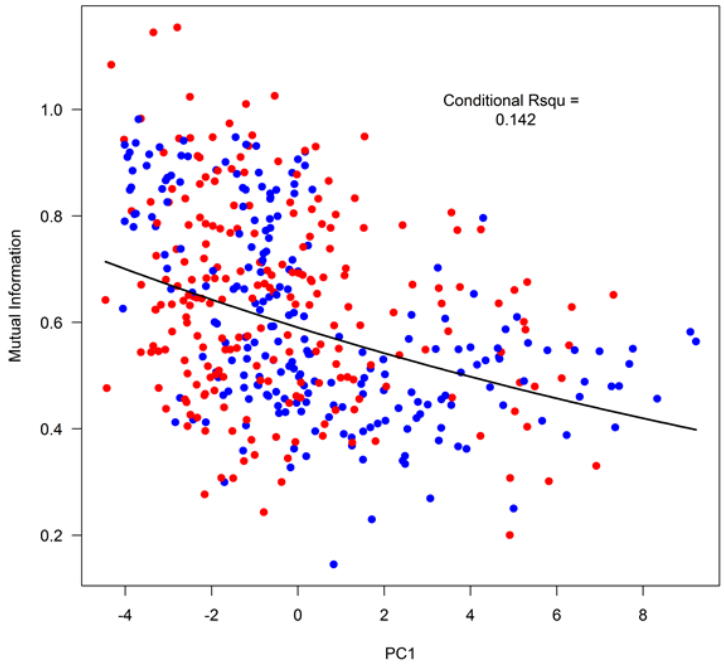



Figure 0.J. The mutual information was lower in food webs that were under higher hydrochemical stress. The distribution of flows was more even in food webs which were under more hydrochemical stress. Lines give the fit of the mixed effects models (see Table 0.1). Blue = stream sites, red = lake sites.

All food webs, at all sites in all years were found to have a higher occurrence of motifs associated with dynamic stability (tri-trophic chain, apparent competition and direct competition) than would be expected by chance (Figure 0.K, Figure 0.L). This was not related to the level of acidity the food web was exposed to.

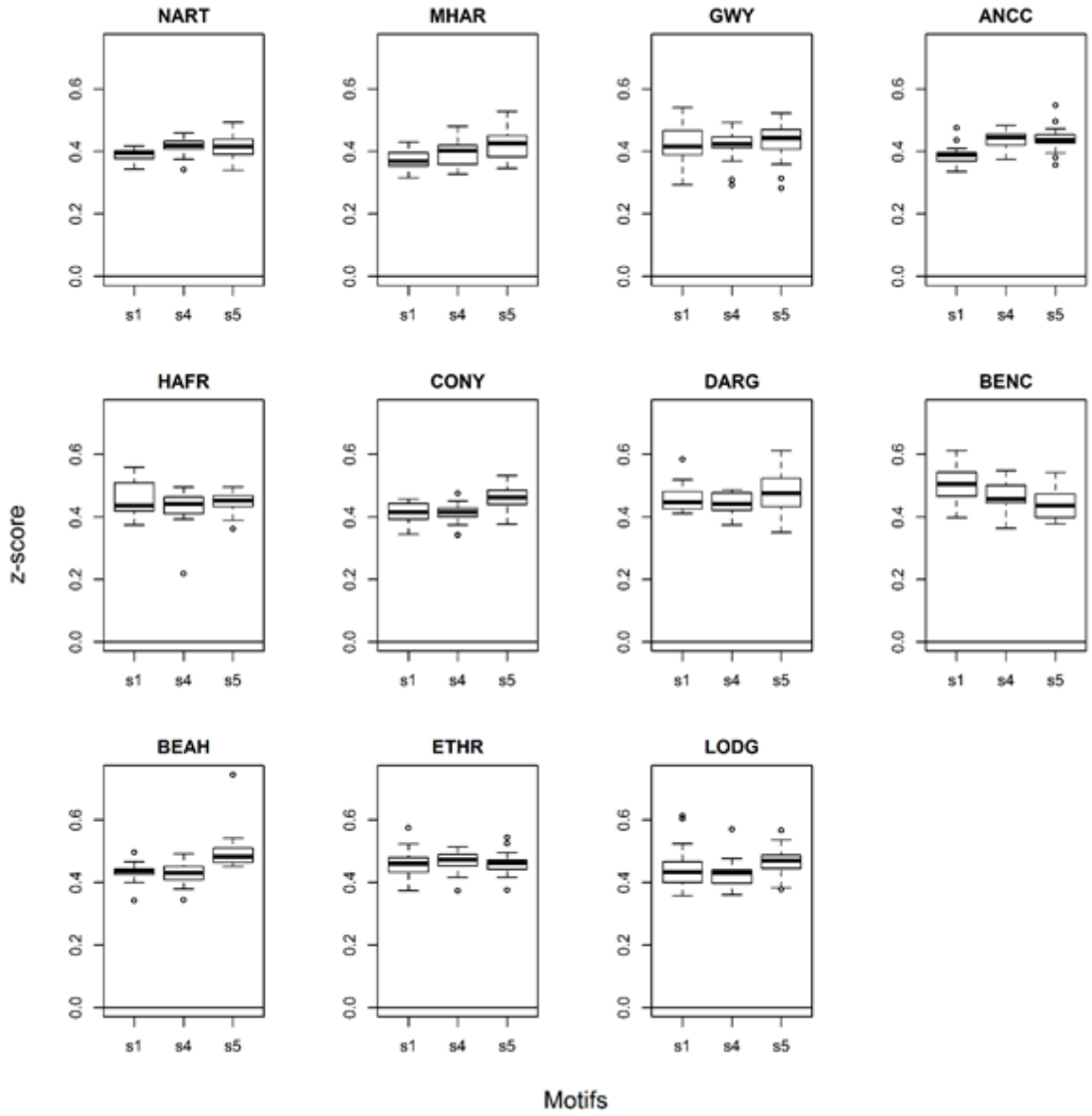



**Figure 0.K**. The z-scores for the occurrence of three three-node motifs at each stream site. s1 = tri-trophic chain, s4 = apparent competition, s5 = direct competition. Boxplots represent the median and interquartile range of the z-scores, one z-scores for each food web over the course on monitoring.

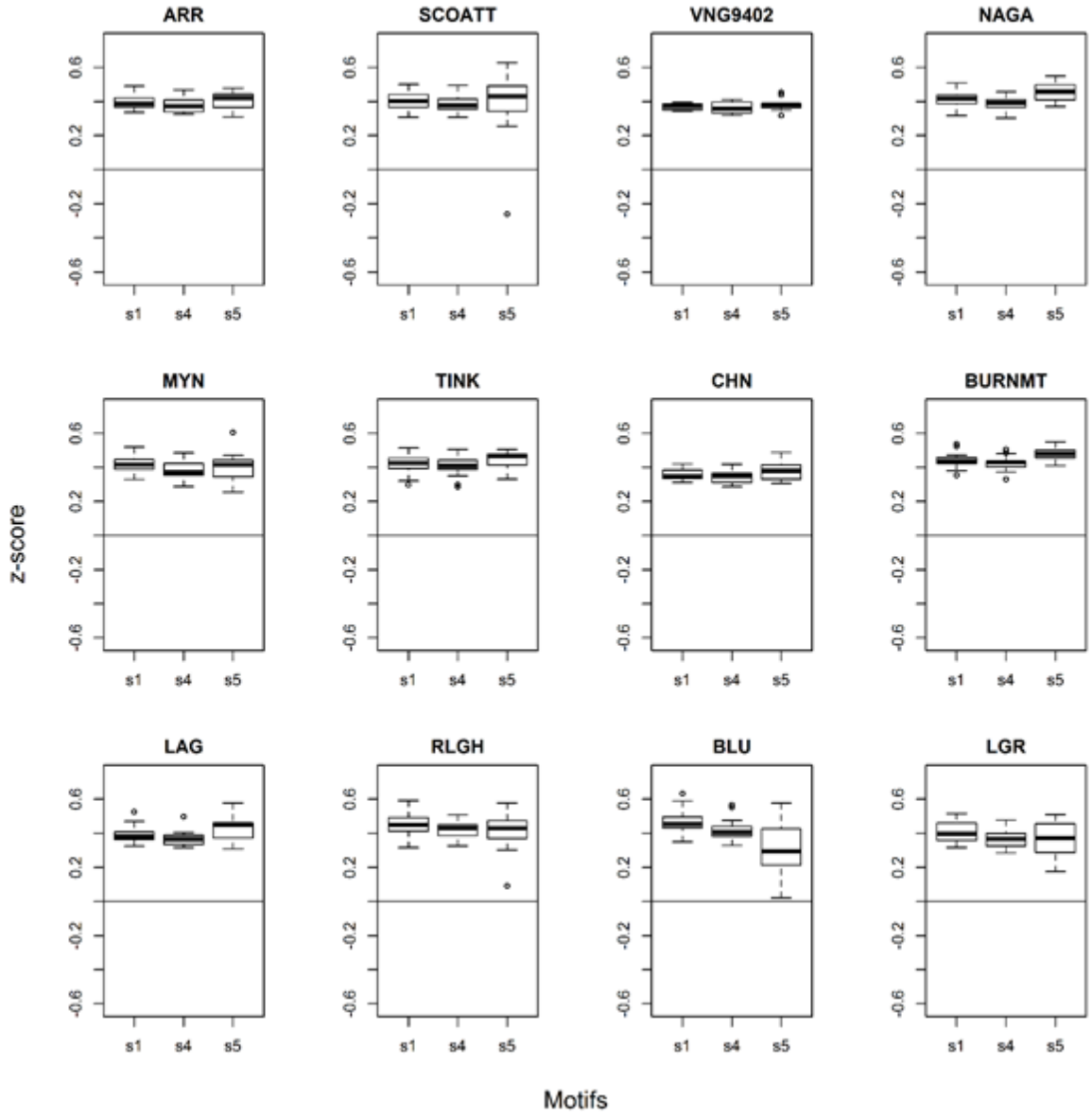



**Figure 0.L**. The z-scores for the occurrence of three three-node motifs at each lake site. s1 = tri-trophic chain, s4 = apparent competition, s5 = direct competition. Boxplots represent the median and interquartile range of the z-scores, one z-scores for each food web over the course on monitoring.

## A.4 Discussion

Contrary to our hypothesis, in general food webs were found to be less stable under acidified conditions. Acidified food webs were found to have greater link turnover (stream food webs only), smaller cores, were less robust to simulated species removal and had more even distribution of biomass flows. Those sites which had changed their hydrochemistry the most over the monitoring period also experienced the greatest variability in species relative abundance. The only stability measure which increased with acidity was the density of connections within the core (Rich-club coefficient), here acidified food webs had a more densely connected core, which may have a stabilising effect. There was no evidence of a relationship between the occurrence of stabilising motifs and the hydrochemical stress of a food web.

These results contrast with previous work (Lundberg *et al.* 2000; Ledger & Hildrew 2005; Kernan *et al.* 2010; Layer *et al.* 2010). It is possible that acidified and non-acidified food webs show different characteristics of stability, for instance acidified food webs might be more persistent and resist the invasion of acid sensitive species, while at the same time be less robust to species loss. The persistence and resilience of these food webs, would be better understood using controlled experiments, where the perturbation strength (and type, i.e. press or pulse perturbation) can be manipulated, and the persistence and recovery time could be measured. Additionally, although the metrics used here should be independent of network size, as they are measured proportional to total network size, it would be useful to demonstrate this conclusively using null-model simulations.

## A.5 References

Borgatti, S.P. & Everett, M.G. (2000). Models of core /periphery structures. *Social networks*, **21**, 375–395.

Borrelli, J.J. (2015). Selection against instability: stable subgraphs are most frequent in empirical food webs. *Oikos*, **124**, 1583–1588.

Brede, M. (2010). Coordinated and uncoordinated optimization of networks. *Physical Review E*, **81**, 066104.

Brose, U., Jonsson, T., Berlow, E.L., Warren, P., Banasek-Richter, C., Bersier, L.F., Blanchard, J.L., Brey, T., Carpenter, S.R., Blandenier, M.F.C., Cushing, L., Dawah, H.A., Dell, T., Edwards, F., Harper-Smith, S., Jacob, U., Ledger, M.E., Martinez, N.D., Memmott, J., Mintenbeck, K., Pinnegar, J.K., Rall, B.C., Rayner, T.S., Reuman, D.C., Ruess, L., Ulrich, W., Williams, R.J., Woodward, G. & Cohen, J.E. (2006). Consumer-resource body-size relationships in natural food webs. *Ecology*, **87**, 2411–2417.

Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M. & West, G.B. (2004). Toward a metabolic theory of ecology. *Ecology*, **85**, 1771–1789.

Csete, M. & Doyle, J. (2004). Bow ties, metabolism and disease. *Trends in biotechnology*, **22**, 446–50.

Cyr, H., Peters, R.H. & Downing, J.A. (1977). Population density and community size structure: comparison of aquatic and terrestrial systems. *Oikos*, **80**, 139–149.

Derényi, I., Farkas, I., Palla, G. & Vicsek, T. (2004). Topological phase transitions of random networks. *Physica A: Statistical Mechanics and its Applications*, **334**, 583–590.

Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002). Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, **5**, 558–567.

Elton, C.S. (1958). *The ecology of invasions by animals and plants*. Methuen and Co. Ltd., London, UK.

Gilljam, D., Thierry, A., Edwards, F.K., Figueroa, D., Ibbotson, A.T., Jones, J.I., Lauridsen, R.B., Petchey, O.L., Woodward, G. & Ebenman, B. (2011). Seeing Double: Size-Based and Taxonomic Views of Food Web Structure. *Advances in Ecological Research*, **45**, 67–133.

Kernan, M., Battarbee, R.W., Curtis, C., Monteith, D.T. & Shillands, E.M. (2010). *Recovery of lakes and streams in the UK from acid rain. The United Kingdom Acid Waters*

*Monitoring Network 20 year interpretative report*. Report to the Department for Environment, Food and Rural Affairs.

Kokkoris, G.D., Jansen, V.A.A., Loreau, M. & Troumbis, A.Y. (2002). Variability in interaction strength and implications for biodiversity. *Journal of Animal Ecology*, **71**, 362–371.

Layer, K., Riede, J.O., Hildrew, A.G. & Woodward, G. (2010). Food Web Structure and Stability in 20 Streams Across a Wide pH Gradient. *Advances in Ecological Research*, **42**, 265–299.

Leaper, R. & Raffaelli, D. (1999). Defining the abundance body-size constraint space: Data from a real food web. *Ecology Letters*, **2**, 191–199.

Ledger, M.E., Brown, L.E., Edwards, F.K., Milner, A.M. & Woodward, G. (2012). Drought alters the structure and functioning of complex food webs. *Nature Climate Change*, **3**, 223–227.

Ledger, M.E. & Hildrew, A.G. (2005). The ecology of acidification and recovery: changes in herbivore-algal food web linkages across a stream pH gradient. *Environmental Pollution*, **137**, 103–118.

Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. (2011). Controllability of complex networks. *Nature*, **473**, 167–73.

Lundberg, P., Ranta, E. & Kaitala, V. (2000). Species loss leads to community closure. *Ecology Letters*, **3**, 465–468.

Ma, A. & Mondragón, R.J. (2015). Rich-Cores in Networks. *Plos One*, **10**, e0119678.

MacArthur, R. (1955). Fluctuations of Animal Populations and a Measure of Community Stability. *Ecology*, **36**, 533–536.

May, R.M. (1972). Will a large complex system be stable? *Nature*, **238**, 413–414.

McCann, K.S. (2000). The diversity-stability debate. *Nature*, **405**, 228–233.

Mellin, C., Huchery, C., Caley, M.J., Meekan, M.G. & Bradshaw, C.J.A. (2010). Reef size and isolation determine the temporal stability of coral reef fish populations. *Ecology*, **91**, 3138–3145.

Neutel, A.-M., Heesterbeek, J. a P., van de Koppel, J., Hoenderboom, G., Vos, A., Kaldeway, C., Berendse, F. & de Ruiter, P.C. (2007). Reconciling complexity with stability in naturally assembling food webs. *Nature*, **449**, 599–602.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H. & Wagner, H. (2012). vegan: Community Ecology Package. *R package version*, **1**, R package version 2.0–4.

Pawar, S., Dell, A.I. & Savage, V.M. (2012). Dimensionality of consumer search space

drives trophic interaction strengths. *Nature*, **486**, 485–9.

Peters, R.H. (1986). *The ecological implications of body size*.

Pimm, S.L. (1984). The complexity and stability of ecosystems. *Nature*, **307**, 321–326.

Rall, B.C., Brose, U., Hartvig, M., Kalinkat, G., Schwarzmüller, F., Vucic-Pestic, O., Petchey, O.L., Schwarzmuller, F., Vucic-Pestic, O. & Petchey, O.L. (2012). Universal temperature and body-mass scaling of feeding rates. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **367**, 2923–34.

Reuman, D.C., Mulder, C., Banašek-Richter, C., Cattin Blandenier, M.F., Breure, A.M., Hollander, H. Den, Kneitel, J.M., Raffaelli, D., Woodward, G., Cohen, J.E., Banasek-Richter, C., Blandenier, M.F.C., Breure, A.M., Den Hollander, H., Kneitel, J.M., Raffaelli, D., Woodward, G. & Cohen, J.E. (2009). Allometry of Body Size and Abundance in 166 Food Webs (H. Caswell, Ed.). *Advances in Ecological Research*, **41**, 1–44.

Reuman, D.C., Mulder, C., Raffaelli, D. & Cohen, J.E. (2008). Three allometric relations of population density to body mass: theoretical integration and empirical tests in 149 food webs. *Ecology Letters*, **11**, 1216–1228.

Strona, G., Nappo, D., Boccacci, F., Fattorini, S. & San-Miguel-Ayanz, J. (2014). A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications*, **5**, 4114.

Tang, S., Pawar, S. & Allesina, S. (2014). Correlation between interaction strengths drives stability in large ecological networks. *Ecology Letters*, **17**, 1094–1100.

Ulanowicz, R.E. (2004). Quantitative methods for ecological network analysis. *Computational Biology and Chemistry*, **28**, 321–339.

Yodzis, P. (1981). The stability of real ecosystems. *Nature*, **289**, 674–676.

Yodzis, P. & Innes, S. (1992). Body Size and Consumer-Resource Dynamics. *The American Naturalist*, **139**, 1151–1175.

Zhou, S. & Mondragon, R.J. (2004). The rich-club phenomenon in the Internet topology. *Ieee Communications Letters*, **8**, 180–182.

# Appendix B | Drought rewires the cores of food webs[1]

[1] Lu, X., **Gray, C**., Brown, L.E., Ledger, M.E., Milner, A.M., Mondragón, R.J., Woodward, G. & Ma, A. (in review) Drought rewires the cores of food webs. *Nature Climate Change*.

## B.1 Summary

Droughts are intensifying across the globe (Hartmann *et al.* 2013; Kendon *et al.* 2013), with potentially devastating implications for fresh water ecosystems (Milly *et al.* 2005; Vorosmarty *et al.* 2010). We used novel network science approaches to investigate drought impacts on stream food webs and explored potential consequences for web robustness to future perturbations. The substructure of the webs was characterised by a core of richly-connected species (Ma & Mondragón 2015) surrounded by poorly-connected peripheral species. Drought caused the partial collapse of the food webs (Ledger *et al.* 2012) but loss of the most extinction-prone peripheral species triggered a substantial rewiring of interactions within the food web cores. These shifts in species interactions in the core conserved the underlying core/periphery substructure and stability of the drought-impacted webs. When we perturbed the webs by simulating species loss *in silico*, the rewired drought webs exhibited comparable robustness to the larger, undisturbed webs. Our research unearths previously unknown compensatory dynamics arising from within the core that can underpin food web stability in the face of environmental perturbations.

## B.2 Main text

Many areas of the world are becoming more prone to drought (Hartmann *et al.* 2013; Kendon *et al.* 2013) and declining precipitation coupled with increasing demand for water could threaten the integrity of freshwater ecosystems across the globe (Milly *et al.* 2005; Vorosmarty *et al.* 2010). In rivers and streams, the elimination of sensitive species could potentially undermine food web structure and functioning (Closs & Lake 1994; Lake 2003; Lytle & Poff 2004), yet how this affects their stability - at both substructural and whole-network levels (Woodward *et al.* 2012) has yet to be fully elucidated. Responses to climate change are frequently interpreted autecologically, through analysis of individual species traits (McKee & Atkinson 2000) but these provide no information on alterations of species functional attributes and conceal potential compensatory behavioural mechanisms, such as resource switching. Synecological approaches that can address changing species interactions in the context of the whole food web (Tylianakis *et al.* 2007; Petchey *et al.* 2010; Woodward *et al.* 2010), and hence the potential trophic mechanisms behind community-level responses (Ebenman & Jonsson 2005; Borrvall & Ebenman 2008), remain scarce. In addition, there are non-random substructures in food webs which could underpin their responses to climate-induced perturbations (Garlaschelli *et al.* 2003). Emerging network science has linked the presence of a cohesive "core" of closely interacting nodes and a loosely connected "periphery"(Borgatti & Everett 2000; Csete & Doyle 2004; Csermely *et al.* 2013; Ma & Mondragón 2015) to the stability of complex (non-ecological) networks (Derényi *et al.* 2004; Brede 2010). The significance of this for food-web responses to an environmental perturbation - drought - is reported here for the first time.

The network "core" is a cohesive group of highly connected nodes that governs the functional attributes of a wide range of complex systems (Borgatti & Everett 2000). It determines system robustness because densely intertwined pathways within the substructure can provide redundancy by buffering external fluctuations (Borgatti &

Everett 2000; Csermely *et al.* 2013) without altering functioning (Kitano 2004); such structures are absent from less robust, regular small-world networks (Thompson *et al.* 2012). Core-size relative to the rest of the web indicates a network's state (Derényi *et al.* 2004; Csete & Doyle 2004; Brede 2010): large cores provide greater scope for redundancy of links and rewiring in the event of node and link failure, whilst small cores indicate vulnerability and systems being under stress.

Here, we quantify experimentally how drought disturbance influences stream food web substructure and model how this then determines robustness to future perturbations. We analysed food webs from a stream mesocosm experiment in which benthic communities subjected to a drought treatment for two years were compared with undisturbed controls (four replicates; eight food webs in total; see Methods). Food webs were constructed from gut contents analysis of all 3,643 individuals collected at the end of the experiment. These exceptionally well-resolved webs encompassed 783 pairwise trophic interactions among 74 trophic elements, consisting of detrital resources, primary producers and a taxonomically diverse array of invertebrate consumers (Table S1). Ecological communities consist of coexisting taxa and species extinction can trigger rippling effects due to their interdependency; as a result, community fragility to disturbance can be influenced by structural properties, such as the distribution of trophic interactions (Ebenman & Jonsson 2005; Borrvall & Ebenman 2008). We hypothesised that the food webs were governed by a core/periphery structure, as detected recently in a range of non-ecological networks (Csete & Doyle 2004; Csermely *et al.* 2013; Ma & Mondragón 2015). Highly connected core species are functionally important because they provide alternative routes for the flux of matter and may therefore buffer the effects of perturbations and enhance stability in food webs. Peripheral species are less integral to the ecosystem in topological sense, and changes in the food web composition and configuration will likely lead to isolation (i.e. extinction) of these species, similar to previous observations in mutualistic networks (Burgos *et al.* 2007). Specialist consumers from

the web periphery will be especially vulnerable to extinction because they are more loosely connected and dependent on fewer resource species. Redundancy in the links within the core could, in theory, provide a means of withstanding the effect of species loss and rebalancing the structure of food webs, thereby conserving overall robustness.

To test our hypotheses, we applied a novel graph profiling technique (Ma & Mondragón 2015) to characterise the cores of our eight highly-resolved food webs (Ledger *et al.* 2011a; Woodward *et al.* 2012). To generate a graph profile for a web, nodes were ranked by their degree (number of links). Starting from the highest degree node, we examined the interconnectedness among the high degree nodes as those of a lower rank were included sequentially. A point is reached whereby the connectivity among the high degree nodes peaks, reflecting the cohesiveness in the core and defining the core boundary followed by generally decreasing connectedness thereafter. The rest of the nodes form the periphery, which is only loosely connected to the core, and contains few or no links among its constituents. We then measured the density of interactions within the core and across the web using the "rich-club" coefficient (Zhou & Mondragon 2004). To gauge the level of organisation in the core/periphery structure between the drought and control treatments, we employed an ensemble of null networks, whereby links were reshuffled randomly while conserving network properties (Maslov *et al.* 2004). Graph profiles obtained from the null models represent network structures that would simply happen by chance, and they were used to benchmark the link patterns of the empirical webs. The further an empirical web deviates from its null models (i.e. a z-score greater or less than 0), the more significant, in statistical terms, its link patterns, which also indicates the level of organisation that has taken place to generate the observed pattern. To examine the effectiveness of the compensatory mechanism provided by the core, we studied network robustness by measuring the rate at which the structural integrity of food webs collapsed (Dunne *et al.* 2002) under two commonly simulated species removal

scenarios: i) random removal and ii) targeted removal of core species (i.e. high degree species).

All eight food webs exhibited a clear core/periphery structure (Figure B.A), here evidenced by a distinct peak in their core profiles and a step-change in interconnectedness from high to low degree species (indicated by a vertical line in Figure B.A, at which the number of links $k_r^+$ is at its maximum, and after which it decreases steadily). The food web cores contained species from all trophic levels (Figure B.A) and accounted for (on average) 50% of the species. The proportion of core species was unchanged by drought (t-test on arcsine transformed proportion data, *d.f.=3, p=0.16*; Table B.1), despite absolute species losses of 25%. Core size was large relative to non-ecological networks (5-30% of total network size (Csermely *et al.* 2013; Ma & Mondragón 2015)), indicating that natural systems may possess far greater linkage redundancy. Species extinction was greatest in the periphery (one tailed t-test on arcsine transformed proportion data, *d.f.=3, p=0.01;* Table B.1), and as expected, species that fell into this category were mainly invertebrate consumers high in the food chain which lost all their resources. Drought caused more species in the core to migrate into the periphery of the web via a reshuffling of interactions, than *vice versa* (one tailed t-test test on arcsine transformed proportion data, *d.f.=3, p=0.01*, Table B.1 and Figure B.B). Despite this drought-induced realignment of species, the preservation of the core/periphery structure (Figure B.B) and its relative size is suggestive of underlying inertia within the webs' substructure.

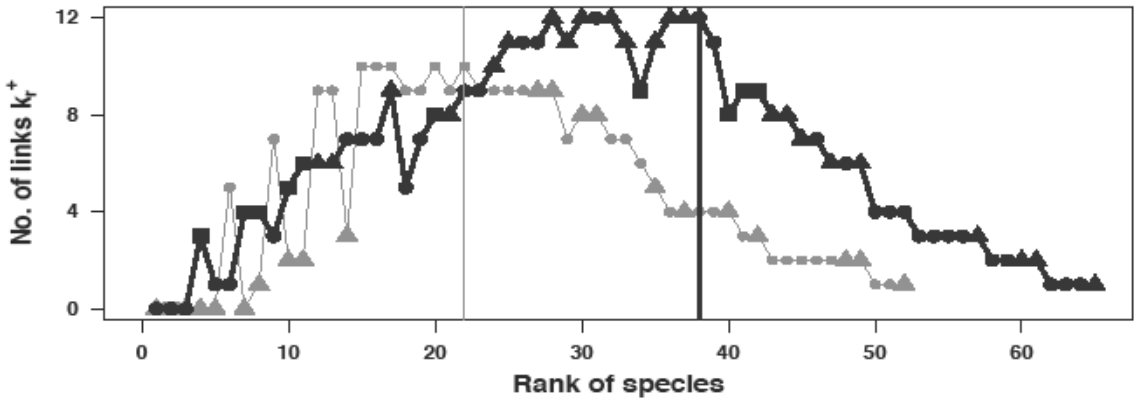



**Figure B.A. Core/periphery structure of control and drought food webs.** Comparisons of one block of control and drought core profiles. Nodes are ranked by their decreasing order of degree and plotted by the number of links with nodes of a higher rank, $k_r^+$. The control web (dark thick line) is plotted alongside its respective drought web (light thin line). Species were classified as *Basal* (circles), *Intermediate* (squares) or *Top* (triangles). The maximum of the curve $k_{r*}^+$, defines the boundary of the core for the control (dark thick line) and drought (light thin line) webs.

**Table B.1. Statistics from two independent samples t-tests.** The effects of drought on the relative core size and robustness were tested using one-tailed t-test on arcsine transformed data. Two-tailed t-test on arcsine transformed data was applied to examine if peripheral species are more likely than core species to go extinct, and if more core species than periphery species realigned after drought. Significant differences are indicated in bold.

| | *df* | *p* | | *df* | *p* |
|---|---|---|---|---|---|
| Relative core size | 3 | 0.16 | More extinction from periphery | 3 | **0.01** |
| Robustness (random) | 3 | 0.89 | More species realigned from core | 3 | **0.01** |
| Robustness (targeted) | 3 | 0.17 | | | |

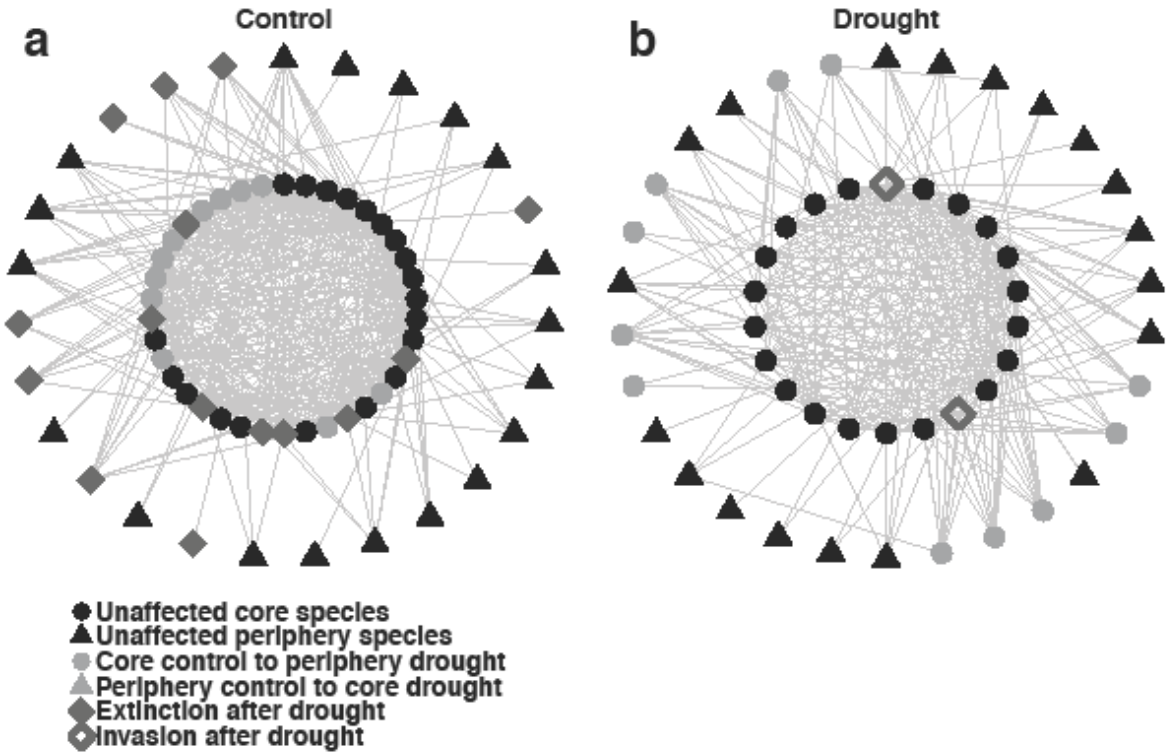



**Figure B.B. Drought caused species re-alignment in substructures.** Comparisons of one block of control (**a**) and drought (**b**) food web structures. Core species in the inner ring are surrounded by periphery species in the outer ring. In this web pair, drought caused 15 species to go extinct (filled diamonds) and 11 core species to shift to the periphery (light circles).

Drought reduced the density of connections within the core (Figure B.Ca), as shown by lower rich-club coefficients, $\phi_r$. This phenomenon in non-ecological networks is a common response to stress (Derényi *et al.* 2004; Brede 2010), and in our case was a result of compensatory re-wiring as core species moved into the periphery: the density of connections in the periphery was unaffected by drought despite peripheral species loss. All webs showed a marked deviation in connectivity from their respective null models within their cores, revealing a systematic, non-random substructure (Figure B.Cb). Drought resulted in a greater decrease in the *z*-score within the core: *i.e.*, link density inside was significantly lower than what would be expected by chance, suggesting even more intense organisation had taken place in response to the drought. This pronounced change in the core supports our hypothesis

about its governing role in the re-structuring of food webs under this stressor.

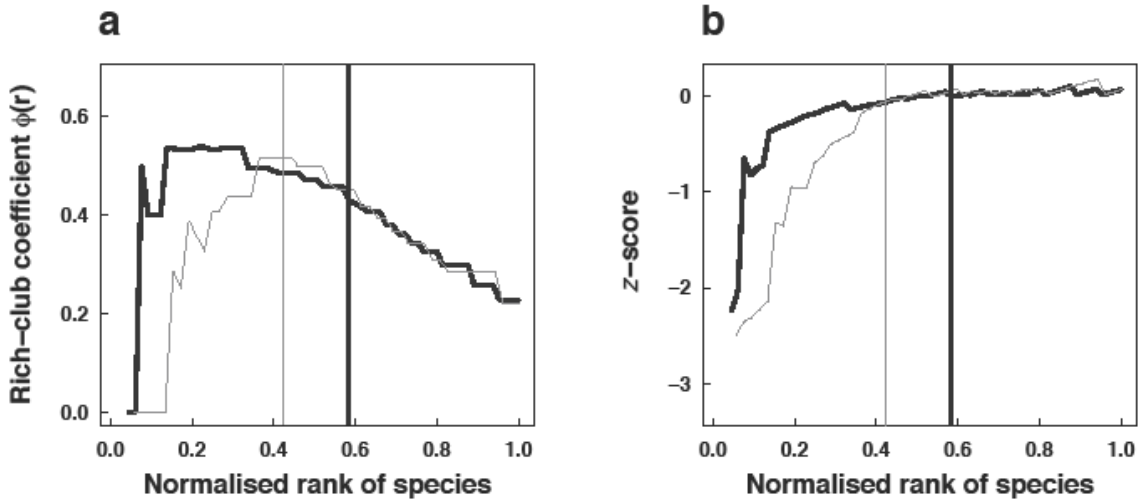



**Figure B.C. Drought reduced the link density in the core and caused further restructuring in the core.** (**a**) The density of connections across the network measured by the rich-club coefficient, $\phi_r$, is shown for one block of control (dark thick line) and drought-disturbed (light thin line) mesocosms. Nodes were ordered by their degree which were then *normalised* by the size of the network. Boundaries of the cores are marked by vertical lines as in **Figure B.A** (**b**) Comparisons of the web pair's deviance in connection density from their respective null models and more negative z-scores indicate greater deviance from the null model.

Food webs were robust to simulated random species removal, and this was unaffected by drought: the amount of primary extinction required for 50% species loss was comparable in both treatments (t-test on arcsine transformed proportion data, *d.f.=3, p=0.89;* Table B.1). This can be explained by the conservation of the overall core/periphery structure and relative core size. As the loss of peripheral species would have skewed the probability of a core species being chosen under random removal in drought webs, the realignments of species from the core to the periphery rebalanced the overall network structure, conferring the same degree of resistance towards these perturbations. When more highly connected species were removed first, drought webs were as robust to species removal as control webs (t-test on arcsine

transformed proportion data, *d.f.=3, p=0.17;* Table B.1). This suggests that although the density of connections within the core was altered by drought, overall network integrity and ability to withstand further perturbations was conserved by species re-alignment. It is conceivable that a threshold core connectance may exist, beyond which this redundancy is lost and the associated food web collapses, echoing ideas suggested by Dunne *et al.* (2002) and Krause *et al.* (2003). Identifying this threshold would allow us to better predict which communities are most at risk from environmental change.

Our results demonstrate that drought disturbance triggered previously unknown substructural changes within real food webs, beyond the direct and obvious species losses that have been reported elsewhere when based on fixed autecological traits (Ledger *et al.* 2012; Woodward *et al.* 2012). While the underlying core/periphery structure was robust to perturbations, the composition and configuration of the food web substructures changed markedly, with a steep reduction in interactions among the remaining core species. The ability to predict which networks of species interactions are most vulnerable to anthropogenic pressures, and the identification of a core of species vital to the functioning and persistence of a community within an ecosystem, would greatly enhance our ability to direct conservation efforts more effectively in the face of environmental perturbations (Ebenman & Jonsson 2005; Borrvall & Ebenman 2008). Traditional network metrics were far less sensitive (Ledger *et al.* 2012) than the novel measures applied in this study, and therefore less useful for gauging changes in food webs exposed to perturbations. Substructural approaches that capture the plastic synecological traits defined by species interactions can help to unearth compensatory shifts within ecological networks, and provide us with a major new way to detect and understand the effects of environmental change on ecological communities.

## B.3 Methods

### B.3.1 Experimental design.

Details of the experimental design and methods used to build the food webs are published elsewhere (Woodward *et al.* 2012; Ledger *et al.* 2013). To summarise the experiment ran for two years (March 2000-February 2002) in outdoor stream mesocosms that consisted of four pairs of channels subjected to either control or drought conditions. All channels were subject to two months of constant flow before a drought treatment (6 days of dewatering per month) was applied to one channel per pair. During the simulated drying periods, surface flows ceased and drying of exposed substrata occurred in patches, whereas the interstices beneath the bed surface remained wet, and small pools persisted at intervals along the length of the dewatered channels (Lancaster & Ledger 2015). Surfaces of exposed substrata dried at natural ambient rates such that the stress experienced by organisms stranded in the mesocosms was consistent with those in adjacent drying stream reaches (Harris 2006). This experimental design simulated periodic drying events occurring during a supra-seasonal drought. Stream drying events have occurred during major droughts in Europe (Parry *et al.* 2012) and are expected to increase in frequency with climate change (Beniston *et al.* 2007). As with all mesocosm experiments, our design necessitated some trade-off between realism and replication (Ledger *et al.* 2008, 2011b). The simulated flows may adequately capture the expected changes in the magnitude and frequency of drying in rivers under climate change but do not necessarily reflect the expected changes in seasonality of these events. At the end of the experiment all invertebrates were collected and identified and gut content analysis was performed: all individuals and their gut contents were identified to genus or species level, where possible. The resultant eight food webs are among the most highly resolved to date, comprising 783 pairwise trophic interactions and 74 trophic elements in the aggregate web. Comparison of the control channel food webs to data collected for 82 'natural' river food webs showed the mesocosm channels contained

realistic webs, with consistent and similar size structures suggesting that patterns of energy flux between mesocosm consumers and resources were good analogues of those in natural systems (Brown *et al.* 2011). Species were categorised into three trophic levels: Basal (B), Intermediate (I) and Top (T). A basal species was defined as a species with no prey; a top-level species was referred to as a species with no predators; and the rest were defined as intermediate species.

### B.3.2 Food web profiling.

The core profiling method identifies a substructure of highly interconnected species by ordering species with respect to the number of connections to other species and the extent to which those connections link to more highly connected species in the web (Ma & Mondragón 2015). Highly interconnected species constitute the web core, with less-connected nodes forming the periphery. Each food web was represented as a binary and undirected network with $S$ nodes (species) and $E$ links (the interaction between species). To obtain a core profile, nodes were ordered in descending order of their degree (i.e. number of links) and a node with a rank $r$ has degree $k_r$. The number of links that a node shares with nodes of a higher rank is $k_r^+$ and the number of links with nodes of a lower rank is therefore $k_r - k_r^+$. Starting with the node with the highest rank, the value of $k_r^+$ fluctuates as nodes from further down the rank are being included. There will be a point $r^*$ where $k_r^+$ reaches its maximum and will always be less than $k_{r*}^+$ thereafter, marking the boundary of the core. To quantify the density of links inside the core, the rich-club coefficient (Zhou & Mondragon 2004) was calculated, which is defined as:

$$\phi_r = \frac{2}{r(r-1)} \sum_{i=1}^{r} k_i^+ = \frac{2E_r}{r(r-1)}$$

where $E_r$ is the number of links shared by the highest ranked r nodes and $r(r-1)/2$ is the maximum number of possible links among these nodes. The connectivity of a core is given by $\phi_{r*}$ whereby a fully connected core has a value of $\phi_{r*}$= 1 and a fully disconnected core gives $\phi_{r*}$= 0. Given that drought webs contain fewer species than

their control counterparts, results could have been skewed by their reduced web size if their absolute values were used: to overcome this the species rank was normalised by the overall web size.

### B.3.3 Null model.

A statistical null model was used to determine the probability of the connectivity observed in the empirical data. For each empirical food web, we applied a randomisation method (Maslov *et al.* 2004) to generate an ensemble of 100 networks by randomly reshuffling the links while conserving the properties of the empirical network, including the number of nodes, the number of links and the degree distribution. This allows us to assess the statistical significance of the patterns of interactions observed in the empirical webs with respect to patterns that would simply occur by chance. To quantify how the link density in the core differs from the random networks, we first referred the rich-club coefficient of the empirical food web and compared that to its null counterpart by calculating the z-score. A $z$-score of 0 means that the empirical data exhibits an organisation of links that is the same as what you would expect from a random case; a value > 0 means that the empirical has a higher than expected density of links, and vice-versa. This effectively describes the degree of organisation of species interactions in the sense that the more improbable a configuration of links is, the more organisation is required to be in place to attain the observed pattern. Again, the rank of species was normalised to compensate for the effect of different web sizes when comparing the control and drought food web pairs.

### B.3.4 Network robustness.

To assess this, we simulated primary species loss in all the food webs by manually removing species (Dunne *et al.* 2002). Firstly, species were chosen randomly and removed from the food web, together with all their associated links, in an iterative manner. We recorded the total species at each step, which accounts for both primary loss and secondary extinction (as a result of species isolation from resource). Robustness was quantified by the amount of primary extinction required for a total

loss of 50% of the species. We repeated this for 100 times for each web and results were averaged. Secondly, species were removed in the descending order of degree which is often considered as the worst case scenario as the most important (connected) nodes are being targeted. Similarly, species were removed in an iterative manner, but the degree order of nodes was re-calculated after each species removal as removing a node and its links may impact on the degree order among the rest of the nodes. Again, robustness was evaluated by the total primary extinction required for a cumulative 50% species loss.

## B.4 References

Beniston, M., Stephenson, D.B., Christensen, O.B., Ferro, C.A.T., Frei, C., Goyette, S., Halsnaes, K., Holt, T., Jylhä, K., Koffi, B., Palutikof, J., Schöll, R., Semmler, T. & Woth, K. (2007). Future extreme events in European climate: An exploration of regional climate model projections. *Climatic Change*, **81**, 71–95.

Borgatti, S.P. & Everett, M.G. (2000). Models of core /periphery structures. *Social networks*, **21**, 375–395.

Borrvall, C. & Ebenman, B. (2008). Biodiversity and persistence of ecological communities in variable environments. *Ecological Complexity*, **5**, 99–105.

Brede, M. (2010). Coordinated and uncoordinated optimization of networks. *Physical Review E*, **81**, 066104.

Brown, L.E., Edwards, F.K., Milner, A.M., Woodward, G. & Ledger, M.E. (2011). Food web complexity and allometric scaling relationships in stream mesocosms: implications for experimentation. *Journal of Animal Ecology*, **80**, 884–895.

Burgos, E., Ceva, H., Perazzo, R.P.J., Devoto, M., Medan, D., Zimmermann, M. & María Delbue, A. (2007). Why nestedness in mutualistic networks? *Journal of Theoretical Biology*, **249**, 307–313.

Closs, G.P. & Lake, P.S. (1994). Spatial and temporal variation in the structure of an intermittent-stream food web. *Ecological Monographs*, **64**, 1–21.

Csermely, P., London, a., Wu, L.-Y. & Uzzi, B. (2013). Structure and dynamics of core/periphery networks. *Journal of Complex Networks*, **1**, 93–123.

Csete, M. & Doyle, J. (2004). Bow ties, metabolism and disease. *Trends in biotechnology*,

**22**, 446–50.

Derényi, I., Farkas, I., Palla, G. & Vicsek, T. (2004). Topological phase transitions of random networks. *Physica A: Statistical Mechanics and its Applications*, **334**, 583–590.

Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002). Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters*, **5**, 558–567.

Ebenman, B. & Jonsson, T. (2005). Using community viability analysis to identify fragile systems and keystone species. *Trends in Ecology & Evolution*, **20**, 568–575.

Garlaschelli, D., Caldarelli, G. & Pietronero, L. (2003). Universal scaling relations in food webs. *Nature*, **423**, 165–8.

Harris, R.M.L. (2006). The effect of experimental drought disturbance on macroinvertebrate assemblages in stream mesocosms. *PhD thesis, University of Birmingham, UK*.

Hartmann, D.L., Tank, A.M.G.K. & Rusticucci, M. (2013). *IPCC Fifth Assessment Report, Climatie Change 2013: The Physical Science Basis*.

Kendon, M., Marsh, T. & Parry, S. (2013). The 2010-2012 drought in England and Wales. *Weather*, **68**, 88–95.

Kitano, H. (2004). Biological robustness. *Nature reviews. Genetics*, **5**, 826–37.

Krause, A.E., Frank, K.A., Mason, D.M., Ulanowicz, R.E. & Taylor, W.W. (2003). Compartments revealed in food-web structure. *Nature*, **426**, 282–285.

Lake, P.S. (2003). Ecological effects of perturbation by drought in flowing waters. *Freshwater Biology*, **48**, 1161–1172.

Lancaster, J. & Ledger, M.E. (2015). Population-level responses of stream macroinvertebrates to drying can be density-independent or density-dependent. *Freshwater Biology*, In press.

Ledger, M.E., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N. & Woodward, G. (2013). Extreme Climatic Events Alter Aquatic Food Webs. A Synthesis of Evidence from a Mesocosm Drought Experiment. *Advances in Ecological Research*, **48**, 343–395.

Ledger, M.E., Brown, L.E., Edwards, F.K., Milner, A.M. & Woodward, G. (2012). Drought alters the structure and functioning of complex food webs. *Nature Climate Change*, **3**, 223–227.

Ledger, M.E., Edwards, F.K., Brown, L.E., Milner, A.M. & Woodward, G. (2011a). Impact of simulated drought on ecosystem biomass production: an experimental

test in stream mesocosms. *Global Change Biology*, **17**, 2288–2297.

Ledger, M.E., Edwards, F.K., Brown, L.E., Milner, A.M. & Woodward, G. (2011b). Impact of simulated drought on ecosystem biomass production: An experimental test in stream mesocosms. *Global Change Biology*, **17**, 2288–2297.

Ledger, M.E., Harris, R.M.L., Armitage, P.D. & Milner, A.M. (2008). Disturbance frequency influences patch dynamics in stream benthic algal communities. *Oecologia*, **155**, 809–819.

Lytle, D. a & Poff, N.L. (2004). Adaptation to natural flow regimes. *Trends in Ecology and Evolution*, **19**, 94–100.

Ma, A. & Mondragón, R.J. (2015). Rich-Cores in Networks. *Plos One*, **10**, e0119678.

Maslov, S., Sneppen, K. & Zaliznyak, A. (2004). Detection of topological patterns in complex networks: correlation profile of the internet. *Physica A: Statistical Mechanics and its Applications*, **333**, 529–540.

McKee, D. & Atkinson, D. (2000). The influence of climate change scenarios on populations of the mayfly Cloeon dipterum. *Hydrobiologia*, **441**, 55–62.

Milly, P.C.D., Dunne, K. a & Vecchia, a V. (2005). Global pattern of trends in streamflow and water availability in a changing climate. *Nature*, **438**, 347–50.

Parry, S., Hannaford, J., Lloyd-Hughes, B. & Prudhomme, C. (2012). Multi-year droughts in Europe: analysis of development and causes. *Hydrology Research*, **43**, 689–706.

Petchey, O.L., Brose, U. & Rall, B.C. (2010). Predicting the effects of temperature on food web connectance. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **365**, 2081–2091.

Thompson, R.M., Brose, U., Dunne, J.A., Hall, R.O., Hladyz, S., Kitching, R.L., Martinez, N.D., Rantala, H., Romanuk, T.N., Stouffer, D.B. & Tylianakis, J.M. (2012). Food webs: reconciling the structure and function of biodiversity. *Trends in Ecology and Evolution*, **27**, 689–697.

Tylianakis, J.M., Tscharntke, T. & Lewis, O.T. (2007). Habitat modification alters the structure of tropical host–parasitoid food webs. *Nature*, **445**, 202–205.

Vorosmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R. & Davies, P.M. (2010). Global threats to human water security and river biodiversity. *Nature*, **468**, 334.

Woodward, G., Brown, L.E.L.E., Edwards, F.K.F.K., Hudson, L.N.L.N., Milner, a. M.A.M., Reuman, D.C.D.C. & Ledger, M.E.M.E. (2012). Climate change impacts in multispecies systems: drought alters food web size structure in a field

experiment. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **367**, 2990–2997.

Woodward, G., Perkins, D.M. & Brown, L.E. (2010). Climate change and freshwater ecosystems: impacts across multiple levels of organization. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **365**, 2093–2106.

Zhou, S. & Mondragon, R.J. (2004). The rich-club phenomenon in the Internet topology. *Ieee Communications Letters*, **8**, 180–182.

# Appendix C | Gene-to-ecosystem impacts of a catastrophic pesticide spill: testing a multilevel bioassessment approach in a river ecosystem[1]

[1] Thompson, M.S.A., Bankier, C., Bell, T., Dumbrell, A.J., **Gray, C**., Ledger, M.E., Lehman, K., McKew, B.A., Sayer, C.D., Shelley, F., Trimmer, M., Warren, S.L. & Woodward, G. (2015). Gene-to-ecosystem impacts of a catastrophic pesticide spill: testing a multilevel bioassessment approach in a river ecosystem. *Freshwater Biology*.

## C.1 Summary

1. Pesticides can have strong deleterious impacts in fresh waters, but understanding how these effects cascade through natural ecosystems, from microbes to apex predators, is limited because research that spans multiple levels of biological organisation is rare.

2. We report how an accidental insecticide spill altered the structure and functioning of a river across levels ranging from genes to ecosystems. We quantified the impacts on assemblages of microbes, diatoms, invertebrates and fish and measured leaf-litter decomposition rates and microbial functional potential at upstream control and downstream impacted sites two months after the spill.

3. Both direct and indirect impacts were evident across multiple levels of organisation and taxa, from the base of the food web to higher trophic levels. At the molecular level, differences in functional gene abundance within the impacted sites reflected a combination of direct and indirect effects of the pesticide, via elevated microbial populations capable of utilising chlorpyrifos as a resource (i.e. direct effect) and oxidising ammonia released by decaying invertebrate carcasses (i.e. indirect effect).

4. At the base of the food chains, diatom taxa found only in the impacted sites were an order-of-magnitude larger in cell-size than the largest comparable taxa in control communities, following the near-extirpation of their consumers. Population biomass of the key detritivore *Gammarus pulex* was markedly lower, as was the rate of litter decomposition in the impacted sites. This was partially compensated for, however, by elevated microbial breakdown, suggesting another indirect food-web effect of the toxic spill.

5. Although many species exhibited population crashes or local extirpation, total macroinvertebrate biomass and abundance were largely unaffected due to a compensatory elevation in small tolerant taxa such as oligochaetes, and/or taxa which were in their adult terrestrial life-stage at the time of the spill meaning they avoided

contact with the polluted waters (e.g. chironomids). Mass-abundance scaling of trophic links between consumers and resources revealed extensive restructuring within the food web.

6. This case study shows that pesticides can affect food-web structure and ecosystem functioning, both directly and indirectly across levels of biological organisation. It also demonstrates how an integrated assessment approach, as adopted here, can elucidate links between micro-biota, macroinvertebrates and fish, for instance, thus improving our understanding of the range of biological consequences of chemical contamination in natural ecosystems.

## C.2 Introduction

Freshwaters are exposed to multiple pesticides and other toxic chemicals at local to global scales (Schinegger *et al.* 2011; Beketov *et al.* 2013; Stehle & Schulz 2015). Ecotoxicological experiments in the laboratory have revealed with great accuracy and precision how these can affect the survival of target species (e.g. *G. pulex*; Xuereb *et al.* 2007), and community- and ecosystem-level responses have been demonstrated in micro- and mesocosm experiments (e.g. Van den Brink *et al.* 1995; Van Wijngaarden *et al.* 1996; Traas *et al.* 2004; Halstead *et al.* 2014) and field surveys (Chung, Wallace & Grubaugh 1993; Triebskorn *et al.* 2003; Malaj *et al.* 2014). In the last decade, new indices of community response have been proposed specifically to detect pesticide pollution (e.g. Liess & Ohe 2005; Schäfer *et al.* 2007; Liess, Schäfer & Schriever 2008) and to link community change to toxicants in the field (e.g. Kefford *et al.* 2010).

Despite these advances, a mechanistic understanding of both the toxic effects of pesticides (i.e. direct) and those mediated via the food web (i.e. indirect) across multiple levels of biological organisation (i.e. from genes to ecosystems) is still limited in natural settings (Kohler & Triebskorn 2013). This is likely because there are relatively few opportunities to understand how pesticides affect whole rivers or lakes,

due to the logistical, ethical, and legal difficulties in conducting such a study in a controlled manner. Here, we address this research gap by quantifying the gene-to-ecosystem consequences of a major pesticide spill that caused widespread kills of invertebrates over 15 km in a large lowland river by combining citizen science biomonitoring data with a suite of non-traditional measures of ecosystem impact.

Invertebrate data were collected by citizen scientists prior to, during and after the spill enabling before-after-control-impact (BACI) assessment. These data enabled the UK Environment Agency to identify chlorpyrifos as the cause of the catastrophic mortality following the spill. Chlorpyrifos is a widely used organophosphate pesticide (insecticide and acaricide) which attacks insect (and arachnid) nervous systems. Since insects are core intermediate species in almost all stream food webs, perturbations to their populations have potential to ripple through the entire food web, as bottom-up effects on the fish assemblage and top-down effects on the microbial communities that drive a range of biogeochemical processes. Specifically, chlorpyrifos can affect microbial, invertebrate and fish populations, both directly and indirectly (see reviews by Barron & Woodburn 1995; Brock, Lahr & Van den Brink 2000; Giddings *et al.* 2014), food-web structure (Traas *et al.* 2004) and can suppress invertebrate-mediated litter breakdown (Maltby & Hills 2008). Placing the potentially subtle effects of pesticides within a coherent multilevel framework requires a combination of structural and functional measures from the microbial community at the base of the food web to apex predators. This has been partially achieved in some studies using mesocosms (e.g. Van den Brink *et al.* 1995; Van Wijngaarden *et al.* 1996; Kersting & Van den Brink 1997; Halstead *et al.* 2014), but rarely in natural settings (Kohler & Triebskorn 2013), and never in a manner that simultaneously captures molecular-level responses through to the full complexity of the food web in the same system.

Here we present data that reveal how chlorpyrifos affected the structure and functioning of the river food web, based on several complementary approaches including the abundance of targeted functional genes, those responsible for the

degradation of chlorpyrifos (Kwak *et al.* 2012), for example, measures of microbial and invertebrate resource use and "trivariate analysis" (*sensu* Cohen *et al.* 2009). This collection of measures across multiple levels of organisation provides a vital bridge between field and laboratory-based findings and highlights the advantages of using a holistic approach to understand chemical stressor impacts in natural ecosystems.

We test the following hypotheses:

The structure (assessed using the abundance of functional gene loci) and functional capacity of the microbial assemblage will change due to direct effects (i.e. the pesticide provides an additional substrate) and indirect effects (i.e. increased organic substrates are derived from decaying invertebrates) of the pesticide.

Compensatory mechanisms will be evident in the food web in the aftermath of the spill, with less pesticide-sensitive, small, opportunistic, vagile, and fast-growing taxa (e.g. chironomids) higher in abundance and/or biomass in the absence of larger, slow-growing taxa (e.g. *Gammarus pulex*), relative to control communities.

Leaf litter breakdown will be impaired by the loss of key detritivores, with microbial activity hence accounting for a greater proportion of total litter breakdown.

The food web will undergo extensive restructuring, particularly in terms of altered mass-abundance scaling relationships of the links between nodes. Local extirpations of intermediate species (e.g. herbivorous insects) will release basal species under top-down control (e.g. benthic algae) while suppressing bottom-up fluxes to higher trophic levels (e.g. fish).

## C.3 Methods

### C.3.1 Study site

The River Kennet is a lowland chalk tributary (catchment area 1200 km$^2$) of the River Thames in southern England, designated as a UK Site of Special Scientific Interest (SSSI). The river is groundwater-dominated, has hard water and is nutrient-rich (Figure C.A; Table C.1). Its diverse fauna is dominated by Gammaridae, Baetidae, Ephemerellidae, Simuliidae and Chironomidae, which support an economically important salmonid game fishery (Wright *et al.* 2002; 2004).

On 1 July 2013, following their routine biomonitoring, a citizen-science group (Action for the River Kennet, ARK) reported a large-scale invertebrate kill along a 15-km stretch of the river. On 2 July 2013, an Environment Agency pollution incident team collected the first samples for, and detected, the organophosphate chlorpyrifos. This insecticide attacks the nervous system of insects by inhibiting acetylcholinesterase, and can be toxic to fish and meiofauna (Carr, Ho & Chambers 1997; DeLorenzo, Scott & Ross 1999). Concentrations of of 0.52-0.82µg L$^{-1}$ were recorded coming from the main tertiary sewage treatment works in Marlborough, Wiltshire, on 2 and 5 July, respectively (Figure C.A), probably resulting from a "down-the-drain" incident. The peak concentration was most likely missed by the sampling team, but even the measured concentration is sufficient to be acutely toxic to arthropods (Giddings *et al.* 2014), particularly over extended periods (i.e. >24 hours; Rubach, Crum & Van den Brink 2011). Chlorpyrifos was also detected at concentrations between 0.06-0.07 µg L$^{-1}$ across the impacted study site on 5 July. By 9 July 2013 the pesticide was undetectable, indicating that a single pulse was received and remained in the water column for a few days.

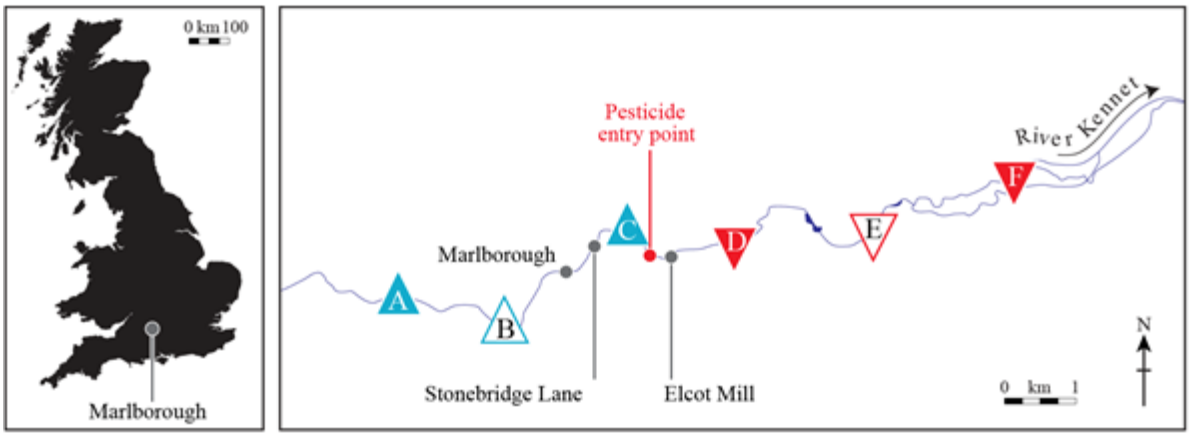



**Figure C.A.** River Kennet (UK) with study sites A-C (upward pointing triangles = control) and D-F (downward pointing triangles = impacted). Data for sites A, C, D and F (filled triangles) are presented here. Monitoring data for aquatic macroinvertebrates were collected by citizen scientists upstream (i.e. control site) at Stonebridge Lane and downstream at Elcot Mill (i.e. impacted site) of Marlborough sewage treatment works, where the pesticide entered the river.

**Table C.1**. Locations of upstream control and downstream impacted sites as well as of water chemistry monitoring stations of the Environment Agency (EA). Mean and range, in brackets, of annual water chemistry concentrations from Environment Agency monitoring data are shown from sites located between control and impacted reaches. Oxidised nitrogen (oxidised N) is the sum of nitrate (NO3-) and nitrite (NO2-).

| Site | Condition | Latitude, Longitude |
|---|---|---|
| A | Control | 51°4170'N, 1°7536'W |
| EA Control | Control | 51°4235'N, 1°7165'W |
| C | Control | 51°4227'N, 1°6982'W |
| D | Impacted | 51°4227'N, 1°6982'W |
| EA Impact | Impacted | 51°4170'N, 1°7536'W |
| F | Impacted | 51°4163'N, 1°7325'W |
| Water chemistry | EA Control | EA Impacted |
| Alkalinity (mg $L^{-1}$) | 250 (187-262) | 243 (189-254) |
| Conductivity (µS $cm^{-1}$) | 626 (449-738) | 609 (492-686) |
| Oxidised N (mg $L^{-1}$) | 6.6 (4.4-7.5) | 6.8 (4.4-7.6) |
| Dissolved oxygen (mg $L^{-1}$) | 9.0 (6.9-10.0) | 9.6 (6.9-10.9) |
| Temperature (°C) | 11.0 (5.7-14.4) | 11.1 (5.7-14.5) |
| pH | 7.6 (7.4-7.8) | 7.9 (7.4-8.1) |
| Ortho-phosphate (mg $L^{-1}$) | 0.08 (0.02-0.36) | 0.08 (0.02-0.34) |

### C.3.2 Contribution of citizen scientists

Citizen scientists from ARK were trained by the Riverfly Partnership to collect and identify aquatic macroinvertebrates and had collected data for multiple sites for several years prior to and following the spill (Fig. S1). During the current study, they collected one monthly kick sample (3-minutes duration) from an upstream control and downstream impacted site (Figure C.A). A standard hand net (1-mm mesh) was used following the Riverfly Monitoring Initiative standard protocol (http://www.riverflies.org). The invertebrates collected were identified live on the bank, without magnification, and abundance ranked per sample as: 0 = 0 individuals; 1-9 = 1; 10-99 = 2; 100-1000 = 3; >1000 = 4, for eight key groups: 1. Cased Trichoptera; 2. caseless Trichoptera; 3. Ephemeridae; 4. Ephemerellidae; 5. Heptageniidae; 6. Baetidae; 7. Plectoptera; 8. Gammaridae, which were summed to give a total score based on the number and diversity of the target taxa. These data provide a critical BACI element to the study, enabling us to track the impact of the spill through both space and time.

Mean annual water chemistry data were obtained for Environment Agency monitoring stations located 2.3 km upstream and 2.7 km downstream from the spill and were similar across the study site (Table C.1). These water chemistry data, combined with the ARK monitoring data of macroinvertebrates, showed no evidence of organic pollution from the sewage treatment works, indicating that sewage was an unlikely cause of the invertebrate mortality event (Fig. S1).

### C.3.3 Sampling protocol

Comprehensive biological sampling began in September 2013, as soon as possible after the chlorpyrifos spill had been identified as the causal agent, using an experimental design comprising three upstream control and three downstream impacted reaches, each 50m long, along a *c*. 6km river stretch (Figure C.A). Sites were c. 1km apart, with similar channel forms and riparian surroundings. Here we present

data from two control and two impacted reaches (Figure C.A) for a suite of structural and functional indicators to test a multilevel bioassessment approach. Three sediment samples, a stone scrape, three Surber samples and depletion electrofishing were used to characterise microbial, diatom, macroinvertebrate and fish structural attributes, respectively. At each site, 10 fine- (0.5mm) and 10 coarse-mesh (10mm) leaf-litter bags were used to determine rates of decomposition driven by microbes alone or by whole communities. In addition, a sample of river water was collected and incubated with a range of substrates to assess microbial functional capacity.

### C.3.4 Microbial functional gene abundance

We used quantitative PCR (qPCR) to examine gene abundance for microbial functional and taxonomic marker genes. 16S rRNA gene abundance was used as a proxy for total bacterial abundance. Direct effects of the chlorpyrifos spill were examined using the organophosphate hydrolase gene (*opd*), which is responsible for the degradation of chlorpyrifos by bacteria; bacterial populations containing this gene have previously been demonstrated to increase in abundance at sites impacted by organophosphate (Kwak *et al.* 2012). Indirect effects were examined by quantifying the abundance of genes coding for enzymes involved in N-cycling: nitrite reductase (*nirS*) and ammonia monoxygenase (*amoA*) from ammonia-oxidising archaea (AOA) and bacteria (AOB) as these are most likely to reflect decomposition of dead arthropods in impacted sites. We hypothesised that decomposition of dead arthropods would result in an increased input of $NH_4^+$ from ammonification of organic N. We focused on *nirS* and *amoA* genes as both nitrification and denitrification pathways are important in removing N from systems and can be coupled when denitrifies reduce the $NO3^-$ produced by the nitrifiers that oxidised $NH_4^+$. By focusing on functions of a range of populations, a change across all populations combined provides an indicator for community-level effects of chlorpyrifos on river microbes. Full details of DNA isolation, primer details and qPCR cycling conditions are available

in the Microbial Functional Gene Abundance section in the Supplementary Material.

### C.3.5 Microbial functional potential

Open-water samples were collected from each site and returned to the laboratory in an ice-chilled cooler. Samples were allowed to settle (>10 min), after which a 100-µL aliquot was pipetted into each well of a Biolog EcoPlate, which contained a single carbon substrate, including carbohydrates, polymers, fatty acids and amino acids. Each well also contained the redox dye tetrazolium, which is reduced during microbial respiration, resulting in a measurable colour change. Each EcoPlate contains 31 substrates plus a no-substrate control in triplicate. Plates were incubated in the dark at 22°C for 5 days, after which colour change was quantified by measuring optical density at 600 nm using a Biotek HT absorbance reader (Biotek, Swindon, UK). For each EcoPlate, we calculated the substrate usage by subtracting the mean of the three no-substrate controls from each measurement. Usage was ranked across the substrates in each replicate, and the ranked optical densities were plotted to visualise broad changes across sites.

### C.3.6 Population abundance, community structure and food web size-scaling

Quantitative depletion electrofishing was undertaken, with population densities estimated using the R package FSA (Ogle 2012) and iterative Maximum Weighted Likelihood statistics (equation S1 and S2 in Supplementary Material; after Carle & Strub 1978). All fishes caught were identified to species and measured by fork length. For each species, individual dry mass was calculated from length using length-mass regression equations generated from a sub-sample (see equations S1 and S2 in Supplementary Material).

Invertebrates were collected ($n$ = 3 samples per site) using a Surber sampler

(0.0625 $m^2$, 335 μm mesh), preserved in 99.8% ethanol, and later sorted from debris, identified to the highest possible taxonomic resolution (usually species), and counted (Table S1). Dry masses of invertebrates were determined from regressions of linear dimensions using published equations (see Table S2); a subset of 60 individuals were measured per species per site, or every individual where abundance was below 60. We distinguished between arthropods (i.e. insect larvae and Crustacea) and other taxa (i.e. Tricladida, Annelida and Mollusca) based on their sensitivity to chlorpyrifos (Raven & George 1989; Giddings *et al.* 2014).

Diatoms were scraped from 8.64 $cm^2$ of the upper surface of one cobble at each site using a 3.6 by 2.4 cm photographic slide as a flexible quadrat and toothbrush, preserved using Lugol's iodine, and prepared using standard methods (Battarbee *et al.* 2001). A minimum of 300 diatom valves were identified to species per sample using the keys of Krammer & Bertalot (1986), Krammer *et al*. (1986), Krammer & Lange-Bertalot (1991a b) and abundances per unit area were determined as in Battarbee (1973). Linear dimensions were measured to the nearest 1μm to estimate diatom biovolume (Table S3; Hillebrand *et al.* 1999). The first 30 specimens of all common ($n$ >30) species were measured and where species were encountered less frequently, all specimens in the count were measured. Carbon content was estimated (Rocha & Duncan 1985) and then converted to dry mass (Sicko-Goad, Schelske & Stoermer 1984).

We used these mass-abundance data from across the different taxa and trophic levels to construct whole-community 'trivariate food webs' - food webs ordinated by overlaying feeding links on the bivariate relationship between species mean body mass and their numerical abundance on a double logarithmic scale - to understand how chlorpyrifos alters food-web structure. Deviations in *MN* among species pairwise links can be used to identify alterations to biomass fluxes in the food web. For instance, altered consumer-resource feeding "link angles" can reveal rates of change in biomass, population production and population consumption between species-pairs,

through to the food web as a whole (*sensu* Cohen *et al.* 2009), and these changes can help us to interpret direct and indirect effects of chlorpyrifos.

Trivariate webs were constructed for all sites. Feeding links were inferred from trophic interactions published in the literature (Table S4). We assumed that if a trophic interaction between two species has been reported in the literature and those same species were present at one of our sites, then that trophic interaction also occurred, as has been validated in other stream food webs (Layer *et al.* 2010; Layer, Hildrew & Woodward 2013).. In a few instances, feeding links were assigned on the basis of taxonomic similarity. For example, if a link had been established from the literature for at least one congener it was assumed that different species within the same genus fed upon the same resources and were consumed by the same consumers. It was necessary to extend this assumption to the family level in some instances where information in the primary literature was scarce (Table S5). This minimises bias between nodes where the quantity of directly observed information varies and allows the method to be reproduced exactly (Gray *et al.* 2014).

### C.3.7 Ecosystem functioning: leaf-litter decomposition

At each site, the decomposition rate of leaf-litter was determined from leaf-packs containing 3.0 g (±0.3 g SD) black alder (*Alnus glutinosa*) incubated in the river for 9 days. Coarse (150 mm by 100 mm, 10mm mesh) and fine (150 mm by 100 mm, 500 μm mesh) mesh-aperture bags were used to determine the fraction of decomposition contributed by microbes (mass loss from fine mesh bags) and invertebrates (difference in mass loss from coarse and fine mesh bags). Leaf breakdown rates were expressed as the exponential decay rate coefficient, $k$ (see equation S3; Woodward *et al.* 2012).

### C.3.8 Data analysis

Trivariate statistics were calculated using the method of Cohen *et al* (2009) in the R package Cheddar (Hudson *et al.* 2012). We used link angles to estimate changes in potential biomass flux between a resource and its consumer. In summary, a link can be viewed as a vector from a resource to its consumer and, considering that invertebrate taxa abundance and/or mass is predicted to decrease at impacted sites, a change in the angle of invertebrate upper- and lower-links would indicate a potential change in biomass flux (Figure C.B).

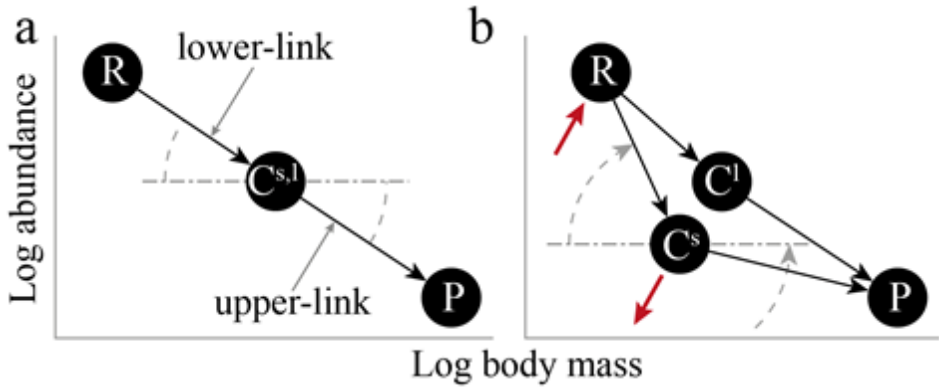



**Figure C.B.** (a) Location of consumers sensitive to pesticides (Cs) and less sensitive to pesticides (Cl) in relation to the consumer resources (R) and predators (P) as viewed on a double-logarithmic scale of body mass versus abundance. (b) Changes within the food web following pesticide exposure can be assessed by using link angles as a proxy for changes in potential biomass flux within the food web: a predicted decrease in Cs MN following pesticide exposure and an increase in R MN due to the release from top-down consumer control can be assessed using the Cs link angles in relation to Cl and control data; a decrease in Cs lower-link angles would indicate a potential reduction in biomass flux between R-Cs; an increase in Cs upper-link angle could indicate hysteresis within the network whereby P is yet to be impacted by the loss of Cs, or that P has increased reliance on other resources, or a combination of the two.

Linear mixed effect models (LMM) were used to test for differences in mean annual water quality, with treatment and date as fixed and random factors, respectively. Differences in biotic response variables (link angles, species and

community abundance and/or biomass, gene abundances and microbial capacity) between treatments were tested using LMM with site and treatment as random and fixed factors, respectively. Where necessary a variance structure was used to account for unequal variance between sites in order to meet model assumptions (after Zuur *et al.* 2009). If data were not normally distributed they were $Log_{10}$ transformed to meet the assumptions of the test. All LMM were performed using the nlme package in R (Pinheiro *et al.* 2011) and estimates were made using restricted maximum likelihood or, when testing for differences in group means (e.g. invertebrate communities within and between treatments), using general linear hypotheses tests in the R package multcomp (Hothorn *et al.* 2014).

## C.4 Results

### C.4.1 Macroinvertebrate monitoring by citizen scientists

Within control sites, *G. pulex* had the highest relative abundance (61%), followed by Baetidae (17%), Ephemerellidae (12%), cased Trichoptera (9%) and Plecoptera (1%). The macroinvertebrate assemblage in the three months prior to the spill was similar but following the spill on July 1st 2013, there was a 99.5% reduction in total abundance from the previous month (Figure C.C). By September, total abundance had increased again, but was dominated by Ephemeroptera instead of *G. pulex*, the latter being the slowest taxa to recover, as recorded by the citizen scientists. When the citizen science macroinvertebrate data and Environment Agency water quality data were combined there was no evidence to suggest that nutrient pollution was the cause of the macroinvertebrate mortality event (results are presented in the Supplementary Material).

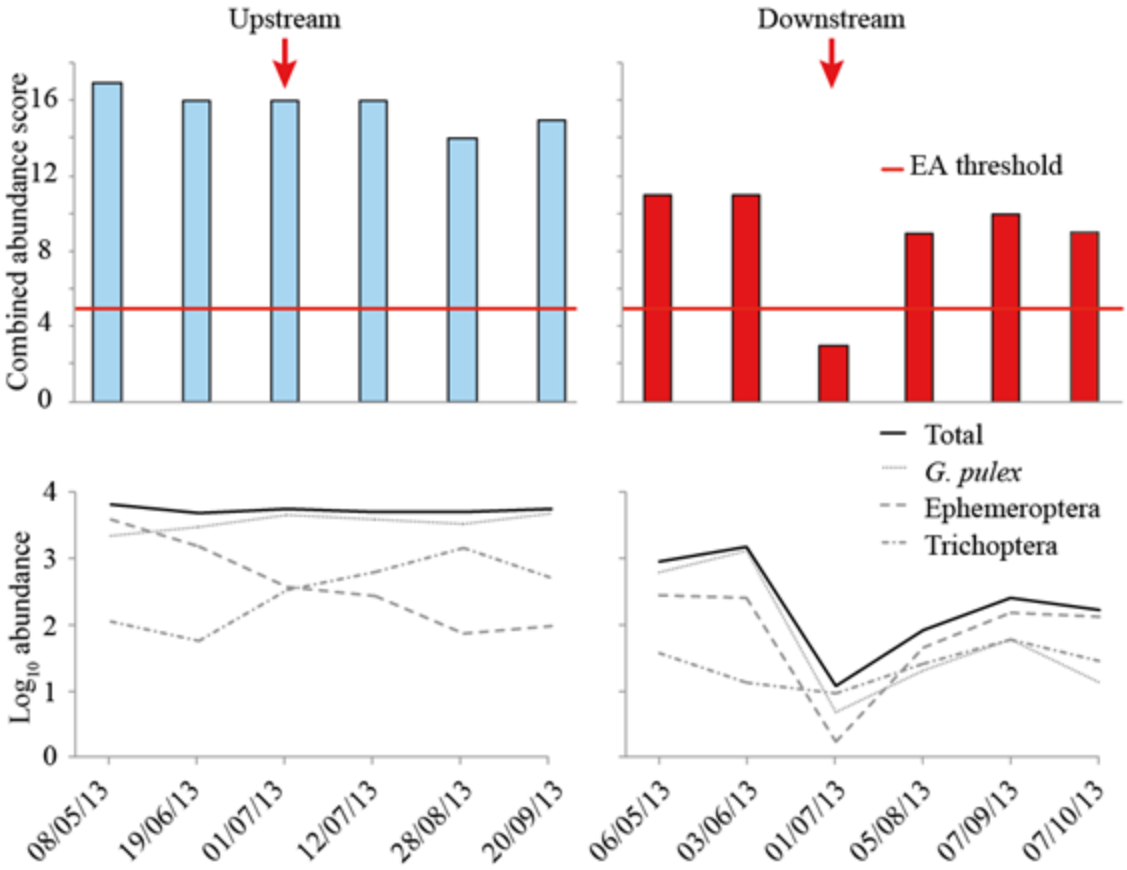



**Figure C.C**. Top: Aquatic macroinvertebrate monitoring data collected by citizen scientists show macroinvertebrate scores before and after the toxic spill (red arrows), based on total abundance of the target taxa. The red line represents an Environment Agency threshold for substantial ecological degradation. Bottom: abundance of key taxa in relation to scores collected from an upstream control at Stonebridge Lane and a downstream impacted site at Elcot Mill (see Figure C.A).

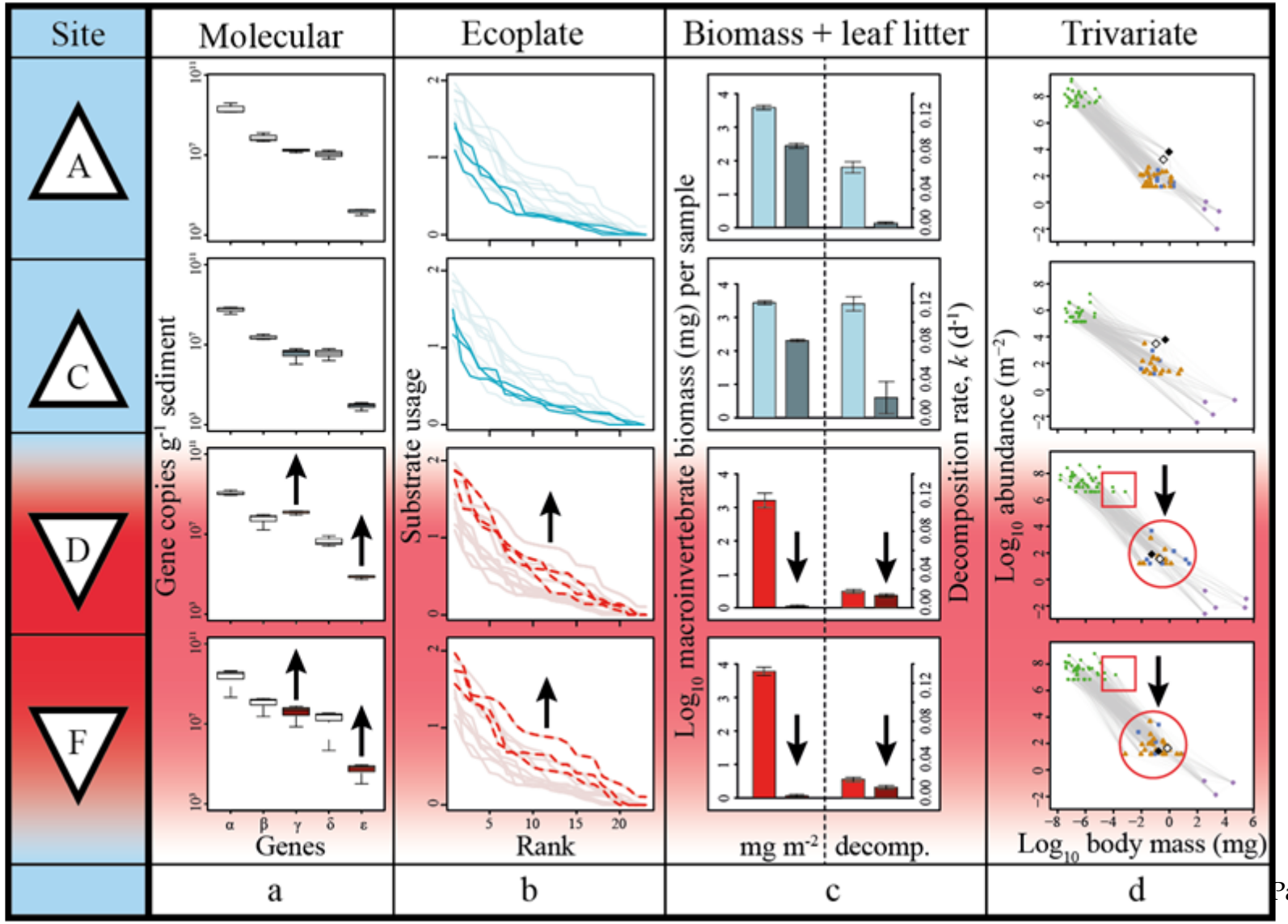

Site
Molecular
Ecoplate
Biomass + leaf litter
Trivariate
A
C
D
F
Gene copies g-1 sediment
Genes
Substrate usage
Rank
Log10 macroinvertebrate biomass (mg) per sample
mg m-2
decomp.
Decomposition rate, k (d-1)
Log10 abundance (m-2)
Log10 body mass (mg)
a
b
c
d

**Figure C.D.** Vertical arrows indicate notable differences between ecological data from control sites A and C and from impacted sites D and F two months after the toxic spill. (a) Molecular results from microbial qPCR assays targeting the (α) 16S rRNA (microbial abundance), (β) *nirS* (nitrite reductase) (γ) *amoA* (ammonia monooxygenase) AOB (ammonia oxidising bacteria), (δ) *amoA* (ammonia monooxygenase) AOA (ammonia oxidising archaea), (ε) *opd* (organophosphorus hydrolase) genes. (b) Ecoplate microbial functional potential on 31 carbon substrates (x-axis) and their usage (y-axis; measured as optical density at 600 nm after 5 days of incubation at 22 °C as defined in the Methods) (c) Biomass of macroinvertebrates (light shading) and a keystone detritivore, *Gammarus pulex* (dark shading), and leaf-litter breakdown rates by all consumers (light shading) and microbes only (dark shading); error bars represent standard error (d) Trivariate mass-abundance food webs: green circles = algae (large species found only in the impacted sites highlighted), yellow symbols = arthropods (decreased relative to controls), blue symbols = other macroinvertebrates, black filled diamond = *G. pulex*, black open diamond = *Baetis*, pink symbols = fishes.

### C.4.2 Microbial functional gene abundance and functional potential

Analyses of gene abundances revealed that ammonia oxidisers (*amoA*), particularly AOBs, were up to 30-fold higher ($t_2$ = 4.99; $p$ = 0.03), and populations capable of utilising organophosphate (*oph*) as a resource were up to 7-fold higher in impacted sites compared with control sites (Figure C.Da; $t_2$ = 6.14; $p$ = 0.02). The elevation in the abundance of these populations suggests both direct (i.e. microbes utilised the insecticide as a resource) and indirect effects (i.e. microbes utilised ammonia released by decaying invertebrates) of chlorpyrifos. However, there was no significant difference in the total abundance of bacteria, nor of the abundance of nitrite reducers or AOAs (Figure C.Da).

The functional microbial assays showed impacted sites had higher overall substrate usage and a shallower rank abundance curve, indicating substantial functional changes in response to the spill. Mean overall carbon usage in the impacted sites differed from that in the control sites (Figure C.Db; $t_2$ = 4.2, $p$ = 0.05), with lower mean substrate usage in the latter. Differences among control and impacted sites suggested elevated rates of substrate usage of simple carbohydrates (e.g. glucose-1-phosphate, $t_2$ = 4.4, $p$ = 0.05; α-D-lactose, $t_2$ = 7.7, $p$ = 0.02) and amino acids in the impacted sites, with little difference in the usage of complex polymers (e.g. Tween 40).

### C.4.3 Macroinvertebrate community structure and ecosystem functioning

Total macroinvertebrate biomass and abundance did not significantly differ between the control and impacted sites ($t_2$ = -1.43; $p$ = 0.29; $t_2$ = -2.11; $p$ = 0.17). However, arthropod biomass was 92.9% lower in impacted sites than arthropod biomass in control sites and 80.4% lower than biomass of less pesticide-sensitive taxa in impacted sites (

Table C.**2**; Figure C.E). In addition, the biomass of macroinvertebrate taxa considered less sensitive to pesticides was 97.2% lower than that of the sensitive arthropods in control sites (Table C.2), thus the former were partly compensating for the loss of the latter within impacted sites. *G. pulex* biomass (99.6%) and abundance (99.2%) and *Baetis* biomass (18.7%) and abundance were lower (95.6%; Figure C.Dc; Figure C.Dd), but chironomid biomass (89.3%) and abundance (92.2%) and oligochaete biomass (85.4%) and abundance was higher in impacted sites compared to control sites (94.5%; Figure C.E; Table C.2). Macroinvertebrate diversity was similar between control and impacted sites ($t_2$ = -0.39; $p$ = 0.74), as was also true for fish diversity (Table C.3), whereas four taxa of large diatoms (*Cymatopleura solea*, *Cymatopleura elliptica, Gyrosigma attenuatum* and *Surirella caproni*) were present only in the impacted sites (Figure C.Dd). Microbial decomposition was higher, whereas total decomposition mediated by both microbes and detritivores was lower, in the impacted sites (Table C.2; Figure C.Dc), probably reflecting the decline of *G. pulex* and partial compensation by increased microbial activity.

**Table C.2**. General linear model tests of the biomass (mg) and abundance of arthropods and other macroinvertebrates (Tricladida, Annelida and Mollusca, which are considered to be less sensitive to chlorpyrifos than arthropods) per sample; *Baetis*, *Gammarus pulex* (i.e. *K*-selected taxa), chironomid and oligochaete (i.e. *r*-selected taxa) biomass and abundance; arthropod-resource and other-resource trivariate lower link angles, *Baetis* and *G. pulex* upper-link angles and both total and microbial leaf-litter breakdown rate between control (C) and impacted (I) sites. Significant *p* values (<0.05) are highlighted in bold.

| $Log_{10}$ (biomass +1) | Estimate | Std. Error | *z* value | *p* |
|---|---|---|---|---|
| C:arthropods - C:other | 1.62 | 0.09 | 17.53 | **<0.001** |
| I:arthropods - I:other | -0.73 | 0.12 | 6.00 | **<0.001** |
| C:arthropods - I:arthropods | 1.17 | 0.23 | 5.19 | **<0.001** |
| C:other - I:other | -1.17 | 0.25 | -4.73 | **<0.001** |
| $Log_{10}$ (abundance +1) | | | | |
| C:arthropods - C:other | 1.28 | 0.19 | 6.82 | **<0.001** |
| I:arthropods - I:other | -0.05 | 0.19 | 0.25 | 0.99 |
| C:arthropods - I:arthropods | 0.56 | 0.24 | 2.37 | 0.06 |
| C:other - I:other | -0.76 | 0.24 | -3.23 | **0.005** |
| $Log_{10}$ (biomass +1) | | | | |
| C:*Baetis* - I:*Baetis* | 0.62 | 0.16 | 4.00 | **<0.001** |
| C:*G. pulex* - I:*G. pulex* | 2.30 | 0.15 | 15.82 | **<0.001** |
| C:chironomids - I:chironomids | -0.93 | 0.15 | -6.38 | **<0.001** |
| C:oligochaetes - I:oligochaetes | -0.81 | 0.15 | -5.49 | **<0.001** |
| $Log_{10}$ (abundance +1) | | | | |
| C:*Baetis* - I:*Baetis* | 1.21 | 0.24 | 4.98 | **<0.001** |
| C:*G. pulex* - I:*G. pulex* | 2.31 | 0.22 | 10.63 | **<0.001** |
| C:chironomids - I:chironomids | -1.14 | 0.22 | -5.24 | **<0.001** |
| C:oligochaetes - I:oligochaetes | -1.12 | 0.23 | -4.92 | **<0.001** |
| Invertebrate-resource lower-link angles | | | | |
| C:arthropods - C:other | -0.08 | 0.02 | -3.8 | **<0.001** |
| I:arthropods - I:other | 0.2 | 0.02 | 10.35 | **<0.001** |
| C:arthropods - I:arthropods | -0.32 | 0.24 | -1.36 | 0.44 |
| C:other - I:other | -0.04 | 0.24 | -0.18 | >0.99 |
| *Baetis* and *G. pulex* upper-link angles | | | | |
| C:*Baetis* – I:*Baetis* | -103.71 | 24.3 | -4.27 | **<0.001** |
| C:*G. pulex* – I:*G. pulex* | -62.8 | 25.73 | -2.44 | **0.03** |
| Leaf litter decomposition (*k*) | | | | |
| I:total - C:total | -0.05 | 0.01 | -6.57 | **<0.001** |
| I:microbial - C:microbial | 0.01 | 0.002 | 5.75 | **<0.001** |

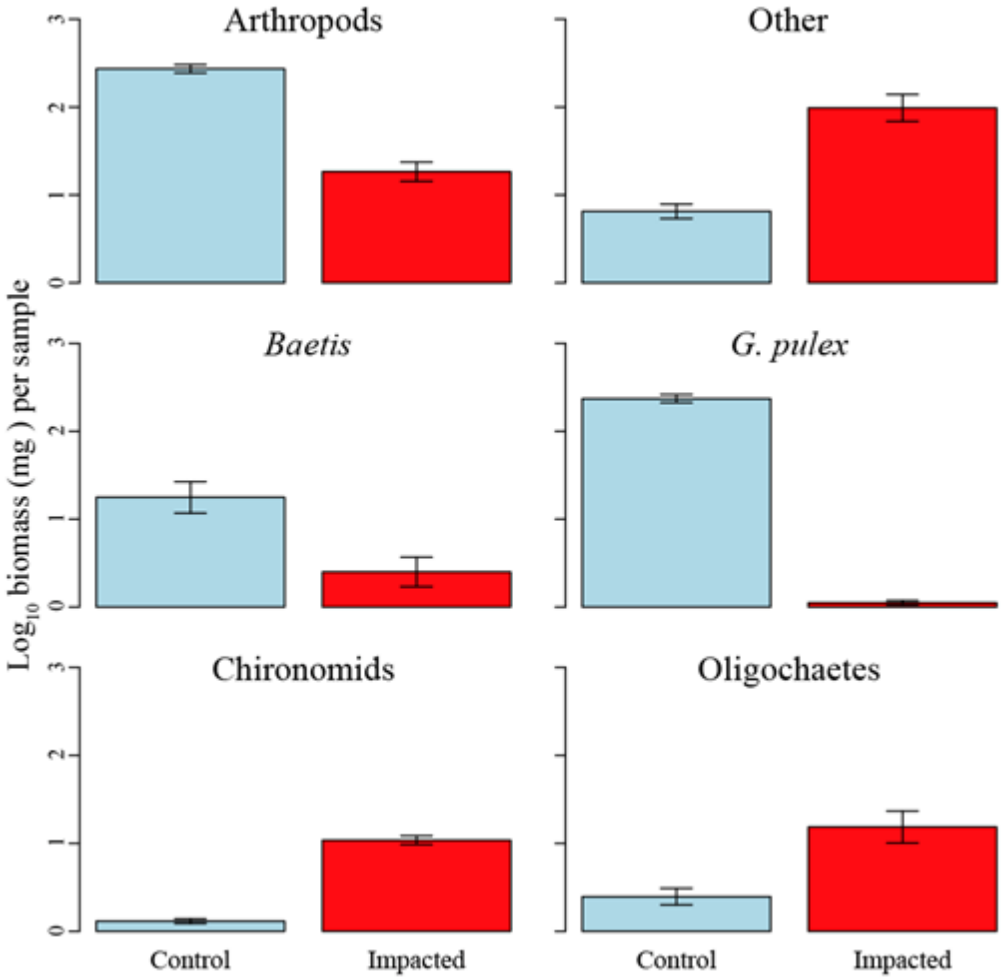



**Figure C.E.** Macroinvertebrate mean biomass (per sample with standard error) at control and impacted sites in the River Kennet.

### C.4.4 Trivariate analysis

Arthropod lower-link angles were less negative (i.e. shallower) than less pesticide-sensitive taxa in the control communities, but more negative (i.e. steeper) within the impacted communities (Table C.2). This indicates altered mass-abundance scaling relationships of the links between nodes (Figure C.B). *G. pulex* and *Baetis* had

the highest biomass and numerical abundance within the control macroinvertebrate community, respectively (Figure C.Dc, Figure C.Dd), and these species upper-link angles (i.e. to their predators) became shallower at impacted sites (Table C.2), thus indicating a potential decrease in biomass flux to fishes from both the detritivore and herbivore food chains. To illustrate the direction of biomass flux through the food web and the connection of a key species to all other taxa via relatively direct and short paths, we constructed an example food chain with *G. pulex* as the focal species (Figure C.F), which showed that even in this complex food web most species are only 1-2 links from all the others, highlighting the potential for perturbations to ripple rapidly through the network. More commonly used whole-network metrics, such as the regression slope and intercept, showed no clear differences that could be ascribed to the pesticide spill (Table C.3).

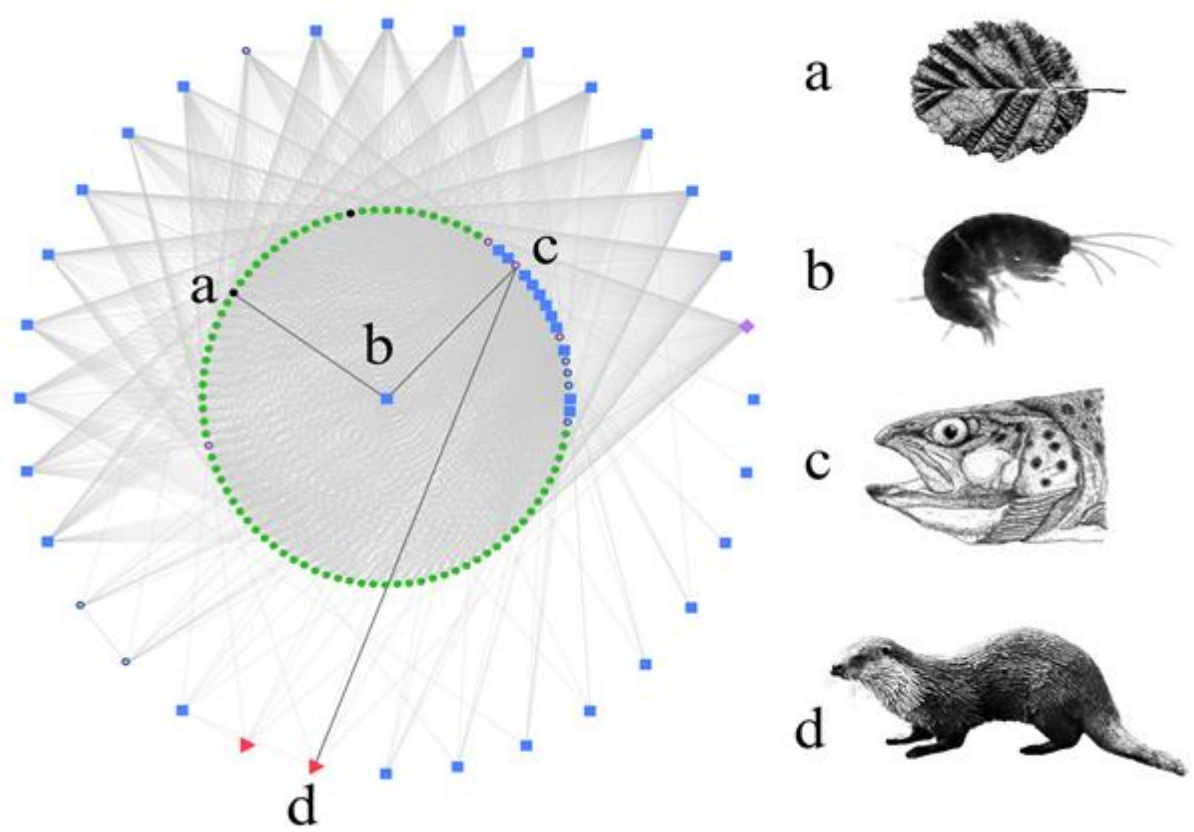



**Figure C.F.** Aggregated network for the River Kennet food web, highlighting an exemplar food chain from the basal resource to the apex predator; a = coarse particulate organic matter (e.g. leaf litter), b = Gammarus pulex, c = brown trout, Salmo trutta, d = Eurasian otter, Lutra lutra. The two concentric circles of nodes represent the shortest food web distances to or from G. pulex – those in the inner circle are a single link removed from G. pulex, those in the outer circle are separated by two links in the shortest path. Here, all species are at most 2 links away from G. pulex, although longer food chains are present in the network, as shown by a-b-c-d. Symbols for nodes represent different trophic elements: green circles represent producers, blue squares: macroinvertebrates, purple diamonds: vertebrate ectotherms, red triangles: endotherms, black circles: abiotic resources. Light blue and light purple circles represent cannibalistic nodes of invertebrates and vertebrate ectotherms, respectively.

**Table C.3.** Properties of the trivariate food webs at control and impacted stream sites.

| | Site A | Site C | Site D | Site F |
|---|---|---|---|---|
| Property | Control | Control | Impacted | Impacted |
| Number of nodes | 68 | 60 | 64 | 73 |
| Number of fish species | 4 | 4 | 5 | 3 |
| Number of macroinvertebrate taxa | 35 | 23 | 20 | 32 |
| Number of diatom taxa | 29 | 33 | 39 | 38 |
| Number of links | 837 | 635 | 739 | 1060 |
| Linkage density | 11.96 | 10.41 | 11.37 | 14.13 |
| Directed connectance | 0.17 | 0.17 | 0.17 | 0.19 |
| Trivariate regression slope | -0.98 | -0.67 | -0.92 | -0.95 |
| Trivariate regression intercept | 1.29 | 1.26 | 1.58 | 1.35 |

## C.5 Discussion

The documented insecticide spill in the River Kennet affected multiple organisational levels, from individual genes, through to food web structure and an ecosystem process. The location of pesticide-sensitive macroinvertebrate consumers relative to their resources in *MN* space shifted markedly, and the collapse in the population of a previously dominant keystone detritivore, *G. pulex*, was especially notable. This was associated with dramatically impaired rates of detritivore-mediated litter decomposition, with potential repercussions for the higher trophic levels. In this highly interconnected food web (Figure C.F) perturbations could potentially not only easily propagate through species interactions, but could also dissipate effectively. These properties could confer resilience on the system as a whole, as alternative feeding paths provide relatively direct "short-circuits" in the food web (Figure C.F). Various compensatory mechanisms and hystereses within the food web were evident following the spill, including elevated microbial decomposer activity in the absence of invertebrate detritivores (Figure C.Dc) and irruptions and growth of less pesticide-sensitive and *r*-selected taxa capable of exploiting new resources (Figure C.E). The functional potential of the microbial assemblage in particular was higher in the impacted sites, as was the abundance of genes associated with organophosphate use and ammonia oxidation in the aftermath of widespread arthropod deaths (Figure C.Da; Figure C.Db). Extended temporal sampling will likely reveal if the sewage treatment work is potentially confounding our interpretation of this result, although there is no suggestion this is the case, as water quality is essentially identical above and below the works (Fig. S1).

Microbial biodiversity accounts for most of a river's biodiversity, drives key ecosystem processes and biogeochemical cycles (e.g. nitrogen cycle) and interacts with higher trophic levels. Our qPCR assays revealed that the abundance of genes associated with the turnover of organophosphate and ammonia was higher in polluted sediment, revealing both direct and indirect effects of the spill on microbial

activities.

Strong links between changes in the structure and functioning of the microbial and invertebrate community were evident, as revealed by the changes in decomposition rates associated with these two major biotic drivers (Gessner & Chauvet 2002; Schäfer *et al.* 2007). The microbial community played a key role in maintaining litter decomposition following the invertebrate losses, and microbial functional potential assessed by Ecoplate assays was also elevated at the impacted sites. The large-scale mortality of invertebrates was likely to have released resources readily available for microbial use, promoting the proliferation of fast-growing bacteria able to use a broad range of substrates. Additional data from more extended sampling will eventually help us to better understand the temporal dynamics of the recovery process, by providing deeper insights into the baseline variability. Even in the current absence of such additional data, our results clearly underline the potential of microbial bioindicators for assessing direct and indirect responses of river ecosystems to environmental impacts.

Employing a highly resolved network-based perspective provided further insights into both direct and indirect effects of the perturbation - from genes to species and from food webs to the ecosystem as a whole - as we were able to connect structural and functional indicators across different levels of biological organisation, as well as improving understanding of the associated responses. For instance, *G. pulex* and *Baetis* represented key nodes in the major detritivore and herbivore food chains, respectively, as is the case in many lowland running waters (Woodward *et al.* 2008; Layer *et al.* 2010), and both populations collapsed in the impacted sites. Our broad multilevel approach revealed how the loss of consumers could result in the release of their resources (or potential competitors), and also how major conduits of energy and biomass flux to the species at the top of the food web, including ecologically important and economically valuable fish species, such as trout, could be compromised.

Microcosm and mesocosm experiments have described ecosystem-level

responses to, and recovery from, combined pesticide and nutrient additions (Traas *et al.* 2004; Halstead *et al.* 2014), and observational field-based research has demonstrated that recovery of the invertebrate community and leaf-litter decomposition was related to aerial mobility of repopulating taxa (Chung *et al.* 1993). Our study represents a novel approach, integrating a broad range of assessment metrics at multiple levels and this has helped us to better understand the effects of a pesticide spill in a natural setting. The same approach is also more widely applicable to assessments of effects caused by other stressors, such as acidification and eutrophication, where interactions within food webs can shape both the ecosystem impact and the rate and trajectory of recovery (e.g. Ledger & Hildrew 2005; Layer *et al.* 2010; Rawcliffe *et al.* 2010). Thus, such an approach offers a way to move beyond partial taxonomic or trait-based views to one that explicitly incorporates species interactions in food webs and ecosystem processes in river bioassessment (Gray *et al.* 2014).

Our study also highlights the value of citizen science in biomonitoring and bioassessment, as it enabled us to place the detailed data specifically and intensively collected after the toxic spill in the context of a wide before-and-after-control-and-impact (BACI) -style "natural experiment", which would have otherwise been impossible to employ in the search for causal relationships. Mobile Ephemeroptera (*Baetis* and Ephemerellidae, both active swimmers with an aerial adult that coincided with the pollution) repopulated the river more quickly than *G. pulex* (Figure C.C), as did the often opportunistic chironomid species and less sensitive non-arthropod taxa such as oligochaetes (Figure C.E). These responses echo those of small *r*-selected taxa preceding the recovery of larger *K*-selected species in previous studies on pesticide contamination (Chung *et al.* 1993; Liess & Schulz 1999; Beketov *et al.* 2008).

It has been hypothesised that ecological inertia can operate within freshwater food webs, creating 'community closure' or recovery trajectories that are not simple reversals of impacts (e.g. Ledger & Hildrew 2005; Layer *et al.* 2011; Layer, Hildrew & Woodward 2013). Impacts on key nodes can alter important aspects of food-web

structure and associated processes, such that although the latter might operate at similar rates, they may be driven by microbes and *r*-selected taxa instead of *K*-selected taxa, as has been reported in response to pesticide contamination (Chung *et al.* 1993) and other stressors (Hladyz *et al.* 2011). Our initial data demonstrate that, while the R. Kennet's ecological structure and functioning were significantly altered by the toxic spill, there were many alternative nodes and links within the food web that could help confer some level of resilience even in the face of catastrophic population losses.

Future work will require well co-ordinated laboratory and field investigations based on matching methodologies to improve understanding of the links between microbiota and larger organisms before, if ever, one can be used as a proxy for the other (e.g. Triebskorn *et al.* 2003). Nonetheless, our study represents a proof-of-concept as to how vastly different metrics might be linked and, as more data are generated over time, potential time × treatment interactions can also be more thoroughly explored. Additional metrics based on, for instance, next-generation sequencing (e.g. Rosi-Marshall *et al.* 2013) or measures of whole-ecosystem respiration (e.g. Young, Matthaei & Townsend 2008), could be incorporated to capture the extent of impacts and recovery trajectories more fully.

Although covering only part of the spectrum of responses reported here, other multimetric bioassessments have yielded comparable results, including how pesticides can indirectly release prey species from predation (Papst & Boyer 1980), constrain consumer populations through loss of resources (Brazner & Kline 1990), affect the structure and functioning of aquatic communities in mesocosms (Downing *et al.* 2008; Relyea 2008; Halstead *et al.* 2014) or alter the structure and functioning of natural stream communities (Chung *et al.* 1993; Schäfer *et al.* 2007). Results from correlational studies also suggest that changes at multiple trophic levels may be related to organic chemical contaminants (mostly pesticides) at the continental scale (Malaj *et al.* 2014). Despite this and the worldwide use of, and projected increase in, pesticides, studies of their effects at the ecosystem-level are rare in natural settings

(Kohler & Triebskorn 2013). The present study contributes to bridging this gap.

## C.6 References

Barron M.G. & Woodburn K.B. (1995) Ecotoxicology of chlorpyrifos. *Reviews of Environmental Contamination and Toxicology* **144**, 1–93.

Battarbee R.W. (1973) A new method for the estimation of absolute microfossil numbers, with reference especially to diatoms. *Limnology and Oceanography* **18**, 647–652.

Battarbee R.W., Jones V.J., Flower R.J. & Cameron N.G. (2001) Diatoms. In: Tracking Environmental Change Using Lake Sediments. Vol. 3: Terrestrial, Algal, and Siliceous Indicators (Eds Smol J.P., Birks H.J.B., & Last W.M.) pp. 155–202. Kluwer Academic Publishers, Dordrecht.

Beketov M.A., Kefford B.J., Schäfer R.B. & Liess M. (2013) Pesticides reduce regional biodiversity of stream invertebrates. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 11039–11043.

Beketov M.A., Schäfer R.B., Marwitz A., Paschke A. & Liess M. (2008) Long-term stream invertebrate community alterations induced by the insecticide thiacloprid: Effect concentrations and recovery dynamics. *Science of the Total Environment* **405**, 96–108.

Brazner J.C. & Kline E.R. (1990) Effects of chlorpyrifos on the diet arid growth of larval fathead minnows, pimephales promelas, in littoral enclosures. *Canadian Journal of Fisheries and Aquatic Science*s **47**, 1157–1165.

Brock T., Lahr J. & Van den Brink P.J. (2000) Ecological risks of pesticides in freshwater ecosystems. Part 2: Insecticides. Alterra, Green World Research, Wageningen, the Netherlands.

Carle F.L. & Strub M.R. (1978) A new method for estimating population size from removal data. *Biometrics*, 621–630.

Carr R.L., Ho L.L. & Chambers J.E. (1997) Selective toxicity of chlorpyrifos to several species of fish during an environmental exposure: biochemical mechanisms. *Environmental Toxicology and Chemistry* **16**, 2369–2374.

Chung K., Wallace J.B. & Grubaugh J.W. (1993) The impact of insecticide treatment on abundance, biomass and production of litterbag fauna in a headwater stream: a study of pretreatment, treatment and recovery. *Limnologica* **23**, 93–106.

Cohen J.E., Schittler D.N., Raffaelli D.G. & Reuman D.C. (2009) Food webs are more than the sum of their tritrophic parts. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 22335–22340.

DeLorenzo M.E., Scott G.I. & Ross P.E. (1999) Effects of the agricultural pesticides atrazine, deethylatrazine, endosulfan, and chlorpyrifos on an estuarine microbial food web. *Environmental Toxicology and Chemistry* **18**, 2824–2835.

Downing A.L., DeVanna K.M., Rubeck-Schurtz C.N., Tuhela L. & Grunkemeyer H. (2008) Community and ecosystem responses to a pulsed pesticide disturbance in freshwater ecosystems. *Ecotoxicology* **17**, 539–548.

Gessner M.O. & Chauvet E. (2002) A case for using litter breakdown to assess functional stream integrity. *Ecological Applications* **12**, 498–510.

Giddings J.M., Williams W.M., Solomon K.R. & Giesy J.P. (2014) Risks to aquatic organisms from use of chlorpyrifos in the United States. *Reviews of Environmental Contamination and Toxicology* **231**, 119–162.

Gray C., Baird D.J., Baumgartner S., Jacob U., Jenkins G.B., O'Gorman E.J., et al. (2014) FORUM: Ecological networks: the missing links in biomonitoring science. *Journal of Applied Ecology* **51**, 1444–1449.

Halstead N.T., McMahon T.A., Johnson S.A., Raffel T.R., Romansic J.M., Crumrine P.W., et al. (2014) Community ecology theory predicts the effects of agrochemical mixtures on aquatic biodiversity and ecosystem properties. *Ecology Letters* **17**, 932–941.

Hillebrand H., Dürselen C.D., Kirschtel D., Pollingher U. & Zohary T. (1999) Biovolume calculation for pelagic and benthic microalgae. *Journal of Phycology* **35**, 403–424.

Hladyz S., Åbjörnsson K., Chauvet E., Dobson M., Elosegi A., Ferreira V., et al. (2011) Stream ecosystem functioning in an agricultural landscape: the importance of terrestrial-aquatic linkages. *Advances in Ecological Research* **44**, 211–276.

Hothorn T., Bretz F., Westfall P., Heiberger R.M. & Schuetzenmeister A. (2014) Multcomp: simultaneous inference in general parametric models. R package version, 1.3–2.

Hudson L.N., Emerson R., Jenkins G.B., Layer K., Ledger M.E., Pichler D.E., et al. (2012) Cheddar: analysis and visualisation of ecological communities in R. *Methods in Ecology and Evolution* **4**, 99–104.

Kefford B.J., Schäfer R.B., Liess M., Goonan P., Metzeling L. & Nugegoda D. (2010) A similarity-index-based method to estimate chemical concentration limits

protective for ecological communities. *Environmental Toxicology and Chemistry* **29**, 2123–2131.

Kersting K. & Van den Brink P.J. (1997) Effects of the insecticide Dursban® 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: responses of ecosystem metabolism. *Environmental Toxicology and Chemistry* **16**, 251–259.

Kohler H.R. & Triebskorn R. (2013) Wildlife ecotoxicology of pesticides: can we track effects to the population level and beyond? *Science* **341**, 759–765.

Krammer K. & Bertalot L.H. (1986) Bacillariophyceae, 1. Teil: Naviculaceae. In Süßwasserflora von Mitteleuropa (eds H. Ettl, J. Gerloff, H. Heynig & D. Mollenhauer). Gustav Fischer Verlag, Stuttgart, Germany.

Krammer K. & Bertalot L.H. (1986) Bacillariophyceae, 1. Teil: Naviculaceae. In Süßwasserflora von Mitteleuropa (Eds H. Ettl, J. Gerloff, H. Heynig & D. Mollenhauer). Gustav Fischer Verlag, Stuttgart, Germany.

Krammer K., Lange-Bertalot H., Krammer K., Lange-Bertalot H., Bate N., Podzorski A., et al. (1986) Bacillariophyceae, 2. Teil: Bacillariaceae, Epithemiaceae, Surirellaceae. In Süßwasserflora von Mitteleuropa (Eds H. Ettl, J. Gerloff, H. Heynig & D. Mollenhauer). Gustav Fischer Verlag, Stuttgart, Germany.

Krammer K. & Lange-Bertalot H. (1991a) Bacillariophyceae, 3. Teil: Centrales, Fragilariaceae, Eunotiaceae. In Süßwasserflora von Mitteleuropa (Eds H. Ettl, J. Gerloff, H. Heynig & D. Mollenhauer). Gustav Fischer Verlag, Stuttgart, Germany.

Krammer K. & Lange-Bertalot H. (1991b) Bacillariophyceae, 4. Teil: Achnanthaceae, Kritische Ergänzungen zu Navicula (Lineolatae) und Gomphonema. In Süßwasserflora von Mitteleuropa (Eds H. Ettl, J. Gerloff, H. Heynig & D. Mollenhauer). Gustav Fischer Verlag, Stuttgart, Germany.

Kwak Y., Kim S.-J., Rhee I.-K. & Shin J.-H. (2012) Application of quantitative real-time polymerase chain reaction on the assessment of organophosphorus compound degradation in in situ soil. *Journal of the Korean Society for Applied Biological Chemistry* **55**, 757–763.

Layer K., Hildrew A.G. & Woodward G. (2013) Grazing and detritivory in 20 stream food webs across a broad pH gradient. *Oecologia* **171**, 459–471.

Layer K., Hildrew A.G., Jenkins G.B., Riede J.O., Rossiter S.J., Townsend C.R., et al. (2011) Long-term dynamics of a well-characterised food web: four decades of acidification and recovery in the Broadstone Stream model system. *Advances in Ecological Research* **44**, 69–117.

Layer K., Riede J.O., Hildrew A.G. & Woodward G. (2010) Food web structure and stability in 20 streams across a wide ph gradient. *Advances in Ecological Research* **42**, 265–299.

Ledger M.E. & Hildrew A.G. (2005) The ecology of acidification and recovery: changes in herbivore-algal food web linkages across a stream pH gradient. *Environmental Pollution* **137**, 103–118.

Liess M. & Ohe von der P.C. (2005) Analyzing effects of pesticides on invertebrate communities in streams. *Environmental Toxicology and Chemistry* **24**, 954–965.

Liess M. & Schulz R. (1999) Linking insecticide contamination and population response in an agricultural stream. *Environmental Toxicology and Chemistry* **18**, 1948–1955.

Liess M., Schäfer R.B. & Schriever C.A. (2008) The footprint of pesticide stress in communities. *Science of the Total Environment* **406**, 484–490.

Malaj E., Ohe von der P.C., Grote M., Kühne R., Mondy C.P., Usseglio-Polatera P., et al. (2014) Organic chemicals jeopardize the health of freshwater ecosystems on the continental scale. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 9549–9554.

Maltby L. & Hills L. (2008) Spray drift of pesticides and stream macroinvertebrates: experimental evidence of impacts and effectiveness of mitigation measures. *Environmental Pollution* **156**, 1112–1120.

Ogle D.H. (2012) FSA: Fisheries stock analysis. R package version 0.2–8.

Papst M.H. & Boyer M.G. (1980) Effects of two organophosphorus insecticides on the chlorophyll a and pheopigment concentrations of standing ponds. *Hydrobiologia* **69**, 245–250.

Pinheiro J., Bates D., DebRoy S. & Sarkar D. (2011) R Development Core Team. 2010. nlme: linear and nonlinear mixed effects models. R package version 3.1-97. R Foundation for Statistical Computing, Vienna, Austria.

Raven P.J. & George J.J. (1989) Recovery by riffle macroinvertebrates in a river after a major accidental spillage of chlorpyrifos. *Environmental Pollution* **59**, 55–70.

Rawcliffe R., Sayer C.D., Woodward G., Grey J., Davidson T.A. & Jones I.J. (2010) Back to the future: using palaeolimnology to infer long-term changes in shallow lake food webs. *Freshwater Biology* **55**, 600–613.

Relyea R.A. (2008) A cocktail of contaminants: how mixtures of pesticides at low concentrations affect aquatic communities. *Oecologia* **159**, 363–376.

Rocha O. & Duncan A. (1985) The relationship between cell carbon and cell volume in

freshwater algal species used in zooplanktonic studies. *Journal of Plankton Research* **7**, 279–294.

Rosi-Marshall E.J., Kincaid D.W., Bechtold H.A., Royer T.V., Rojas M. & Kelly J.J. (2013) Pharmaceuticals suppress algal growth and microbial respiration and alter bacterial communities in stream biofilms. *Ecological Applications* **23**, 583–593.

Rubach M.N., Crum S.J. & Van den Brink P.J. (2011) Variability in the dynamics of mortality and immobility responses of freshwater arthropods exposed to chlorpyrifos. *Archives of Environmental Contamination and Toxicology* **60**, 708–721.

Schäfer R.B., Caquet T., Siimes K., Mueller R., Lagadic L. & Liess M. (2007) Effects of pesticides on community structure and ecosystem functions in agricultural streams of three biogeographical regions in Europe. *Science of the Total Environment* **382**, 272–285.

Schinegger R., Trautwein C., Melcher A. & Schmutz S. (2011) Multiple human pressures and their spatial patterns in European running waters. *Water and Environment Journal* **26**, 261–273.

Sicko-Goad L.M., Schelske C.L. & Stoermer E.F. (1984) Estimation of intracellular carbon and silica content of diatoms from natural assemblages using morphometric techniques. *Limnology and Oceanography* **29**, 1170–1178.

Stehle S. & Schulz R. (2015) Agricultural insecticides threaten surface waters at the global scale. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 5750–5755.

Traas T.P., Janse J.H., Van den Brink P.J., Brock T.C.M. & Aldenberg T. (2004) A freshwater food web model for the combined effects of nutrients and insecticide stress and subsequent recovery. *Environmental Toxicology and Chemistry* **23**, 521–529.

Triebskorn R., Adam S., Behrens A., Beier S., Böhmer J., Braunbeck T., et al. (2003) Establishing causality between pollution and effects at different levels of biological organization: The VALIMAR project. *Human and Ecological Risk Assessment* **9**, 171–194.

Van den Brink P.J., van Donk E., Gylstra R., Crum S.J. & Brock T.C. (1995) Effects of chronic low concentrations of the pesticides chlorpyrifos and atrazine in indoor freshwater microcosms. *Chemosphere* **31**, 3181–3200.

Van Wijngaarden R., Van den Brink P.J., Crum S.J., Brock T., Leeuwangh P. & Voshaar O.J.H. (1996) Effects of the insecticide Dursban® 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: I. Comparison of short-term

toxicity between the laboratory and the field. *Environmental Toxicology and Chemistry* **15**, 1133–1142.

Woodward G., Gessner M.O., Giller P.S., Gulis V., Hladyz S., Lecerf A., et al. (2012) Continental-scale effects of nutrient pollution on stream ecosystem functioning. *Science* **336**, 1438–1440.

Woodward G., Papantoniou G., Edwards F. & Lauridsen R.B. (2008) Trophic trickles and cascades in a complex food web: impacts of a keystone predator on stream community structure and ecosystem processes. *Oikos* **117**, 683–692.

Wright J.F., Clarke R.T., Gunn R., Kneebone N.T. & Davy-Bowker J. (2004) Impact of major changes in flow regime on the macroinvertebrate assemblages of four chalk stream sites, 1997–2001. *River Research and Applications* **20**, 775–794.

Wright J.F., Gunn R.J.M., Winder J.M., Wiggers R., Vowles K., Clarke R.T., et al. (2002) A comparison of the macrophyte cover and macroinvertebrate fauna at three sites on the River Kennet in the mid 1970s and late 1990s. *Science of the Total Environment* **282**-283, 121–142.

Xuereb B., Noury P., Felten V., Garric J. & Geffard O. (2007) Cholinesterase activity in Gammarus pulex(Crustacea Amphipoda): characterization and effects of chlorpyrifos. *Toxicology* **236**, 178–189.

Young R.G., Matthaei C.D. & Townsend C.R. (2008) Organic matter breakdown and ecosystem metabolism: functional indicators for assessing river ecosystem health. *Journal of the North American Benthological Society* **27**, 605–625.

Zuur A., Ieno E.N., Walker N., Saveliev A.A. & Smith G.M. (2009) Mixed Effects Models and Extensions in Ecology with R. Springer Science & Business Media.

## C.7 Supporting Information

Additional Supporting Information may be found in the online version of this article:

**Appendix S1**. Supplementary material. Table S1. Mean numerical abundance at control and impacted sites for diatoms in the trivariate food webs shown in Figure C.Dd.

**Table S2**. Mean numerical abundance at control and impacted sites for macroinvertebrates and fishes in the trivariate food webs shown in Figure C.Dd.

**Table S3**. Shapes of diatom species used to calculate biovolumes (Hillebrand et al., 1999).

**Table S4.** Equations used to calculate macroinvertebrate individual dry mass (DM).

**Table S5**. Sources of feeding interactions derived from the primary literature.

**Table S6.** The taxonomic resolution (i.e. generality) assigned to each node in the networks to create links between nodes.

**Figure S1**. UK Environment Agency water chemistry data and macroinvertebrate data collected by citizen scientists between July 2012 and July 2014. Water chem- istry samples were collected from an upstream control (grey; adjacent to site B) and a downstream impacted monitoring station (black; adjacent to site E); citizen science macroinvertebrate samples were collected from a control site at Stonebridge Lane and an impacted site at Elcot Mill (Figure C.A).

# Appendix D | Freshwater conservation and biomonitoring of structure and function: genes to ecosystems[1]

[1] **Gray, C**., Bista, I., Creer, S., Demars, B.O.L., Falciani, F., Monteith, D.T., Sun, X. & Woodward, G. (2015). Freshwater conservation and biomonitoring of structure and function: genes to ecosystems. Aquatic Functional Biodiversity: Ecological and Evolutionary Approaches (eds A. Belgrano, G. Woodward & U. Jacob). Elsevier.

## D.1 Summary

Biomonitoring and conservation of freshwaters to date have fallen short of incorporating a fully ecological and evolutionary perspective. Due to this, the predictive capacity of current biomonitoring approaches is restricted and will have a limited ability to adapt in the face of rapid and global habitat modification and climate change. We briefly outline the present state of biomonitoring as well as some of its limitations. We then address how incorporating an ecological and evolutionary approach to biomonitoring and conservation will allow us to better understand interactions between the evolution and ecology of a species. This approach alongside the incorporation of measures of ecosystem functioning and aided by new technologies such as novel molecular markers or the use of microbes, may facilitate the future development of a more comprehensive and effective biomonitoring framework.

## D.2 Current focus of aquatic biomonitoring and conservation

Freshwater biomonitoring, i.e. the repeated, quantitative assessment of surface waters using the presence and/or abundance of groups of organisms of known environmental sensitivity, currently provides a staple tool in aquatic management and conservation, and underpins wide-reaching environmental legislation including the European Union Water Framework Directive (EU WFD), Environmental Quality Standards for Surface Water in China (GB 3838-2002) and the Clean Water Act in the United States of America. Its scientific origins can be traced back to societal changes during the industrialisation of the developed world and simultaneous scientific developments in epidemiology and biological taxonomy - the impacts of rising human populations on the chemical and microbiological quality of urban water supplies necessitated the development of rapid and robust methods to assess risks to public health.

The history of aquatic biomonitoring is extensively reviewed elsewhere (e.g. Metcalfe 1989; Rosenberg & Resh 1993; Friberg et al. 2011) and so will not be discussed in detail here but essentially biomonitoring hinges on two basic concepts: first that aquatic organisms tend to be unevenly distributed across environmental gradients, and should therefore have value as indicators of ecosystem state, and second, the biota provide a more temporally integrated indication of ecosystem quality than many abiotic measurements, such as spot sampled water chemistry.

Three key developments over the course of the 20th century had major impacts on routine environmental assessment by regulatory authorities (Metcalfe 1989). First, Kolkwitz and Marsson (1902; 1909), introduced what became the "saprobien system", in which groups of organisms were directly linked with perceived discrete levels of organic contamination and by inference, oxygen availability of waters. Second, biological diversity indices became popular around the middle of the century, based largely on the premise that species richness and evenness is reduced with increasing environmental disturbance. Finally, biotic indices that combined these methodologies

(such as the Trent Biotic Index, Chandler's Score System and the Biological Monitoring Working Party) were developed. Despite the widespread adoption of these indices (in particular the Average Score per Taxon (ASPT) approach), many surface waters are more likely to be compromised by other anthropogenic stressors, such as acidification, toxins, climate change, atmospheric deposition of reactive nitrogen and habitat modification.

During the 1980s, the need to understand the causes behind surface water acidification stimulated investigation of diatoms as palaeobiological assessment tools (Renberg & Hellberg 1982; Battarbee & Charles 1986). These ubiquitous and chemically sensitive unicellular algae preserve well in lake sediments, thus enabling palaeo-ecologists to reconstruct the environmental history of a water body from sediment cores. Statistical approaches based on weighted averaging procedures were developed to predict (or hindcast) lake chemistry on the basis of spatially derived "training sets" describing the chemical "optima" and tolerances of individual species (e.g. Birks et al. 1990). This approach has proved highly effective in the reconstruction of lake pH and has been applied to infer historical change in other environmental parameters with more mixed success. More recently various community-based multivariate regression approaches have been developed to interpret the environmental significance of trends in contemporarily monitored biota, including diatoms and macroinvertebrates (Monteith et al. 2005; Murphy et al. 2012) and to specifically address the extent to which biological trends can be explained by changes in water quality with time (Halvorsen et al. 2003).

In recent years, more effective water treatment regimes and environmental regulations have improved surface water quality with respect to both organic pollution and water acidity in much of the developed world. The focus of biomonitoring has consequently begun to shift from basic quantification of environmental damage to consideration of how much surface water quality, with respect to these key drivers, still deviates from a desired "reference" condition relative

to a "pristine" state. The bio-assessment tool RIVPACS (River Invertebrate Prediction and Classification System) pioneered this field, by quantifying the differences in the macroinvertebrate assemblage between a site under investigation relative to its "expected" assemblage at unimpacted, but otherwise comparable sites. This approach and its derivatives now underpin most freshwater biomonitoring schemes across Europe (e.g. Simpson et al. 2005; Murphy et al. 2013) and other parts of the world (Simpson & Norris, 2000).

Unfortunately, despite these advances, assigning appropriate reference conditions and current status is still problematic, as pre-industrial (i.e. pre 1800) target conditions are very difficult to model with confidence (Battarbee et al. 2005), as there are rarely useful palaeoecological data from running waters because therr sediments are well-mixed and there are also mismatches between palaeo and contemporay data in standing waters as the two rarely overlap in time, so ground-truthing is difficult. A notable exception is from some of the longer-term biomonitoring schemes, such as the United Kingdom Acid Waters Monitoring Network (Monteith et al 2005; Battarbee et al 2014), where, after several decades of lake biomonitoring using sediment traps, we are now finally able to compare contemporaneously collected data directly with palaeoecological data (Figure D.A). This has raised intriguing questions about stressor impacts: for instance in the Acid Waters Monitoring Network (AWMN) data, the lack of evidence of clear recovery among diatom communities along the acidification trajectory evident in the sediment core records (despite improvements in water chemistry), points to hystereses in these ecosystems, and to the potential ecological importance of other factors that could be setting new environmental states that and may not be reversed in the foreseeable future (Battarbee et al. 2013). The growing realisation that a return to a historical pre-impacted state may be unrealistic is now forcing us to consider shifting environmental baselines when assessing conservation and restoration, and determination of when an alternative state is acceptable with respect to its function, biodiversity and the ecosystem services it provides (UK

National Ecosystem Assessment, 2011; Millenium Ecosystem Assessment 2005). While this palaeoecological reference approach to aquatic monitoring is limited to lake ecosystems (as running water sediments are turned over), there is considerable potential to extend it to other biological proxies and biogeochemical indicators, such as pigments and stable isotopes, and pressures other than acidification (e.g. Smol 2009).

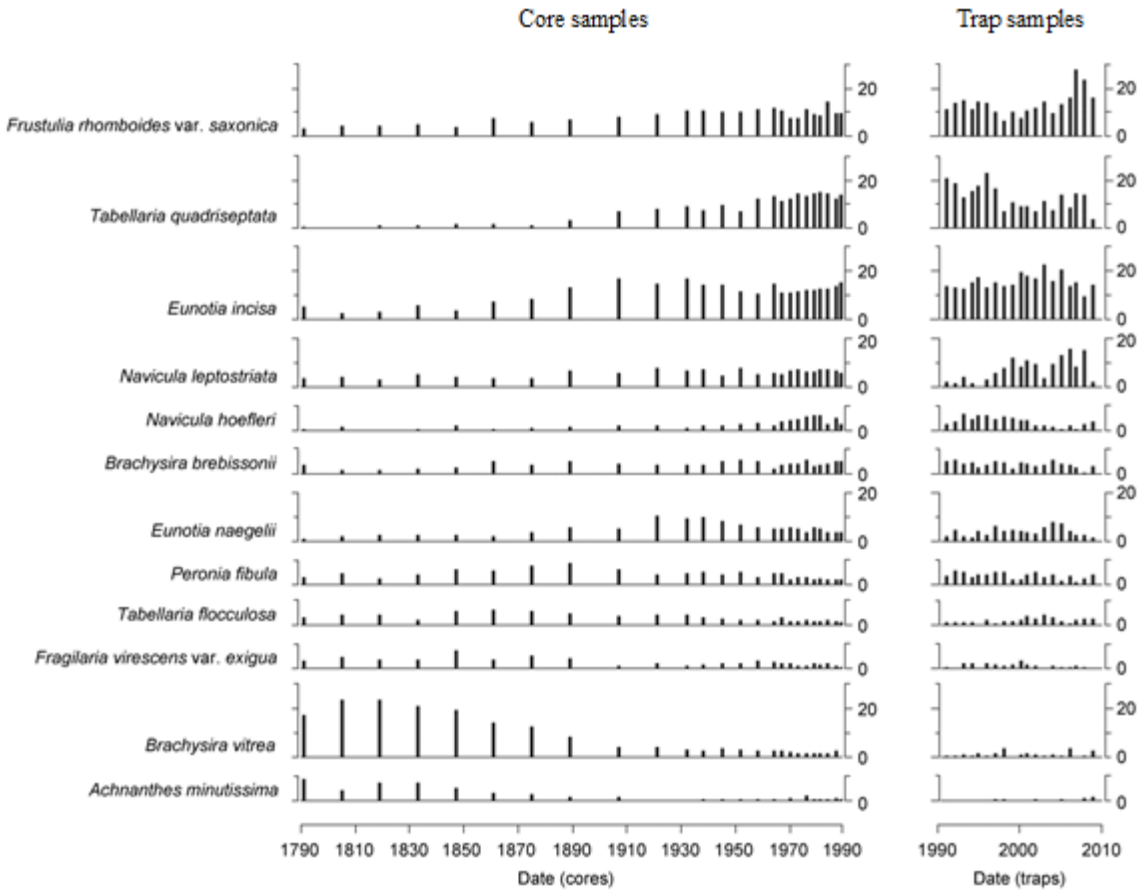



**Figure D.A**. Linking a sites contemporary biomonitoring data to its historical, reference condition (redrawn from data presented in (Battarbee et al. 2014). Percentage relative abundances of diatom species found in sediments of a UK upland lake (Round Loch of Glenhead). Species abundances in historical sediment core samples (left) shift from left to right reflecting increased water acidity during the industrial revolution. Abundances of the same species in contemporary sediment trap assemblages (right) indicate some recent reversal (decline) of some particularly acid-loving species, e.g. Tabellaria quadriseptata, as acidity has declined. However other species that increased during acidification are continuing to increase in abundance while others that were common prior to acidification show little indication of recovery.

All these approaches focus on linking attributes of biological assemblages to a system's chemical or physical state and they have made important contributions to environmental assessment, policy and legislation across ecological and evolutionary timescales. The power of these methodologies can be largely attributed to the wide variation between taxa in tolerance to specific pressures, in particular the bio-availability of oxygen, hydrogen and aluminium ions. Newly emerging environmental threats, such as the many facets of climate change, contamination from organic micropollutants and nanoparticles etc., may not be quite so readily assessed by similar direct environment-taxa calibration-based approaches (Figure D.B) (Friberg et al. 2011). In some cases, other ecosystem metrics, other than the relative abundance of taxa, may yield clearer insights into significant environmental shifts (e.g. Layer et al. 2011). There is, therefore, a growing need to determine how best to assess the impact of these emerging stressors, both in isolation and in combination. Also, the structural biodiversity-centric focus of these traditional methods now needs to be augmented with more explicitly functional measures, to provide complementary insights into the impacts of stressors in freshwater ecosystems (e.g. Woodward et al. 2012).

In addition to largely lacking these explicitly functional ecosystem-level metrics, another common limitation of current taxonomic-based biomonitoring schemes is that, although there is an implicit evolutionary signal embedded within them, (i.e. in terms of the phylogenetic relatedness of the various indicator taxa, which constrains their functional traits), there is still no explicit recognition of the role of adaptation to new stressors and the potential for evolutionary rescue from stressors within species populations: and evolutionary responses can occur surprisingly quickly in many freshwater taxa (e.g. Melian et al. 2011). This could cause mismatches between the reference and impacted conditions, if species are able to adapt to new conditions, rather than acting as passive ciphers that are simply overlain on an environmental template (e.g. Bell & Gonzalez 2011). This has resulted in a paradox of

biomonitoring, in which speciation is the mechanism that produces the response variables we measure but which is then ignored when relating species distributions to environmental conditions. Although research is beginning to fill this gap in understanding (e.g. Thuiller et al. 2011; Vonlanthen et al. 2012) currently in biomonitoring, this "inconvenient truth" is either ignored or attempts are made to circumvent it by removing the phylogenetic signal from the data (e.g. via trait-based approaches).

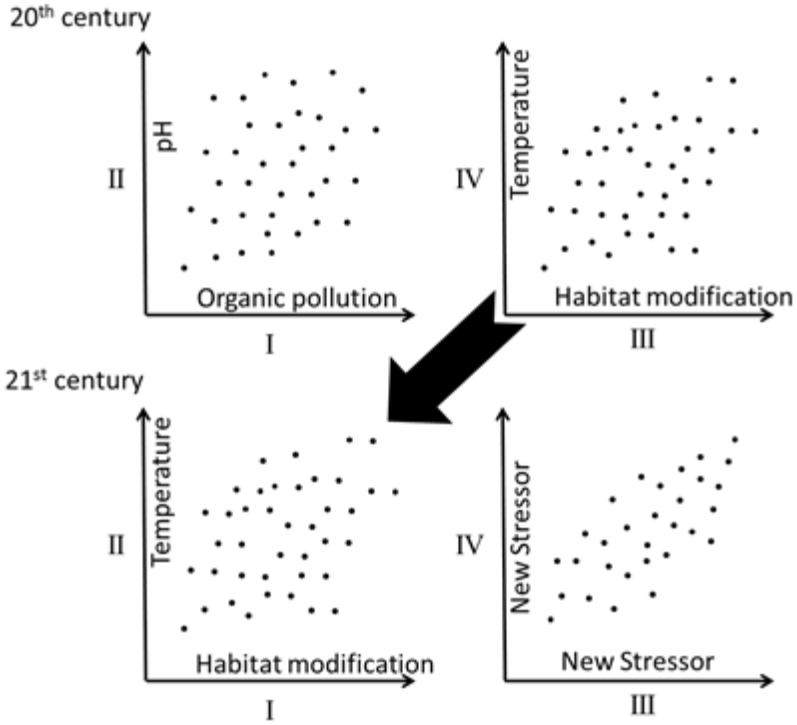



**Figure D.B**. A hypothetical ordination to show the changes in the main drivers of habitat degradation in freshwaters in the developed world over time. In the developed world, over time increasing temperature and habitat modification have become the significant drivers of change in the principal components (Axes I-IV) of community composition, replacing the more historical stressors of organic pollution and pH change. However these historical stressors are still the major causes of habitat degradation in developing countries.

## D.3 State of the art in the science of biomonitoring: from species traits to community and ecosystem

The earliest attempts to combine ecological and evolutionary approaches to biomonitoring included the use of additional measures of biodiversity including phylogenetic diversity (or taxonomic distinctness) and functional diversity conditioned by evolution (e.g. May 1990; Paradis, Claude & Strimmer 2004; Webb, Ackerly & Kembel 2011), though most of the emphasis has been on the former, not the latter. A problem with focusing solely on taxonomy is that if species redundancy is high, as appears to be the case in many freshwaters (e.g. McKie et al. 2008; Perkins et al. 2010; Reiss et al. 2010; Reiss et al. 2011), then species loss is likely to only have strong effects when entire functional guilds are lost: but it is these that we still have limited understanding of due to the longstanding reliance on more traditional measures of biodiversity (e.g., species richness). The realised species trait (or gene) profile at a local scale provides the means to link the potential effects of anthropogenic pressures on species (population) distribution and dynamics: i.e., the trait profile itself may therefore be used for diagnostic purposes (Statzner & Bêche 2010). It is possible, however, that non-causal relationships between individual species traits and contemporary environmental conditions exist (e.g. Poff et al. 2006; Horrigan & Baird 2008) because some traits may represent an evolutionary legacy rather than current adaptation (Gould & Lewontin 1979). Empirical studies have confirmed the large role played by phylogeny or taxonomic distinctness in freshwater ecosystems (Willby, Abernethy & Demars 2000; Poff et al. 2006; Demars et al. 2012) from the structural perspective, but their functional attributes remain far less well-understood.

To interpret biomonitoring results (patterns in species composition) it is crucial to unravel its underlying mechanistic basis (processes which determine this pattern, both anthropogenically mediated or not). Species are not randomly distributed in time (e.g. Lyell & Deshayes 1830) or space (e.g. Humboldt 1849) and Demars and Edwards (2009) recently pointed out that even as far back as in the 19th Century Darwin (1872)

argued that environmental variables only played a subordinate role in the determination of species distribution. He offered a mechanistic explanation (pp. 318–319): immigration of individuals from a species (individuals) pool controlled by dispersal barriers and descent with modification regulated through natural selection, with competition being the most important pressure. He attributed the wide distribution of freshwater organisms to favourable means of dispersal (Darwin, 1872, pp. 323–330, 343–347, e.g., Pollux and Santamaria et al., 2005) and lessened competition (Darwin, 1872, pp.346, e.g., Greulich and Bornette, 2003) in aquatic habitats. This debate of whether species distribution is more controlled by niche assembly (resource heterogeneity) or dispersal assembly is still on-going (Demars & Harper 2005; Heino 2013). Moreover, numerous null models have reproduced biomonitoring patterns of species assembly: e.g. random (Tokeshi 1990), niche (Tokeshi 1993), neutral (Bell 2001; Hubbell 2001), metabolic scaling (Allen, Brown & Gillooly 2002), fractal (Lennon et al. 2007), maximum entropy (Harte 2011).

The general consensus is that patterns in species composition and community structure emerge from the interactions of chance, dispersal and resource heterogeneity in evolving meta-communities (Venail et al. 2008). This is supported by empirical studies using autocorrelation, spatial distances/isolation and dispersal abilities to infer proportion of resource (niche) versus dispersal community assembly (Moilanen & Hanski 2001; Demars & Harper 2005; Moilanen et al. 2005; Moilanen, Leathwick & Elith 2008; Bonada, Dolédec & Statzner 2012). Essentially, this is explicitly adding the otherwise overlooked dynamical component to biomonitoring data, which are often seen as static snapshots whereby species simply map onto the environmental template. It also starts to recognise the inherent role of dispersal and selection for particular functional traits, rather than simply focusing on the phylogenetic tree in isolation.

Every species can be characterised by not only its taxonomic identity but also its biological (response) functional traits, which may be translated into functional

(effect) traits (Engelhardt 2006; Kerkhoff & Enquist 2006; López-Urrutia et al. 2006; Enquist et al. 2007) and eventually into ecosystem services (e.g. García-Llorente et al. 2011). Mapping traits onto the tree of life reveals a convergence (independent appearance of a trait in separate clades) or divergence (appearance of a trait in a single clade) in evolution. This is highly relevant in the context of the insurance hypothesis or portfolio effect, whereby high species (or genetic) richness maintains high and constant ecosystem (or population) productivity and services in a stochastic environment (Yachi & Loreau 1999; Schindler et al. 2010).

**BOX 1: Categorizing continuous variables in biomonitoring**

Figure D.C maps an example of a continuous ecological variable (habitat quality) onto discrete man-made categories. This human need to categorize complexity can be seen in many aspects of ecology, not just in the biomonitoring and conservation fields. Whether it's the difficulties encountered when classifying all of life on earth into discrete species (e.g. Mayden 1997), or the questionable practice of assigning 'typologies' to a given lake or river (e.g. Friberg *et al.* 2011), the motivation comes from our historically poor ability to process large amounts of complex information. However, this process of classification and simplification has allowed us to make some informed generalisations and useful interpretations that otherwise would not be possible. Nevertheless, with the advent of rapidly accelerating computing power the challenge has now shifted away from our previous inability to process complex information, to the interpretation of complex information into simple messages. With expanding analytical ability comes the need to preserve as much ecological information as possible, which will allow a deeper understanding and more informed interpretations to develop the next necessary steps forward in biomonitoring science; the shift of focus away from the simple monitoring of species composition towards the monitoring of ecosystem functions and services.

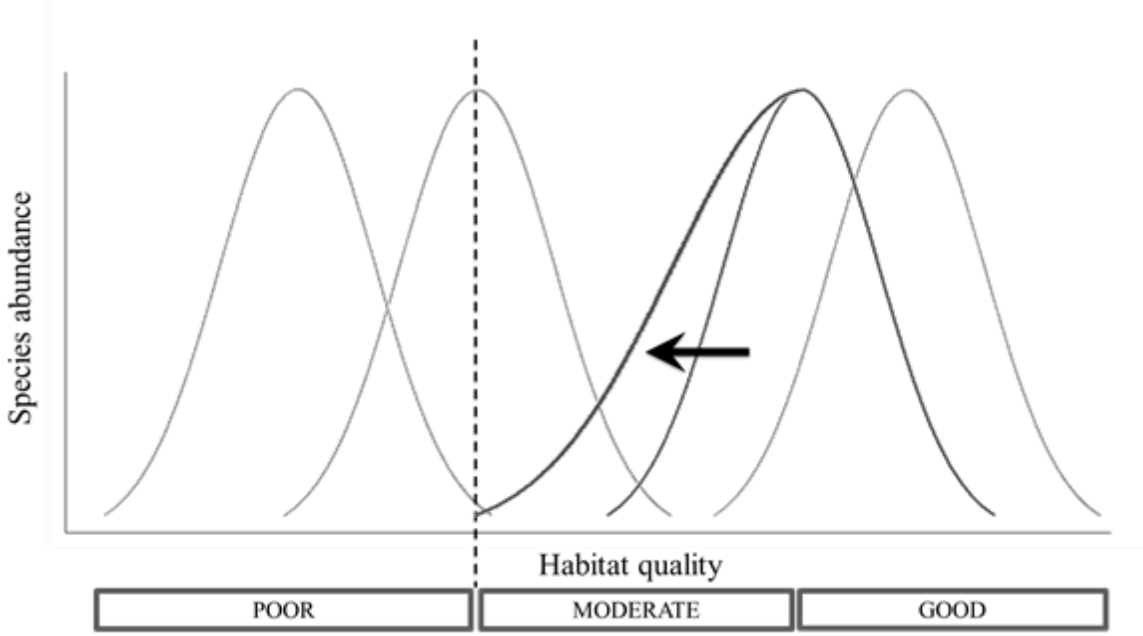



**Figure D.C.** Hypothetical graph showing fluctuations in four species abundances across a habitat quality gradient, alongside the discrete criteria of habitat quality (good, moderate, poor) that these continuous variables are categorized into. The dashed line shows species loss, whereas the solid black arrow shows sub-lethal effects to a particular species population.

The ecology of a species sets the scene in which evolution operates, whilst evolution may influence ecological dynamics by altering the frequency of phenotypes that are available to interact: thus, there are potentially important eco-evolutionary feedbacks, which are only now starting to be recognised (e.g. Melián et al. 2011; Moya-Larano et al. 2012). The ability of a species to adapt to a changing environment is key to how it responds to stressors: species are not simply present or absent if environmental conditions are favourable or unfavourable (Text Box 1). According to the old adage, there are three options - "adapt, perish or move" - that a species is faced with in a changing environment, yet biomonitoring and conservation schemes have largely ignored the first.

An important issue here is that neither ecological nor evolutionary responses occur solely at the population level of organisation: no species is an island, and its interactions with those around it will determine both species-specific and the wider community's responses to changing conditions (e.g. Rybicki & Landwehr 2007). This explains why models derived from bioclimate envelopes and extrapolations from traditional biomonitoring techniques often fail to predict species responses in the real world, because their synecology (the ecology of communities of interacting organisms) is ignored (Woodward et al. 2010; Friberg et al. 2011). The use of trait-based approaches helps to grapple with issues related to functional biodiversity at the autecological level, but it fails to embrace the more complex, higher-level synecological functional roles that species play within multispecies systems such as food webs, which may have seemingly unpredictable emergent properties (Woodward 2009). This can be exemplified by mismatches between real-time or experimental data that track transient dynamics, versus space-for-time substitutions where the different communities across the environmental gradient may already be at equilibrium (e.g. Layer et al. 2010; Layer et al. 2011). Unfortunately, such data are still rare, but where they are available there is compelling evidence that the functional role of species within the food web can have important indirect and direct

consequences that would be missed by relying on static data: a classic example is the seeming paradox of invertebrate abundance declining over several decades of deacidification, yet this response makes sense when the top-down effects of predators on the prey assemblage are included (Layer et al. 2011).

Figure D.D synthesises current thinking in the role of ecology and evolution of species distribution on which taxonomic, functional and phylogenetic diversity determine the dynamics of ecosystem functioning and services, and highlights how they can be integrated in future biomonitoring approaches.

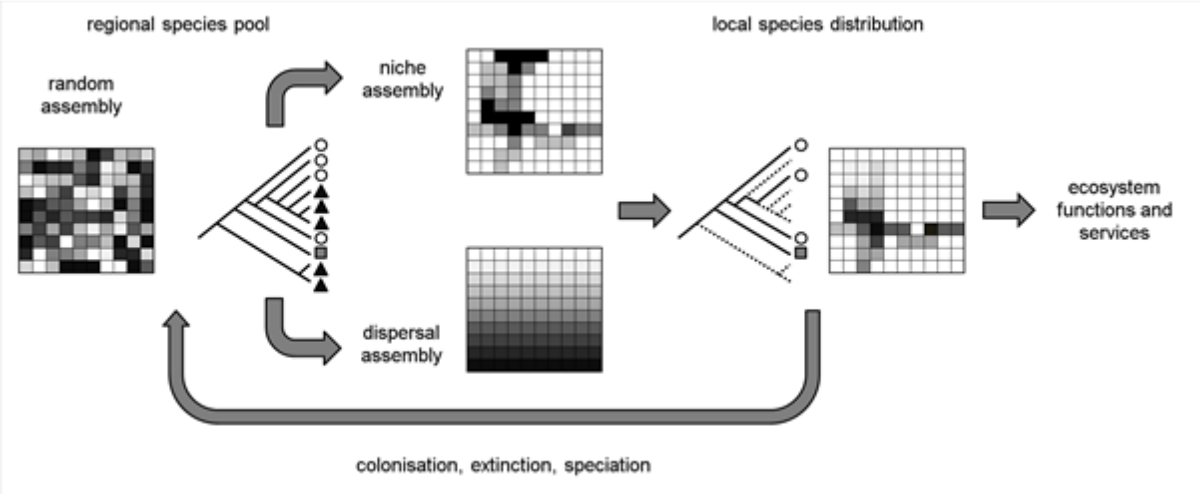



**Figure D.D.** Ecology and evolution of species distribution generates diversity patterns in species (grids), species traits (symbols) and phylogeny (trees). From a hypothetical null model (e.g. random assemblage) and species pool at regional scale, species are sorted through the effects of niche assembly (heterogeneity of resources) and species dispersal into patterns of local species distribution. Over time, local extinction, colonisation and speciation alter the regional species pool and associated phylogeny and trait diversity. The dimensions of diversity: taxonomic, abundance, functional and phylogenetic, determine the dynamics of ecosystem functions and services.

Functional diversity provides a more direct link between species richness and ecosystem functioning, and ultimately the provision of goods and services (Naeem 2002; Woodward 2009). Two essential functions are primary production and

decomposition, which provide the two key energy inputs into any food web, thus ultimately driving the whole system's trophic dynamics, stability and productivity. Production and decomposition thereby provide a variety of services, including the production of fish in fisheries and for recreational angling, or the processing of pollutants and waste products to produce clean water. These vital ecosystem processes are, however, not routinely measured in current biomonitoring techniques. Decomposition rates have been measured in some large-scale studies, but these too are still largely ignored in routine biomonitoring, and the responses remain complex and poorly understood (Woodward et al. 2012). Some functional measures, such as organic matter decomposition, has been the focus of attention (e.g. Young, Matthaei & Townsend 2008), and methods for standardising this measure across ecosystems have been developed (e.g. Kampfraath & Hunting et al. 2012) crucially allowing comparisons between studies, but these methods are yet to be adopted into biomonitoring schemes.

Functional indicators, and especially direct measures of ecosystem processes, should also play a larger role to quantify ecosystem services (Millenium Ecosystem Assessment 2005), which are being advocated increasingly for economic valuations of conservation, management and restoration projects (Costanza et al. 1997; Everard & McInnes 2013). Many ecosystem processes are either services in their own right (e.g. carbon sequestration, nutrient cycling), or they underpin them (e.g. invertebrate production supporting fisheries), and include hydraulic retention (water transient storage), sedimentation rate, and greenhouse gas transfer. The magnitude and rate of many of these processes are sensitive to anthropogenic pressures, highlighting the scope to use functional indicators as diagnostic tools (Odum 1969; Schindler 1987; Sweeney et al. 2004; Mulholland et al. 2008; Yvon-Durocher et al. 2010; Demars et al. 2011).

Important insights into ecological and evolutionary responses to stressors, as well as their functional consequences could be inferred from the large amounts of geo-

referenced and dated lists of taxa currently filling a multitude of databases in local regulatory and conservation agencies, natural history and conservation societies. Many databases are now being assembled that contain some or all of these elements (e.g. FishBase [Frose & Pauly 2010] and Freshwater Life - http://www.freshwaterlife.org - supported by the Freshwater Biological Association). Scientists are collating decades of research to assemble species traits (and genes) in a phylogenetic context. Combining this with environmental data available from a wide range of government agencies and research bodies, and by organising this information into user-friendly databases (e.g. the Global Biotraits Database http://biotraits.ucla.edu/index.php) and connecting them to infer processes from patterns offers great potential for future research (e.g. Demars & Harper 2005; Demars & Trémolières 2009).

The success of the next generation of biomonitoring will not come solely from assembling and interrogating these vast new databases to obtain new response variables, but also from explicitly testing ecological hypotheses and synthesising different branches of science, e.g. eco-enzymatic stoichiometry which allows us to link the elemental composition of microbial communities to their nutrient content and biomass production (Sinsabaugh, Hill & Shah 2009; Hill et al. 2012). Integrating biomonitoring schemes with experimental and modelling approaches will be crucial: combining whole ecosystem experiments with long-term monitoring can reveal spectacular responses to environmental change, although such large-scale, long-term studies are still very much in the minority. Classic examples include the work of Likens et al. (1977) at the Hubbard Brook Experimental Forest, Schindler (1990), Carpenter et al. (2001) at the Experimental Lakes Area (ELA) in Canada (http://www.experimentallakesarea.ca) and Slavik et al. (2004) at the Kuparuk River station of the Long-Term Ecological Research (LTER) network. Other work has made use of these long-term data to develop new dynamical models to link biodiversity change to ecosystem functioning, such as Petchey et al. (2004) study based on the

extensive time series data from the UKs Environmental Change Network. Recently, the American LTER network has been complemented by the National Ecological Observatory Network, NEON (http://www.neoninc.org/news/lterandneon), and the STReam Experimental Observatory Network (STREON, part of NEON) is now the one of the most ambitious long term biomonitoring schemes. It combines comparative surveys across the USA with experimental design (nutrient enrichment and removal of large consumers) that extends previous LYNX programs (Mulholland et al. 2008). In the United Kingdom, the AWMN has also been very effective in providing scientific insights and influencing policy (Hildrew 2009; Layer et al. 2010; Friberg et al. 2011; Layer, Hildrew & Woodward 2013). Moreover, the value of AWMN has increased progressively over the three decades since its inception, as more subtle long-term trends, such as responses to climate change, are now able to be detected. The challenge is now to establish international networks with global coverage to tackle planet scale issues (e.g. Global Lakes Ecological Observatory Network, GLEON), which are also integrated with regional and local monitoring. Long-term monitoring can enable us to detect early warning signals of ecosystem shifts (Scheffer et al. 2009), but it is often difficult to extract research funding for such strategic research, which often appears to fail to meet the "novelty" criteria of many research councils' remits.

## D.4 Future advances and new perspectives – genes to ecosystems

Over the last 20 years huge progress has been made in understanding biodiversity-ecosystem functioning (B-EF) relationships, with an increasing emphasis on freshwater systems over the last decade in particular (Loreau, Naeem & Inchausti 2002; Woodward 2009; Loreau 2010; Reiss et al. 2010). Whilst biomonitoring and conservation have tended to focus on the biodiversity end of the relationship, the

functioning part of the equation and its relationship with biodiversity has been largely ignored in the more applied fields of freshwater ecology (but see Dangles et al. 2004; Cardinale 2011). However, the lack of functional insights is changing, and many emerging legislative and regulatory frameworks are recognising the need for more functional approaches (e.g. the Water Framework Directive). The main finding of B-EF research to date has been the prevalence of high levels of redundancy. Species loss may have initially little impact, but once a critical threshold is passed when entire functional groups are lost, the impacts can be extremely powerful and sensitive to further species loss (Cardinale et al. 2006). These experiments have also revealed evidence of idiosyncratic species responses being important, harking back to earlier ideas about keystone species, where they have both strong and unique influences on a process. Despite these advances, there are still some glaring gaps in our knowledge: few studies have included more than one trophic level; most have measured just one process rather than functioning as a whole, and they have been conducted primarily in small experimental arenas over short timescales (Woodward 2009). As such, many B-EF experiments lack the complexity of natural systems, though attempts are now being made to address these shortcomings (Reiss et al. 2010). In the context of moving from an understanding of B-EF to B-ES (biodiversity-ecosystem services) relationships, there is a huge gap to be bridged in terms of the spatiotemporal scales that are important for the latter, as ecosystem services tend to be manifested at much larger landscape scales, where source-sink, metacommunity and food web dynamics, as well as eco-evolutionary processes (e.g. Melián et al. 2011), are likely to be important.

The application of network-based approaches can be especially powerful here, as there is a strong food web context to where ecosystem services are located, as well as a clear trophic gradient in the scope for insurance and adaptation, which increases down the web's food chains (Figure D.E). Certain stressors are associated with particular nodes in the web (e.g. biomagnification of organochlorine pesticides in apex

predators; antibiotics with the microbial loop at the base of the web), as well as different organisational levels (e.g. food web modules; functional groups, the network as a whole) acting as multiple biosensors. For instance, allometries in food web properties from the level of pairwise links, to tritrophic food chains, to the system's entire constraint space have been used recently to evaluate responses of experimental stream food webs to drought (Woodward et al. 2012; Ledger et al. 2013): these revealed that many of the more commonly used network metrics (such as connectance) were relatively robust to perturbations, whereas others were much more sensitive (e.g. allometric scaling of pairwise links and food chains). The food web provides an intuitive prism through which to view both the lower and higher levels of organisation and how they respond to stressors, as it makes the interactions between species explicit in the response variables, whereas most biomonitoring and conservation approaches focus solely on (a few) nodes, and not the links between them at the system scale (Woodward, Gray & Baird 2013). Considerable work has been done in freshwaters in terms of understanding how food webs respond to stressors, including acidification (e.g. Ledger & Hildrew 2005; Layer et al. 2010; Layer et al. 2011), eutrophication (e.g. Rawcliffe et al. 2010) and hydrological change (e.g. Ledger et al. 2012; Ledger et al. 2013). Such combinations of studies illustrate effectively that studying the feedbacks between the environment and the functioning of the whole system that are mediated by the food web can be extremely powerful, and may even induce regime shifts (Jones & Sayer 2003; Scheffer & Carpenter 2003).

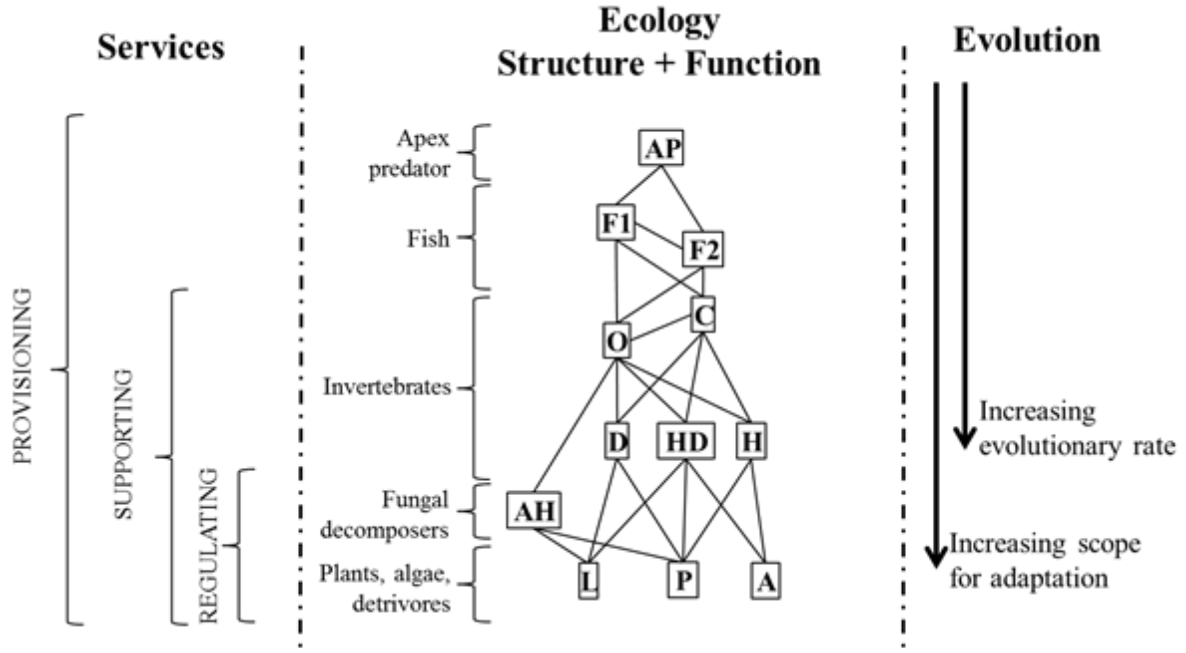



**Figure D.E.** Mapping services onto the food web. When monitoring services we need to monitor the appropriate level of scale. The effects of stressors upon services won't show at all levels of the food web, although may magnify through the food web, or cause trophic cascades. AP = apex predator, F = fish, C = carnivore, omnivore, D = detrivore, HD = herbivore/detritivore, H = herbivore, AH = aquatic hyphomycete, L = leaf-litter, P = plant, A = algae. Adapted from Abrahams et al. (2013) Figure 4.1.

Eco-evolutionary dynamics and feedbacks within the food web can be much faster than previously thought (e.g. Melián et al. 2011), and impacts on the epigenome can lead to quicker adaptation than traditional adaptation of the genome, via genetic plasticity (Johnson & Tricker 2010). Consequently, we are starting to perceive how species evolve in the context of both the biotic and abiotic environment, and how feedbacks and newly-discovered mechanisms can accelerate evolutionary responses (Moya-Larano et al. 2012). In addition to the discovery of these ecological and evolutionary interactions in recent years there have been rapid technological advances in Next Generation Sequencing (NGS, Text box 2) and associated molecular techniques (Hajibabaei et al. 2011; Hajibabaei 2012). This has allowed for significant advances in broadening the coverage of the tree of life and for adopting an eco-

evolutionary approach to biomonitoring in freshwaters: emerging NGS approaches include new generations of molecular markers, the ability to characterise microbes in situ, allowing them to be used to monitor the function of ecosystems as well as determining the function of microbes, metazoans and macrofaunal communities directly (Purdy et al. 2010).

## D.5 Novel molecular and microbial approaches

An organism's molecular state results from its interaction with the environment, and so measuring specific molecular machinery components can provide clues as to which stressors are present in the environment. The first generation of molecular markers (Figure D.F) were developed from hypothesis-driven research and based on biochemical, histological, morphological and physiological changes in nucleic acids and proteins measured with conventional techniques (Ryan & Hightower 1996). The number of such biomarkers is relatively small but they include some very effective examples, such as the general xenobiotic response marker CYP1A (Celander 2011) the endocrine disruption marker vitellogenin (Celander 2011) and the metal stress marker metallothionein (Amiard et al. 2006). However, the hypothesis driven approach to biomarker discovery suffers from an important conceptual flaw, at least in this implementation: single genes whose expression is modulated in a highly specific manner are extremely rare.

In the last ten years, new functional genomics technologies have provided a potential solution to this issue. Since they allow the measurement of the expression of tens of thousands of genes, proteins and metabolites in single experiments, they provide the means to develop multi-gene signatures from the unbiased screening of genome-wide expression data (Van Aggelen et al. 2010; Figure D.F).

The challenge of identifying specific molecular signatures hidden within hundreds of thousands of noisy variables has driven the development of statistical

methods for the identification of molecular components that are differentially expressed in two or more sample groups (i.e. stressed versus controls). Although effective, this approach has limitations: in particular it cannot identify synergistic effects between variables, it has a relatively low statistical power, and biological interpretation is challenging. The introduction of more complex modelling techniques that can assess the predictive power of combinations of biomarkers (Li et al. 2010), has been a significant step forward, particularly when applied to linking phenotypic responses (e.g. physiology) to molecular responses, especially in a network context. Ultimately this has allowed the identification of more effective and ecologically relevant biomarkers (Ankley et al. 2010).

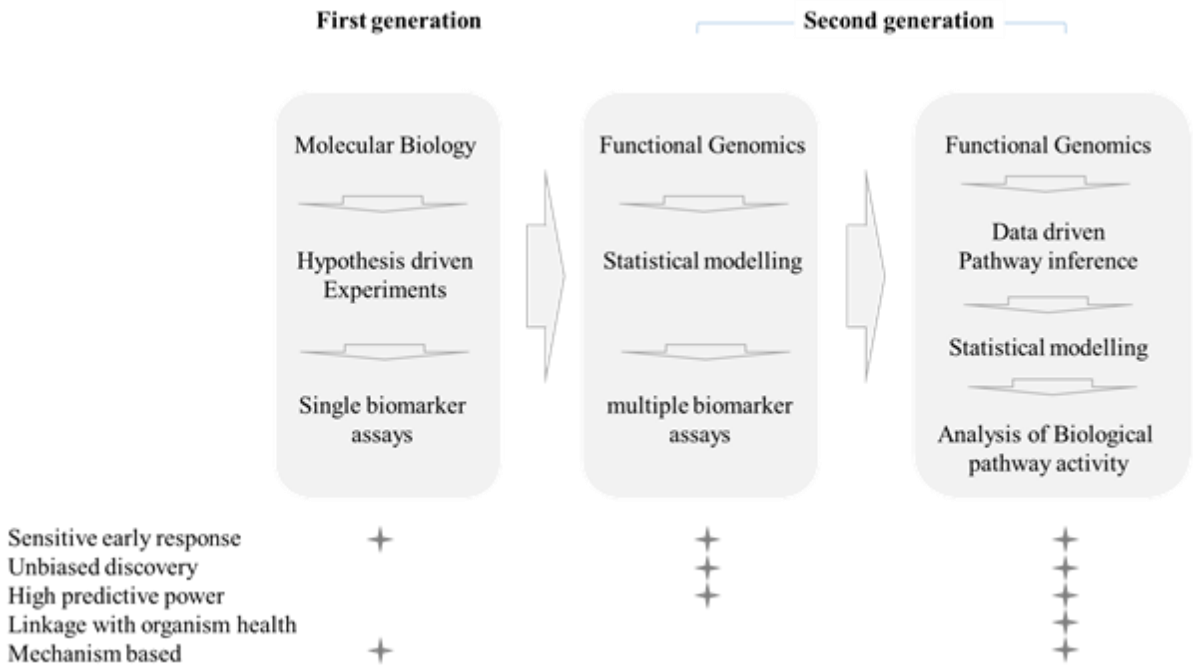



**Figure D.F.** The evolution of Biomarker discovery from the first generation approaches which use single genes whose expression is modulated by specific stressors, to the most recent advances which allow the discovery of multicomponent molecular signatures.

Despite the potential of these approaches, the vast number of possible combinations of individual measurements drastically limits their ability to explore a large portion of the solution space and therefore makes it extremely difficult to capture biologically relevant pathways that respond specifically to particular stressors. One way to address this challenge is reverse engineering, a branch of Systems Biology that aims to reconstruct the underlying structure of a biological pathway from experimental data. This has been tremendously effective in biomedical research for identifying pathways predictive of clinical response, drug resistance and novel therapeutic targets (Perkins et al. 2011). Again the biomedical-biomonitoring analogy can be used here to extend such approaches to environmental assessment. Because of the complexity of the datasets acquired using omics technologies, any reverse engineering approach must start from the identification of a high-level structure of the underlying biological networks and then progress to identifying more refined sub-networks, which are associated with important phenotypic responses, such as changes in reproductive ability following stress. Although in its infancy, this approach has already been applied by a number of groups for identifying novel stress pathways (Williams et al. 2011).

Overall, the use of these approaches allows the identification of more effective biomarkers than the ones based on differential expression or and has opened up the possibility to develop specific multi-component molecular signatures that are truly representative of a large number of stressors and with high specificity.

The use of biomarkers as a biomonitoring tool relies on inferences from molecular analyses. Returning to the more traditional approach of biomonitoring of using taxa themselves, and given that NGS technologies have finally enabled us to identify microbes in field conditions, these taxa represent ideal candidates for assessing how stressors alter community structure and ecosystem functioning. The pioneering "everything is everywhere, but the environment selects" theory proposed by Baas Becking (1934) suggests that the presence of all microorganisms is ubiquitous,

but our ability to detect them via direct observation is limited by varying densities: i.e., rare microbes may be present but unobserved in ecological samples (de Wit & Bouvier 2006). Consequently, the presence of different microbial species should be dictated by the difference in environmental conditions, rather than distance and biogeography (Zarraonaindia, Smith & Gilbert 2013). If this is true, it could provide a truly globally comparable framework for bioassessment and monitoring. Opposing theories exist, however, which suggest that microbial diversity is shaped by geography as well as the environment (Martiny et al. 2006; O'Malley 2008). The key question is whether the environment enhances the presence of certain microorganisms in different locations, allowing us to compare components of the microbial community for the monitoring of ecosystems. High-throughput technologies with increased detection capabilities can assist here and there is huge potential for these to be exploited by ecologists for monitoring purposes (Green, Bohannan & Whitaker 2008; Purdy et al. 2010; Poisot, Pequin & Gravel 2013; Woodward, Gray & Baird 2013).

Microorganisms play important functional roles in the major biogeochemical cycles at local to global scales, as well as the recycling of nutrients and overall ecosystem functioning (Cotner & Biddanda 2002; Nemergut et al. 2011), and many of these are also either ecosystem services in their own right or key processes that support important services (e.g. carbon sequestration). Moreover, microbial communities are themselves influenced by environmental conditions. Accordingly, bacteria have been suggested as good indicators of environmental change due to some of their attractive biomonitoring properties, such as high diversity (thus broad range of environmental susceptibility), potential ubiquity, short life cycles and minimal disturbance of the site during sampling (Lear et al. 2009, see Figure D.G).

However, until recently their use was hindered by the inability to study them in situ as only 5% of species are considered to be cultivable with standard techniques (Amann, Ludwig & Schleifer 1995; Curtis, Sloan & Scannell 2002) leading to narrowly focused approaches of single species analysis, such as the targeting of specific

ecotypes of pathogens, rather than whole-community detection (Hellawell 1986; Port et al. 2012). High throughput sequencing is already replacing historical fingerprinting approaches (Text box 2) and has been used for the characterization of whole communities from a large variety of sources, from both terrestrial and aquatic systems (Roesch et al. 2007; Cole, Konstandinidis & Farris 2010; Gilbert & Dupont 2011; Foote et al. 2012; Port et al. 2012). Following sequence-based approaches, specific and identifiable microorganisms can be linked to environmental status and used as sensors for the assessment of anthropogenic threats such as eutrophication, acidification, climate change, and land use changes (Port et al. 2012; Yergeau et al. 2012; Heino 2013). In aquatic ecosystems, whole bacterial cell analysis can also be used for the assessment of pollution effects (Lear et al. 2009) and detection of antibiotics in the water (Port et al. 2012).

Recent studies from terrestrial and marine systems (Pommier, Douzery & Mouillot 2012; Sun et al. 2012) suggest that bacterial communities are sensitive indicators of contaminant stress and also support the theory that presence of microorganisms is more related to environmental conditions than dispersal or geography. However, in a freshwater study Lear et al. (2012) found microbial communities did not differ among different environmental pressures, whereas invertebrate sampling was the more effective monitoring tool, suggesting that either the studied microbial communities were unaffected by contaminants, or the discriminatory power of the molecular fingerprinting approaches used was insufficient.

Yergeau et al. (2012) used NGS of the 16S rRNA gene to determine the effect of pollution related to oil sands mining on nearby aquatic microbial community structure. Their findings suggest that the microbial community structure was significantly altered by the distance from mining sites and support the potential use of Bacteria and Archaea as bioindicators of pollution. Furthermore, Kisand et al. (2012), were able to compare the microbial community composition of a highly

impacted area, like the port of Genoa with that of a protected area (low anthropogenic impact), through metagenomic analysis of the microbial communities from water samples. Distinct microbial diversity and abundance counts were detected among the different sites which can be related to the differences of environmental conditions, again demonstrating the potential for use of metagenomics for monitoring of aquatic ecosystems.

**Box 2: What is next generation sequencing/omics?**

The terms "next-generation" sequencing (NGS) or -omic technologies have been in use since a landmark paper (Margulies et al. 2005) detailed the use of 454 massively parallel pyrosequencing. Since then, the development of NGS platforms, accompanied by exponential increases in throughput and decreasing costs has completely transformed the field of DNA sequencing.

For investigating functional diversity, the NGS "-omic" approaches can conveniently be broken down into discrete categories of relevance to different levels of biological organisation. At the individual level, transcriptomic analyses measure differential gene expression via the analysis of expressed total RNA from specific tissues. At the community level, metagenetic or metabarcoding (Fonseca et al. 2010b; Bik et al. 2012; Taberlet et al. 2012) studies estimate environmental taxonomic richness by the en masse sequencing of environmental DNA samples (Sun et al. 2012). Shotgun metagenomic studies instead randomly sequence fragments of the total genomes present in an environmental DNA extraction(Knight et al. 2012), providing insights into both the functional and taxonomic capability of a given environment. Finally, metatranscriptomics enables researchers to investigate the actively transcribed mRNA from a community, giving an insight into the total gene expression from a local ecosystem (Filiatrault 2011; Gilbert & Hughes 2011).

As with microarray studies, gene expression is likely to change significantly at both short (Gilbert & Hughes 2011) and large spatial and temporal scales, so transcriptomic analyses need to be designed around carefully and explicitly framed questions that account for environmental gene expression and short half-life of mRNA (i.e., transcript analyses are often not associated with protein composition) (Moran et al. 2013). These broad -omic categories are summarized in Figure 7.

For ecological studies, a potential disadvantage of these approaches lies in the fact that most platforms incorporate various forms of clonal amplification in the sequencing approaches, thereby introducing potential quantitative biases into datasets. New "third-generation" sequencers and technologies (Ribeiro et al. 2012; Schneider & Dekker 2012; GridION™ and MinION™) that use single molecule sequencing approaches and therefore lack any clonal amplification step prior to sequencing could produce truly quantitative data, although these are currently tailored to analysing shorter numbers of very long reads and many had not reached market maturity at the time of writing.

*Continued…*

**Box 2: Continued**

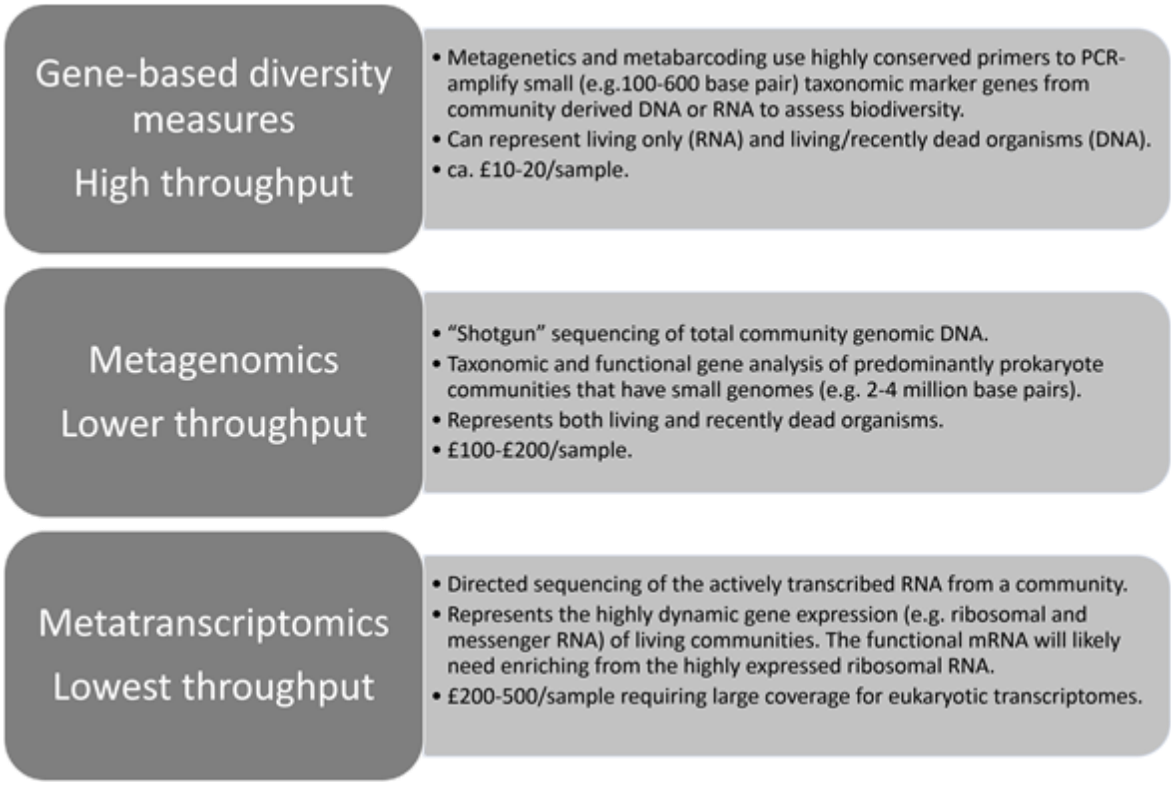



**Figure D.G**. The many -omics approaches to sequencing life, from individuals to whole community techniques that can be adapted to each scenario. Methods applicable to a variety of scales are presented with their respective advantages and disadvantages.

## D.6 The functional analysis of microbes, metazoans and macrofaunal communities

Ecologists are increasingly striving to improve predictive power by not only identifying what organisms are present, but also by asking what are they doing? The majority of functional ecological studies use organismal trait information (Tilman et al. 1997; Petchey & Gaston 2006; Hagen et al. 2012) to provide a metric for quantitative analysis, but these cannot accurately reflect all of the functional attributes of individuals and species in complex ecological communities. In theory, the -omic

toolbox can be employed to address this and to understand functional diversity in ways that have not been previously possible, although synergies with traditional ecology and taxonomy are essential if we are to fully understand the connections between biodiversity and ecosystem functioning and how they respond to stressors (Loreau et al. 2001).

If we consider a hypothetical freshwater ecosystem, with both benthic and aquatic habitats, these can be studied first independently but then combined by investigating both the taxonomic and functional diversity of the entire community using the -omic toolbox (see Text Box 2) tailored to organismal genome size and complimented by biogeochemical and nutrient cycling analyses. Starting with the microbial fraction, taxonomy marker genes such as 16S (Caporaso et al. 2011), ITS (Nilsson et al. 2008) and 18S (Fonseca et al. 2010a; Pawlowski et al. 2012) can be used for the high throughput assessment of bacteria, archaea, fungi and meiobiota respectively from multiple samples. Phylogenetic diversity can then be used throughout all gene marker schemes as a proxy for functional diversity, by employing algorithms such as UniFrac (Lozupone & Knight 2005; Caporaso et al. 2010; Fierer et al. 2012). Metagenomic and metatranscriptomic analysis can be employed to investigate the functional capability and specific functioning of the prokaryotic size fraction characterised by organisms with small genomes (e.g. 2-4Mb) and their relatively small transcriptomes. Metatranscriptomic analyses are likely to be robust in simple communities of eukaryotic organisms where just a few species dominate (Durkin et al. 2012), but given the current limits of sequencing power, achieving effective coverage of replicated samples of complex eukaryotic communities (Bailly et al. 2007), whose transcriptomes can be very large (e.g. 20Mb), is still limited. Similarly, metagenomic sequencing of eukaryotic communities is unlikely to reach the appropriate depth of coverage for ecological synthesis, simply because eukaryotic genomes can be very large (the human genome alone is over 3Gb in size).

Within prokaryotic communities, a new approach (PiCrust) (Langille et al.

2013) has emerged that links marker gene 16S studies to functional diversity maps environmental 16S reads to their closest ancestors with full genome sequences and predicts ancestral states of functional gene ontologies. Initial analyses suggest this outperforms low coverage shotgun metagenomic analyses in well-characterised communities, but further testing and examples will undoubtedly provide further insight. Nevertheless, the model provides a route between high throughput studies and full genome capability that may also eventually feature in the eukaryotic biosphere as more genomes are sequenced.

Advances that are likely to be provided by the -omic toolbox regarding the functional diversity of eukaryotic communities (e.g. protists, fungi, meiobiota and macrofauna) are likely to be achieved by linking genotype phenotype data with the analysis of food webs and networks (Barberan et al. 2012; Rodriguez-Lanetty et al. 2013). The Barcode of Life Project (Ratnasingham & Hebert 2007) strives for the provision of standardised and carefully curated DNA barcode data for organisms based on official barcode markers. So far, almost 200,000 species have been "barcoded". Importantly, this endeavour provides a link between a standardised genotype and the taxonomy and ecology of the barcoded species. At the start of the barcoding movement, sequencing technologies were not mature enough to consider assessing multiple communities of organisms, but recently, a multitude of "meta-barcoding" studies (Epp et al. 2012; Taberlet et al. 2012) have shown that approaches used for microbial communities can be conveniently transferred to macrofaunal communities. If the featured species in the meta-barcoding datasets have barcode reference data, this can provide a very powerful link to the functional attributes of the organisms comprising the sequenced communities. The maturation of the field of meta-barcoding not only provides a huge boost for our ability to assess large numbers of macrofaunal samples simultaneously (Ji et al. 2013), but also re-asserts the need for the generation of reference barcode libraries to provide the necessary links between -omic technologies and functional ecology. Moreover, since gene marker-based studies

do not respect the boundaries between living and recently deceased, or even ingested species, dietary and food web analyses can be conveniently performed using either individual, or species-based sequencing of gut contents to investigate trophic interactions (Pompanon et al. 2012).

Overlying these possibilities is the further opportunity to deduct functional relationships using the analysis of ecological networks at multispecies levels of organisation (Ings et al. 2009; Hagen et al. 2012). Following marker-based approaches and even metagenomic analyses, the resulting data is a familiar taxon-by-sample frequency matrix of genotype occurrence (Ji et al. 2013), that can be related back to phenotype occurrence (i.e. species). The quantitative nature of the associations can be estimated on the basis of the mode of evolution and genomic content of the markers used (while acknowledging potential PCR bias), but the co-occurrence incidence matrices will reflect the distribution of species in space and time. Such power potentially enables us to delimit co-occurring ecological networks (in space and/or time) and how individual networks respond to external drivers. Moreover, some components of the sequence data matrices will be annotated to a high degree of accuracy (e.g. species level for barcoded metabarcoding data) and for all other groups, potentially genus, order, family etc., but at least phylum, enabling the researcher to characterise biological interactions (parasitisim, predation, commensalism, mutualism, competition etc.) and ecological processes (Faust & Raes 2012). The additional strength of -omic high-throughput marker based approaches is that with the now routine analysis of ca. 50 complex samples simultaneously, a high degree of replication and sample coverage can be achieved on scales that are simply not possible using traditional approaches for either microbial or macrofaunal samples. The combination of these emerging technologies and approaches promises a possible means of truly integrating ecological and evolutionary perspectives to responses to stressors across all the major domains of life in aquatic (and terrestrial) ecosystems.

## D.7 Concluding remarks

With an ever-increasing human population the need to monitor and predict our effects on the natural world has never been more important. In the developed world the predominant stressors have changed, presenting new challenges to biomonitoring science (Figure D.B), while developing nations such as India or China are facing the same stressors the western world was exposed to in the 20th century but on a far greater scale (Abate 1995; Yagishita 1995; Aggarwal et al. 2001). An eco-evolutionary approach to biomonitoring will allow us to better understand the dynamics between the selective forces of evolution and the ecology of species. The ability of a community to adapt to change is key to its response to a particular stressor (Woodward et al. 2010; Moya-Larano et al. 2012), and this needs to be considered alongside biomonitoring results. With new technologies such as the rise of new molecular markers (e.g. Van Aggelen et al. 2010; Williams et al. 2011), to the use of microbes (e.g. Lear et al. 2009) and the advances in NGS techniques (Text Box 2) there is a great variety in approaches now available to monitor the functional response of aquatic communities to environmental stress.

A shift in the culture surrounding legislative biomonitoring, governance and stakeholder implementation will be required before these advanced and promising approaches can be integrated into current protocols. There will likely be far fewer "traditional" taxonomists as NGS technologies take over, but many more bioinformaticians will be needed to process and analyse the NGS samples. The rate-limiting step in biomonitoring will shift from the slow and laborious process of identifying individuals through microscopy (data acquisition) to limitations in the efficiency with which large volumes of data can be processed. It is not impossible to imagine a future where remote sensing stations which monitors environmental DNA or RNA and send sequence data back to the laboratory via telemetry, as weather stations do now; unmanned and automated transmitting results back to a central point. As bioinformatics solutions to data analysis and synthesis continue to develop

over time and its huge potential to the biomonitoring world, it is likely to be simply a matter of 'when' and not 'if' this revolution will take place on a truly global scale.

## D.8 References

Abate, T. (1995) Swedish Scientists Take Acid-Rain Research to Developing-Nations. *Bioscience,* **45,** 738-740.

Abrahams, C., Brown, L., Dale, K., Edwards, F., Jeffries, M., Klaar, M., Ledger, M.E., May, L., Milner, A., Murphy, J., Robertson, A. & Woodward, G. (2013) The Impact of Extreme Events on Freshwater Ecosystems. (ed. I. Jones). London.

Aggarwal, S.G., Chandrawanshi, C.K., Patel, R.M., Agarwal, S., Kamavisdar, A. & Mundhara, G.L. (2001) Acidification of surface water in central India. *Water Air and Soil Pollution*, **130,** 855-862.

Allen, A.P., Brown, J.H. & Gillooly, J.F. (2002) Global biodiversity, biochemical kinetics, and the energetic-equivalence rule. *Science*, **297,** 1545-1548.

Amann, R.I., Ludwig, W. & Schleifer, K.H. (1995) Phylogenetic Identification and in-Situ Detection of Individual Microbial-Cells without Cultivation. *Microbiological Reviews*, **59,** 143-169.

Amiard, J.-C., Amiard-Triquet, C., Barka, S., Pellerin, J. & Rainbow, P. (2006) Metallothioneins in aquatic invertebrates: their role in metal detoxification and their use as biomarkers. *Aquatic toxicology,* **76,** 160-202.

Ankley, G.T., Bennett, R.S., Erickson, R.J., Hoff, D.J., Hornung, M.W., Johnson, R.D., Mount, D.R., Nichols, J.W., Russom, C.L., Schmieder, P.K., Serrrano, J.A., Tietge, J.E. & Villeneuve, D.L. (2010) Adverse Outcome Pathways: A Conceptual Framework to Support Ecotoxicology Research and Risk Assessment. *Environmental Toxicology and Chemistry,* **29,** 730-741.

Baas Becking, L.G.M. (1934) *Geobiologie of inleiding tot de milieukunde*. WP Van Stockum & Zoon.

Bailly, J., Fraissinet-Tachet, L., Verner, M.C., Debaud, J.C., Lemaire, M., Wesolowski-Louvel, M. & Marmeisse, R. (2007) Soil eukaryotic functional diversity, a metatranscriptomic approach. *Isme Journal*, **1,** 632-642.

Barberan, A., Bates, S.T., Casamayor, E.O. & Fierer, N. (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. *Isme Journal*, **6,** 343-351.

Battarbee, R., Shilland, E., Simpson, G., Salgado, J., Goldsmith, B., Gray, W. & Turner, S. (2013) Surface-water Acidity in the River Dart Area. Report for the West Country Rivers Trust.

Battarbee, R.W. & Charles, D.F. (1986) Diatom-based pH reconstruction studies of acid lakes in Europe and North America: A synthesis. *Water, Air, and Soil Pollution*, **30,** 347-354.

Battarbee, R.W., Monteith, D.T., Juggins, S., Evans, C.D., Jenkins, A. & Simpson, G.L. (2005) Reconstructing pre-acidification pH for an acidified Scottish loch: A comparison of palaeolimnological and modelling approaches. *Environmental Pollution*, **137,** 135-149.

Battarbee, R.W., Simpson, G.L., Shilland, E.M., Flower, R.J., Kreiser, A., Yang, H., Clarke, amp & Gina (2014) Recovery of UK lakes from acidification: An assessment using combined palaeoecological and contemporary diatom assemblage data. *Ecological Indicators,* **37, Part B,** 365-380.

Bell, G. (2001) Neutral macroecology. *Science,* **293,** 2413-2418.

Bell, G. & Gonzalez, A. (2011) Adaptation and Evolutionary Rescue in Metapopulations Experiencing Environmental Deterioration. *Science,* **332,** 1327-1330.

Bik, H.M., Porazinska, D.L., Creer, S., Caporaso, J.G., Knight, R. & Thomas, W.K. (2012) Sequencing our way towards understanding global eukaryotic biodiversity. *Trends in Ecology & Evolution,* **27,** 233-243.

Birks, H.J.B., Line, J.M., Juggins, S., Stevenson, A.C. & Terbraak, C.J.F. (1990) Diatoms and Ph Reconstruction. *Philosophical Transactions of the Royal Society B-Biological Sciences,* **327,** 263-278.

Bonada, N., Dolédec, S. & Statzner, B. (2012) Spatial autocorrelation patterns of stream invertebrates: exogenous and endogenous factors. *Journal of Biogeography,* **39,** 56-68.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K. & Gordon, J.I. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature methods,* **7,** 335-336.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N. & Knight, R. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences,* **108,** 4516-4522.

Cardinale, B.J. (2011) Biodiversity improves water quality through niche partitioning. *Nature,* **472,** 86-89.

Cardinale, B.J., Srivastava, D.S., Duffy, J.E., Wright, J.P., Downing, A.L., Sankaran, M. & Jouseau, C. (2006) Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature,* **443,** 989-992.

Carpenter, S.R., Cole, J.J., Hodgson, J.R., Kitchell, J.F., Pace, M.L., Bade, D., Cottingham, K.L., Essington, T.E., Houser, J.N. & Schindler, D.E. (2001) Trophic cascades, nutrients, and lake productivity: whole-lake experiments. *Ecological Monographs,* **71,** 163-186.

Celander, M.C. (2011) Cocktail effects on biomarker responses in fish. *Aquatic toxicology,* **105,** 72-77.

Cole, J.R., Konstandinidis, K. & Farris, R.J. (2010) Microbial Diversity and Phylogeny: extending from rRNAs to genomes. *Environmental Molecular Microbiology,* pp. 1-21. Caister Academic Press, Norfolk, UK.

Costanza, R., d'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'neill, R.V. & Paruelo, J. (1997) The value of the world's ecosystem services

and natural capital. *Nature,* **387,** 253-260.

Cotner, J.B. & Biddanda, B.A. (2002) Small players, large role: microbial influence on biogeochemical processes in pelagic aquatic ecosystems. *Ecosystems,* **5,** 105-121.

Curtis, T.P., Sloan, W.T. & Scannell, J.W. (2002) Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences of the United States of America,* **99,** 10494-10499.

Dangles, O., Gessner, M.O., Guerold, F. & Chauvet, E. (2004) Impacts of stream acidification on litter breakdown: implications for assessing ecosystem functioning. *Journal of Applied Ecology,* **41,** 365-378.

Darwin, C. (1872) *On the origin of species by means of natural selection*. John Murray, London.

de Wit, R. & Bouvier, T. (2006) 'Everything is everywhere, but, the environment selects'; what did Baas Becking and Beijerinck really say? *Environmental Microbiology,* **8,** 755-758.

Demars, B.O. & Trémolières, M. (2009) Aquatic macrophytes as bioindicators of carbon dioxide in groundwater fed rivers. *Science of the Total Environment,* **407,** 4752-4763.

Demars, B.O.L. & Edwards, A.C. (2009) Distribution of aquatic macrophytes in contrasting river systems: A critique of compositional-based assessment of water quality. *Science of the Total Environment,* **407,** 975-990.

Demars, B.O.L. & Harper, D.M. (2005) Distribution of aquatic vascular plants in lowland rivers: separating the effects of local environmental conditions, longitudinal connectivity and river basin isolation. *Freshwater Biology,* **50,** 418-437.

Demars, B.O.L., Kemp, J.L., Friberg, N., Usseglio-Polatera, P. & Harper, D.M. (2012) Linking biotopes to invertebrates in rivers: Biological traits, taxonomic composition and diversity. *Ecological Indicators,* **23,** 301-311.

Demars, B.O.L., Manson, J.R., Olafsson, J.S., Gislason, G.M., Gudmundsdottir, R., Woodward, G., Reiss, J., Pichler, D.E., Rasmussen, J.J. & Friberg, N. (2011) Temperature and the metabolic balance of streams. *Freshwater Biology,* **56,** 1106-1121.

Durkin, C.A., Marchetti, A., Bender, S.J., Truong, T., Morales, R., Mock, T. & Armbrust, E.V. (2012) Frustule-related gene transcription and the influence of diatom community composition on silica precipitation in an iron-limited environment. *Limnology and Oceanography,* **57,** 1619-1633.

Engelhardt, K.A. (2006) Relating effect and response traits in submersed aquatic macrophytes. *Ecological Applications,* **16,** 1808-1820.

Enquist, B.J., Kerkhoff, A.J., Stark, S.C., Swenson, N.G., McCarthy, M.C. & Price, C.A. (2007) A general integrative model for scaling plant growth, carbon flux, and functional trait spectra. *Nature,* **449,** 218-222.

Epp, L.S., Boessenkool, S., Bellemain, E.P., Haile, J., Esposito, A., Riaz, T., Erseus, C.,

Gusarov, V.I., Edwards, M.E., Johnsen, A., Stenoien, H.K., Hassel, K., Kauserud, H., Yoccoz, N.G., Brathen, K., Willerslev, E., Taberlet, P., Coissac, E. & Brochmann, C. (2012) New environmental metabarcodes for analysing soil DNA: potential for studying past and present ecosystems. *Molecular Ecology,* **21,** 1821-1833.

Everard, M. & McInnes, R. (2013) Systemic solutions for multi-benefit water and environmental management. *Science of the Total Environment,* **461,** 170-179.

Faust, K. & Raes, J. (2012) Microbial interactions: from networks to models. *Nature Reviews Microbiology,* **10,** 538-550.

Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., Owens, S., Gilbert, J.A., Wall, D.H. & Caporaso, J.G. (2012) Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences of the United States of America,* **109,** 21390-21395.

Filiatrault, M.J. (2011) Progress in prokaryotic transcriptomics. *Current Opinion in Microbiology,* **14,** 579-586.

Fonseca, V.G., Carvalho, G.R., Sung, W., Johnson, H.F., Power, D.M., Neill, S.P., Packer, M., Blaxter, M.L., Lambshead, P.J.D. & Thomas, W.K. (2010a) Second-generation environmental sequencing unmasks marine metazoan biodiversity. *Nature communications,* **1,** 98.

Fonseca, V.G., Carvalho, G.R., Sung, W., Johnson, H.F., Power, D.M., Neill, S.P., Packer, M., Blaxter, M.L., Lambshead, P.J.D., Thomas, W.K. & Creer, S. (2010b) Second-generation environmental sequencing unmasks marine metazoan biodiversity. *Nature communications,* **1**.

Foote, A.D., Thomsen, P.F., Sveegaard, S., Wahlberg, M., Kielgast, J., Kyhn, L.A., Salling, A.B., Galatius, A., Orlando, L. & Gilbert, M.T.P. (2012) Investigating the Potential Use of Environmental DNA (eDNA) for Genetic Monitoring of Marine Mammals. *Plos One,* **7**.

Friberg, N., Bonada, N., Bradley, D.C., Dunbar, M.J., Edwards, F.K., Grey, J., Hayes, R.B., Hildrew, A.G., Lamouroux, N., Trimmer, M. & Woodward, G. (2011) Biomonitoring of Human Impacts in Freshwater Ecosystems: The Good, the Bad and the Ugly. *Advances in Ecological Research,* **44,** 1-68.

Froese, R. & Pauly, D. (2010) FishBase. International Center for Living Aquatic Resources Management.

García-Llorente, M., Martín-López, B., Díaz, S. & Montes, C. (2011) Can ecosystem properties be fully translated into service values? An economic valuation of aquatic plant services. *Ecological Applications,* **21,** 3083-3103.

Gilbert, J.A. & Dupont, C.L. (2011) Microbial Metagenomics: Beyond the Genome. *Annual Review of Marine Science, Vol 3,* **3,** 347-371.

Gilbert, J.A. & Hughes, M. (2011) Gene expression profiling: metatranscriptomics. *High-Throughput Next Generation Sequencing*, pp. 195-205. Springer.

Gould, S.J. & Lewontin, R.C. (1979) Spandrels of San-Marco and the Panglossian Paradigm - a Critique of the Adaptationist Program. *Proceedings of the Royal Society Series B-Biological Sciences,* **205,** 581-598.

Green, J.L., Bohannan, B.J.M. & Whitaker, R.J. (2008) Microbial biogeography: From taxonomy to traits. *Science,* **320,** 1039-1043.

Greulich, S. & Bornette, G. (2003) Being evergreen in an aquatic habitat with attenuated seasonal contrasts - a major competitive advantage? *Plant Ecology,* **167,** 9-18.

Hagen, M., Kissling, W.D., Rasmussen, C., De Aguiar, M.A.M., Brown, L.E., Carstensen, D.W., Alves-Dos-Santos, I., Dupont, Y.L., Edwards, F.K., Genini, J., Guimaraes, P.R., Jenkins, G.B., Jordano, P., Kaiser-Bunbury, C.N., Ledger, M.E., Maia, K.P., Marquitti, F.M.D., Mclaughlin, O., Morellato, L.P.C., O'Gorman, E.J., Trojelsgaard, K., Tylianakis, J.M., Vidal, M.M., Woodward, G. & Olesen, J.M. (2012) Biodiversity, Species Interactions and Ecological Networks in a Fragmented World. *Advances in Ecological Research, Vol 46, Pt 1,* **46,** 89-210.

Hajibabaei, M. (2012) The golden age of DNA metasystematics. *Trends in genetics : TIG,* **28,** 535-537.

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G.A.C. & Baird, D.J. (2011) Environmental Barcoding: A Next-Generation Sequencing Approach for Biomonitoring Applications Using River Benthos. *Plos One,* **6**.

Halvorsen, G.A., Heegaard, E., Fjellheim, A. & Raddum, G.G. (2003) Tracing recovery from acidification in the western Norwegian Nausta watershed. *Ambio,* **32,** 235-239.

Harte, J. (2011) *Maximum entropy and ecology: a theory of abundance, distribution, and energetics*. Oxford University Press.

Heino, J. (2013) The importance of metacommunity ecology for environmental assessment research in the freshwater realm. *Biological Reviews,* **88,** 166-178.

Hellawell, J.M. (1986) Biological indicators of freshwater pollution and environmental management.

Hildrew, A.G. (2009) Sustained Research on Stream Communities: A Model System and The Comparative Approach. *Advances in Ecological Research,* **41,** 175-312.

Hill, B.H., Elonen, C.M., Seifert, L.R., May, A.A. & Tarquinio, E. (2012) Microbial enzyme stoichiometry and nutrient limitation in US streams and rivers. *Ecological Indicators,* **18,** 540-551.

Horrigan, N. & Baird, D.J. (2008) Trait patterns of aquatic insects across gradients of flow-related factors: a multivariate analysis of Canadian national data. *Canadian Journal of Fisheries and Aquatic Sciences,* **65,** 670-680.

Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press.

Humboldt, A. (1849) Cosmos: A Sketch of a Physical Description of the Universe. *EC Otté, I, London: Henry G. Bohn*.

Ings, T.C., Montoya, J.M., Bascompte, J., Bluthgen, N., Brown, L., Dormann, C.F., Edwards, F., Figueroa, D., Jacob, U., Jones, J.I., Lauridsen, R.B., Ledger, M.E., Lewis, H.M., Olesen, J.M., van Veen, F.J.F., Warren, P.H. & Woodward, G. (2009) Ecological networks - beyond food webs. *Journal of Animal Ecology,* **78,** 253-269.

Ji, Y., Ashton, L., Pedley, S.M., Edwards, D.P., Tang, Y., Nakamura, A., Kitching, R., Dolman, P.M., Woodcock, P. & Edwards, F.A. (2013) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters,* **16,** 1245-1257.

Johnson, L.J. & Tricker, P.J. (2010) Epigenomic plasticity within populations: its evolutionary significance and potential. *Heredity,* **105,** 113-121.

Jones, J.I. & Sayer, C.D. (2003) Does the fish-invertebrate-periphyton cascade precipitate plant loss in shallow lakes? *Ecology,* **84,** 2155-2167.

Kampfraath, A.A., Hunting, E.R., Mulder, C., Breure, A.M., Gessner, M.O., Kraak, M.H.S. & Admiraal, W. (2012) DECOTAB: a multipurpose standard substrate to assess effects of litter quality on microbial decomposition and invertebrate consumption. Freshwater Science, 31, 1156-1162.

Kerkhoff, A.J. & Enquist, B.J. (2006) Ecosystem allometry: the scaling of nutrient stocks and primary productivity across plant communities. *Ecology Letters,* **9,** 419-427.

Kisand, V., Valente, A., Lahm, A., Tanet, G. & Lettieri, T. (2012) Phylogenetic and Functional Metagenomic Profiling for Assessing Microbial Biodiversity in Environmental Monitoring. *Plos One,* **7**.

Knight, R., Jansson, J., Field, D., Fierer, N., Desai, N., Fuhrman, J.A., Hugenholtz, P., van der Lelie, D., Meyer, F., Stevens, R., Bailey, M.J., Gordon, J.I., Kowalchuk, G.A. & Gilbert, J.A. (2012) Unlocking the potential of metagenomics through replicated experimental design. *Nature Biotechnology,* **30,** 513-520.

Kolkwitz, R. & Marsson, M. (1902) Grundsätze für die biologische Beurteilung des Wassers nach seiner Flora und Fauna. *Mitteilungen der königlichen Prüfanstalt für Wasserversorgung und Abwasserbeseitigung,* **1,** 3-72.

Kolkwitz, R. & Marsson, M. (1909) Ökologie der tierischen Saprobien. Beiträge zur Lehre von der biologischen Gewässerbeurteilung. *Internationale Revue der gesamten Hydrobiologie und Hydrographie,* **2,** 126-152.

Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile, D.E., Thurber, R.L.V., Knight, R., Beiko, R.G. & Huttenhower, C. (2013) Predictive functional profiling of microbial communities using

16S rRNA marker gene sequences. *Nature Biotechnology*, **31,** 814-+.

Layer, K., Hildrew, A., Monteith, D. & Woodward, G. (2010) Long-term variation in the littoral food web of an acidified mountain lake. *Global Change Biology,* **16,** 3133-3143.

Layer, K., Hildrew, A.G., Jenkins, G.B., Riede, J.O., Rossiter, S.J., Townsend, C.R. & Woodward, G. (2011) Long-Term Dynamics of a Well-Characterised Food Web: Four Decades of Acidification and Recovery in the Broadstone Stream Model System. *Advances in Ecological Research,* **44,** 69-117.

Layer, K., Hildrew, A.G. & Woodward, G. (2013) Grazing and detritivory in 20 stream food webs across a broad pH gradient. *Oecologia,* **171,** 459-471.

Lear, G., Ancion, P.Y., Harding, J. & Lewis, G.D. (2012) Use of bacterial communities to assess the ecological health of a recently restored stream. *New Zealand Journal of Marine and Freshwater Research,* **46,** 291-301.

Lear, G., Boothroyd, I.K.G., Turner, S.J., Roberts, K. & Lewis, G.D. (2009) A comparison of bacteria and benthic invertebrates as indicators of ecological health in streams. *Freshwater Biology,* **54,** 1532-1543.

Ledger, M.E., Brown, L.E., Edwards, F., Milner, A.M. & Woodward, G. (2012) Drought alters the structure and functioning of complex food webs. *Nature Climate Change*.

Ledger, M.E. & Hildrew, A.G. (2005) The ecology of acidification and recovery: changes in herbivore-algal food web linkages across a stream pH gradient. *Environmental Pollution,* **137,** 103-118.

Ledger, M.E., Milner, A., Brown, L., Edwards, F., Hudson, L. & Woodward, G. (2013) Extreme Climatic Events Alter Aquatic Food Webs. A Synthesis of Evidence from a Mesocosm Drought Experiment. *Advances in Ecological Research,* **48,** 343-395.

Lennon, J.J., Koleff, P., GreenwooD, J.J.D. & Gaston, K.J. (2001) The geographical structure of British bird distributions: diversity, spatial turnover and scale. *Journal of Animal Ecology,* **70,** 966-979.

Lennon, J.J., Kunin, W.E., Hartley, S. & Gaston, K.J. (2007) Species distribution patterns, diversity scaling and testing for fractals in southern African birds. *Scaling biodiversity* (eds D. Storch, P.A. Marquet & J.H. Brown). Cambridge University Press Cambridge, UK.

Li, Y., Wang, N., Perkins, E.J., Zhang, C. & Gong, P. (2010) Identification and optimization of classifier genes from multi-class earthworm microarray dataset. *Plos One,* **5,** e13715.

Likens, G.E., Bormann, F.H., Pierce, R.S., Eaton, J.S. & Johnson, N.M. (1977) *Biogeo-chemistry of a forested ecosystem*. Springer-Verlag, New York.

López-Urrutia, Á., San Martin, E., Harris, R.P. & Irigoien, X. (2006) Scaling the metabolic balance of the oceans. *Proceedings of the National Academy of Sciences,* **103,** 8739-8744.

Loreau, M. (2010) *From populations to ecosystems: Theoretical foundations for a new ecological synthesis (MPB-46)*. Princeton University Press.

Loreau, M., Naeem, S. & Inchausti, P. (2002) *Biodiversity and ecosystem functioning: synthesis and perspectives*. Oxford University Press.

Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J.P., Hector, A., Hooper, D.U., Huston, M.A., Raffaelli, D., Schmid, B., Tilman, D. & Wardle, D.A. (2001) Ecology - Biodiversity and ecosystem functioning: Current knowledge and future challenges. *Science,* **294,** 804-808.

Lozupone, C. & Knight, R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Applied and Environmental Microbiology,* **71,** 8228-8235.

Lyell, C. & Deshayes, G.P. (1830) *Principles of Geology: being an attempt to explain the former changes of the earth's surface, by reference to causes now in operation*. John Murray.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J. & Chen, Z. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature,* **437,** 376-380.

Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Krumins, J.A., Kuske, C.R., Morin, P.J., Naeem, S., Ovreas, L., Reysenbach, A.L., Smith, V.H. & Staley, J.T. (2006) Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology,* **4,** 102-112.

May, R.M. (1990) Taxonomy as Destiny. *Nature,* **347,** 129-130.

Mayden, R.L. (1997) A hierarchy of species concepts: The denouement in the saga of the species problem. *Systematics Association Special Volume Series; Species: The units of biodiversity* (eds M.F. Claridge, H.A. Dawah & M.R. Wilson), pp. 381-424.

McKie, B.G., Woodward, G., Hladyz, S., Nistorescu, M., Preda, E., Popescu, C., Giller, P.S. & Malmqvist, B. (2008) Ecosystem functioning in stream assemblages from different regions: contrasting responses to variation in detritivore richness, evenness and density. *Journal of Animal Ecology*, **77,** 495-504.

Melián, C., Vilas, C., Baldo, F., González-Ortegón, E., Drake, P. & Williams, R.J. (2011) Eco-evolutionary dynamics of individual-based food webs. *Advances in Ecological Research,* **45,** 225-268.

Metcalfe, J.L. (1989) Biological Water-Quality Assessment of Running Waters Based on Macroinvertebrate Communities - History and Present Status in Europe. *Environmental Pollution,* **60,** 101-139.

Millenium Ecosystem Assessment, (2005) Ecosystems and Human Well-being: Current State and Trends, Volume 1. Washington.

Moilanen, A., Franco, A.M., Early, R.I., Fox, R., Wintle, B. & Thomas, C.D. (2005) Prioritizing multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society B: Biological Sciences,* **272,** 1885-1891.

Moilanen, A. & Hanski, I. (2001) On the use of connectivity measures in spatial ecology. *Oikos,* **95,** 147-151.

Moilanen, A., Leathwick, J. & Elith, J. (2008) A method for spatial freshwater conservation prioritization. *Freshwater Biology,* **53,** 577-592.

Monteith, D.T., Hildrew, A.G., Flower, R.J., Raven, P.J., Beaumont, W.R.B., Collen, P., Kreiser, A.M., Shilland, E.M. & Winterbottom, J.H. (2005) Biological responses to the chemical recovery of acidified fresh waters in the UK. *Environmental Pollution,* **137,** 83-101.

Moran, M.A., Satinsky, B., Gifford, S.M., Luo, H.W., Rivers, A., Chan, L.K., Meng, J., Durham, B.P., Shen, C., Varaljay, V.A., Smith, C.B., Yager, P.L. & Hopkinson, B.M. (2013) Sizing up metatranscriptomics. *Isme Journal,* **7,** 237-243.

Moya-Larano, J., Verdeny-Vilalta, O., Rowntree, J., Melguizo-Ruiz, N., Montserrat, M. & Laiolo, P. (2012) Climate Change and Eco-Evolutionary Dynamics in Food Webs. *Advances in Ecological Research,* **47,** 1-80.

Mulholland, P.J., Helton, A.M., Poole, G.C., Hall, R.O., Hamilton, S.K., Peterson, B.J., Tank, J.L., Ashkenas, L.R., Cooper, L.W. & Dahm, C.N. (2008) Stream denitrification across biomes and its response to anthropogenic nitrate loading. *Nature,* **452,** 202-205.

Murphy, J.F., Davy-Bowker, J., McFarland, B. & Ormerod, S.J. (2013) A diagnostic biotic index for assessing acidity in sensitive streams in Britain. *Ecological Indicators,* **24,** 562-572.

Murphy, J.F., Winterbottom, J.H., Orton, S., Simpson, G.L., Shilland, E.M. & Hildrew, A.G. (2012) Evidence of recovery from acidification in the macroinvertebrate assemblages of UK fresh waters: A 20-year time series. *Ecological Indicators*.

Naeem, S. (2002) Disentangling the Impacts of Diversity on Ecosystem Functioning in Combinatorial Experiments. *Ecology,* **83,** 2925-2935.

Nemergut, D.R., Costello, E.K., Hamady, M., Lozupone, C., Jiang, L., Schmidt, S.K., Fierer, N., Townsend, A.R., Cleveland, C.C., Stanish, L. & Knight, R. (2011) Global patterns in the biogeography of bacterial taxa. *Environmental Microbiology,* **13,** 135-144.

Nilsson, R.H., Kristiansson, E., Ryberg, M., Hallenberg, N. & Larsson, K.-H. (2008) Intraspecific ITS variability in the kingdom Fungi as expressed in the international sequence databases and its implications for molecular species identification. *Evolutionary bioinformatics online,* **4,** 193.

O'Malley, M.A. (2008) Everything is everywhere but the environment selects: ubiquitous

distribution and ecological determinism in microbial biogeography. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences,* **39,** 314-325.

Odum, E.P. (1969) The strategy of ecosystem development. An understanding of ecological succession provides a basis for resolving man's conflict with nature. *Science,* **164,** 262-269.

Paradis, E., Claude, J. & Strimmer, K. (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics,* **20,** 289-290.

Pawlowski, J., Audic, S., Adl, S., Bass, D., Belbahri, L., Berney, C., Bowser, S.S., Cepicka, I., Decelle, J. & Dunthorn, M. (2012) CBOL protist working group: barcoding eukaryotic richness beyond the animal, plant, and fungal kingdoms. *PLoS biology,* **10,** e1001419.

Perkins, D.M., McKie, B.G., Malmqvist, B., Gilmour, S.G., Reiss, J. & Woodward, G. (2010) Environmental Warming and Biodiversity-Ecosystem Functioning in Freshwater Microcosms: Partitioning the Effects of Species Identity, Richness and Metabolism. *Advances in Ecological Research,* **43,** 177-209.

Perkins, E.J., Chipman, J.K., Edwards, S., Habib, T., Falciani, F., Taylor, R., Van Aggelen, G., Vulpe, C., Antczak, P. & Loguinov, A. (2011) Reverse Engineering Adverse Outcome Pathways. *Environmental Toxicology and Chemistry,* **30,** 22-38.

Petchey, O.L., Downing, A.L., Mittelbach, G.G., Persson, L., Steiner, C.F., Warren, P.H. & Woodward, G. (2004) Species loss and the structure and functioning of multitrophic aquatic systems. *Oikos,* **104,** 467-478.

Petchey, O.L. & Gaston, K.J. (2006) Functional diversity: back to basics and looking forward. *Ecology Letters,* **9,** 741-758.

Poff, N.L., Olden, J.D., Vieira, N.K.M., Finn, D.S., Simmons, M.P. & Kondratieff, B.C. (2006) Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogenetic relationships. *Journal of the North American Benthological Society,* **25,** 730-755.

Poisot, T., Pequin, B. & Gravel, D. (2013) High-Throughput Sequencing: A Roadmap Toward Community Ecology. *Ecology and Evolution,* **3,** 1125-1139.

Pollux, B.J.A., Santamaria, L. & Ouborg, N.J. (2005) Differences in endozoochorous dispersal between aquatic plant species, with reference to plant population persistence in rivers. *Freshwater Biology,* **50,** 232-242.

Pommier, T., Douzery, E.J.P. & Mouillot, D. (2012) Environment drives high phylogenetic turnover among oceanic bacterial communities. *Biology Letters,* **8,** 562-566.

Pompanon, F., Deagle, B.E., Symondson, W.O., Brown, D.S., Jarman, S.N. & Taberlet, P. (2012) Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology,* **21,** 1931-1950.

Port, J.A., Wallace, J.C., Griffith, W.C. & Faustman, E.M. (2012) Metagenomic Profiling of Microbial Composition and Antibiotic Resistance Determinants in Puget Sound. *Plos One,* **7**.

Purdy, K.J., Hurd, P.J., Moya-Larano, J., Trimmer, M., Oakley, B.B. & Woodward, G. (2010) Systems Biology for Ecology: From Molecules to Ecosystems. *Advances in Ecological Research,* **43,** 87-149.

Ratnasingham, S. & Hebert, P.D.N. (2007) BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes,* **7,** 355-364.

Rawcliffe, R., Sayer, C.D., Woodward, G., Grey, J., Davidson, T.A. & Jones, J.I. (2010) Back to the future: using palaeolimnology to infer long-term changes in shallow lake food webs. *Freshwater Biology,* **55,** 600-613.

Reiss, J., Bailey, R.A., Cassio, F., Woodward, G. & Pascoal, C. (2010) Assessing the Contribution of Micro-Organisms and Macrofauna to Biodiversity-Ecosystem Functioning Relationships in Freshwater Microcosms. *Advances in Ecological Research,* **43,** 151-176.

Reiss, J., Bailey, R.A., Perkins, D.M., Pluchinotta, A. & Woodward, G. (2011) Testing effects of consumer richness, evenness and body size on ecosystem functioning. *Journal of Animal Ecology,* **80,** 1145-1154.

Renberg, I. & Hellberg, T. (1982) The Ph History of Lakes in Southwestern Sweden, as Calculated from the Subfossil Diatom Flora of the Sediments. *Ambio,* **11,** 30-33.

Ribeiro, F.J., Przybylski, D., Yin, S.Y., Sharpe, T., Gnerre, S., Abouelleil, A., Berlin, A.M., Montmayeur, A., Shea, T.P., Walker, B.J., Young, S.K., Russ, C., Nusbaum, C., MacCallum, I. & Jaffe, D.B. (2012) Finished bacterial genomes from shotgun sequence data. *Genome Research,* **22,** 2270-2277.

Rodriguez-Lanetty, M., Granados-Cifuentes, C., Barberan, A., Bellantuono, A.J. & Bastidas, C. (2013) Ecological Inferences from a deep screening of the Complex Bacterial Consortia associated with the coral, Porites astreoides. *Molecular Ecology,* **22,** 4349-4362.

Roesch, L.F., Fulthorpe, R.R., Riva, A., Casella, G., Hadwin, A.K.M., Kent, A.D., Daroub, S.H., Camargo, F.A.O., Farmerie, W.G. & Triplett, E.W. (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *Isme Journal,* **1,** 283-290.

Rosenberg, D.M. & Resh, V.H. (1993) *Freshwater biomonitoring and benthic macroinvertebrates*. Chapman and Hall, New York.

Ryan, J.A. & Hightower, L.E. (1996) Stress proteins as molecular biomarkers for environmental toxicology. *Exs,* **77,** 411-424.

Rybicki, N.B. & Landwehr, J.M. (2007) Long-term changes in abundance and diversity of macrophyte and waterfowl populations in an estuary with exotic macrophytes and

improving water quality. *Limnology and Oceanography,* **52,** 1195-1207.

Scheffer, M., Bascompte, J., Brock, W.A., Brovkin, V., Carpenter, S.R., Dakos, V., Held, H., Van Nes, E.H., Rietkerk, M. & Sugihara, G. (2009) Early-warning signals for critical transitions. *Nature,* **461,** 53-59.

Scheffer, M. & Carpenter, S.R. (2003) Catastrophic regime shifts in ecosystems: linking theory to observation. *Trends in Ecology & Evolution,* **18,** 648-656.

Schindler, D.E., Hilborn, R., Chasco, B., Boatright, C.P., Quinn, T.P., Rogers, L.A. & Webster, M.S. (2010) Population diversity and the portfolio effect in an exploited species. *Nature,* **465,** 609-612.

Schindler, D.W. (1987) Detecting ecosystem responses to anthropogenic stress. *Canadian Journal of Fisheries and Aquatic Sciences,* **44,** s6-s25.

Schindler, D.W. (1990) Experimental perturbations of whole lakes as tests of hypotheses concerning ecosystem structure and function. *Oikos***,** 25-41.

Schneider, G.F. & Dekker, C. (2012) DNA sequencing with nanopores. *Nature Biotechnology,* **30,** 326-328.

Simpson, G.L., Shilland, E.M., Winterbottom, J.M. & Keay, J. (2005) Defining reference conditions for acidified waters using a modern analogue approach. *Environmental Pollution,* **137,** 119-133.

Sinsabaugh, R.L., Hill, B.H. & Shah, J.J.F. (2009) Ecoenzymatic stoichiometry of microbial organic nutrient acquisition in soil and sediment. *Nature,* **462,** 795-798.

Simpson, J.C. & Norris, R.H. (2000) Biological assessment of river quality: Development of AUSRIVAS models and outputs. Freshwater Biological Association, The Ferry House, Far Sawrey, Ambleside, Cumbria, LA22 0LP, UK.

Slavik, K., Peterson, B., Deegan, L., Bowden, W., Hershey, A. & Hobbie, J. (2004) Long-term responses of the Kuparuk River ecosystem to phosphorus fertilization. *Ecology,* **85,** 939-954.

Smol, J.P. (2009) *Pollution of lakes and rivers: a paleoenvironmental perspective*. John Wiley & Sons.

Statzner, B. & Bêche, L.A. (2010) Can biological invertebrate traits resolve effects of multiple stressors on running water ecosystems? *Freshwater Biology,* **55,** 80-119.

Sun, M.Y., Dafforn, K.A., Brown, M.V. & Johnston, E.L. (2012) Bacterial communities are sensitive indicators of contaminant stress. *Mar Pollut Bull,* **64,** 1029-1038.

Sweeney, B.W., Bott, T.L., Jackson, J.K., Kaplan, L.A., Newbold, J.D., Standley, L.J., Hession, W.C. & Horwitz, R.J. (2004) Riparian deforestation, stream narrowing, and loss of stream ecosystem services. *Proceedings of the National Academy of Sciences of the United States of America,* **101,** 14132-14137.

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. & Willerslev, E. (2012) Towards next-

generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology,* **21,** 2045-2050.

Thuiller, W., Lavergne, S., Roquet, C., Boulangeat, I., Lafourcade, B. & Araujo, M.B. (2011) Consequences of climate change on the tree of life in Europe. *Nature,* **470,** 531-534.

Tilman, D., Knops, J., Wedin, D., Reich, P., Ritchie, M. & Siemann, E. (1997) The influence of functional diversity and composition on ecosystem processes. *Science,* **277,** 1300-1302.

Tokeshi, M. (1990) Niche Apportionment or Random Assortment - Species Abundance Patterns Revisited. *Journal of Animal Ecology,* **59,** 1129-1146.

Tokeshi, M. (1993) Species Abundance Patterns and Community. *Advances in Ecological Research,* **24,** 111.

UK National Ecosystem Assessment, (2011) The UK National Ecosystem Assessment: Synthesis of the Key Findings. Cambridge.

Van Aggelen, G., Ankley, G.T., Baldwin, W.S., Bearden, D.W., Benson, W.H., Chipman, J.K., Collette, T.W., Craft, J.A., Denslow, N.D. & Embry, M.R. (2010) Integrating omic technologies into aquatic ecological risk assessment and environmental monitoring: hurdles, achievements, and future outlook. *Environmental health perspectives,* **118,** 1.

Venail, P., MacLean, R., Bouvier, T., Brockhurst, M., Hochberg, M. & Mouquet, N. (2008) Diversity and productivity peak at intermediate dispersal rate in evolving metacommunities. *Nature,* **452,** 210-214.

Vonlanthen, P., Bittner, D., Hudson, A.G., Young, K.A., Muller, R., Lundsgaard-Hansen, B., Roy, D., Di Piazza, S., Largiader, C.R. & Seehausen, O. (2012) Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature,* **482,** 357-U1500.

Webb, C., Ackerly, D. & Kembel, S. (2011) Software for the analysis of phylogenetic comunity structure and character evolution (with phylomatic and ecovolve) user's manual version 4.2. Arnold Arboretum of Harvard University.

Willby, N.J., Abernethy, V.J. & Demars, B.O.L. (2000) Attribute-based classification of European hydrophytes and its relationship to habitat utilization. *Freshwater Biology,* **43,** 43-74.

Williams, T.D., Turan, N., Diab, A.M., Wu, H.F., Mackenzie, C., Bartie, K.L., Hrydziuszko, O., Lyons, B.P., Stentiford, G.D., Herbert, J.M., Abraham, J.K., Katsiadaki, I., Leaver, M.J., Taggart, J.B., George, S.G., Viant, M.R., Chipman, K.J. & Falciani, F. (2011) Towards a System Level Understanding of Non-Model Organisms Sampled from the Environment: A Network Biology Approach. *Plos Computational Biology,* **7**.

Woodward, G. (2009) Biodiversity, ecosystem functioning and food webs in fresh waters: assembling the jigsaw puzzle. *Freshwater Biology,* **54,** 2171-2187.

Woodward, G., Benstead, J.P., Beveridge, O.S., Blanchard, J., Brey, T., Brown, L.E., Cross,

W.F., Friberg, N., Ings, T.C., Jacob, U., Jennings, S., Ledger, M.E., Milner, A.M., Montoya, J.M., O'Gorman, E.J., Olesen, J.M., Petchey, O.L., Pichler, D.E., Reuman, D.C., Thompson, M.S.A., Van Veen, F.J.F. & Yvon-Durocher, G. (2010) Ecological Networks in a Changing Climate. *Advances in Ecological Research,* **42,** 71-138.

Woodward, G., Brown, L.E., Edwards, F.K., Hudson, L.N., Milner, A.M., Reuman, D.C. & Ledger, M.E. (2012) Climate change impacts in multispecies systems: drought alters food web size structure in a field experiment. *Philosophical Transactions of the Royal Society B-Biological Sciences,* **367,** 2990-2997.

Woodward, G., Gray, C. & Baird, D.J. (2013) A critique of biomonitoring in an age of globalisation and emerging environmental threats. *Limnetica*.

Yachi, S. & Loreau, M. (1999) Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proceedings of the National Academy of Sciences,* **96,** 1463-1468.

Yagishita, M. (1995) Establishing an acid deposition monitoring network in East Asia. *Water, Air, and Soil Pollution,* **85,** 273-278.

Yergeau, E., Lawrence, J.R., Sanschagrin, S., Waiser, M.J., Korber, D.R. & Greer, C.W. (2012) Next-Generation Sequencing of Microbial Communities in the Athabasca River and Its Tributaries in Relation to Oil Sands Mining Activities. *Applied and Environmental Microbiology,* **78,** 7626-7637.

Young, R.G., Matthaei, C.D. & Townsend, C.R. (2008) Organic matter breakdown and ecosystem metabolism: functional indicators for assessing river ecosystem health. Journal of the North American Benthological Society, 27, 605-625.

Yvon-Durocher, G., Jones, J.I., Trimmer, M., Woodward, G. & Montoya, J.M. (2010) Warming alters the metabolic balance of ecosystems. *Philosophical Transactions of the Royal Society B-Biological Sciences,* **365,** 2117-2126.

Zarraonaindia, I., Smith, D.P. & Gilbert, J.A. (2013) Beyond the genome: community-level analysis of the microbial world. *Biology & philosophy***,** 1-22.

# Appendix E | Supplementary material for Chapter 3

Table A. Results from the $\chi^2$ contingency test (see main text). p-values significant at the 0.1 level are highlighted in bold.

| | ave.pH | ave.ANC | ave.DOC | ave.L.Al |
|---|---|---|---|---|
| meanTH | 0.371314 | 1 | 0.530735 | 0.412294 |
| maxTH | 0.241879 | 1 | 0.350825 | 0.245877 |
| E | **0.090955** | 1 | 1 | 0.654173 |
| Vulnerability | 1 | 0.634683 | 1 | 0.381809 |
| Generality | 1 | **0.097951** | 0.145427 | 1 |
| sd.V | 0.418791 | 1 | 0.22039 | 0.668166 |
| sd.G | **0.087956** | 1 | 0.577211 | 0.662169 |
| redundancy | 0.690655 | 0.635182 | 0.570715 | 0.68066 |

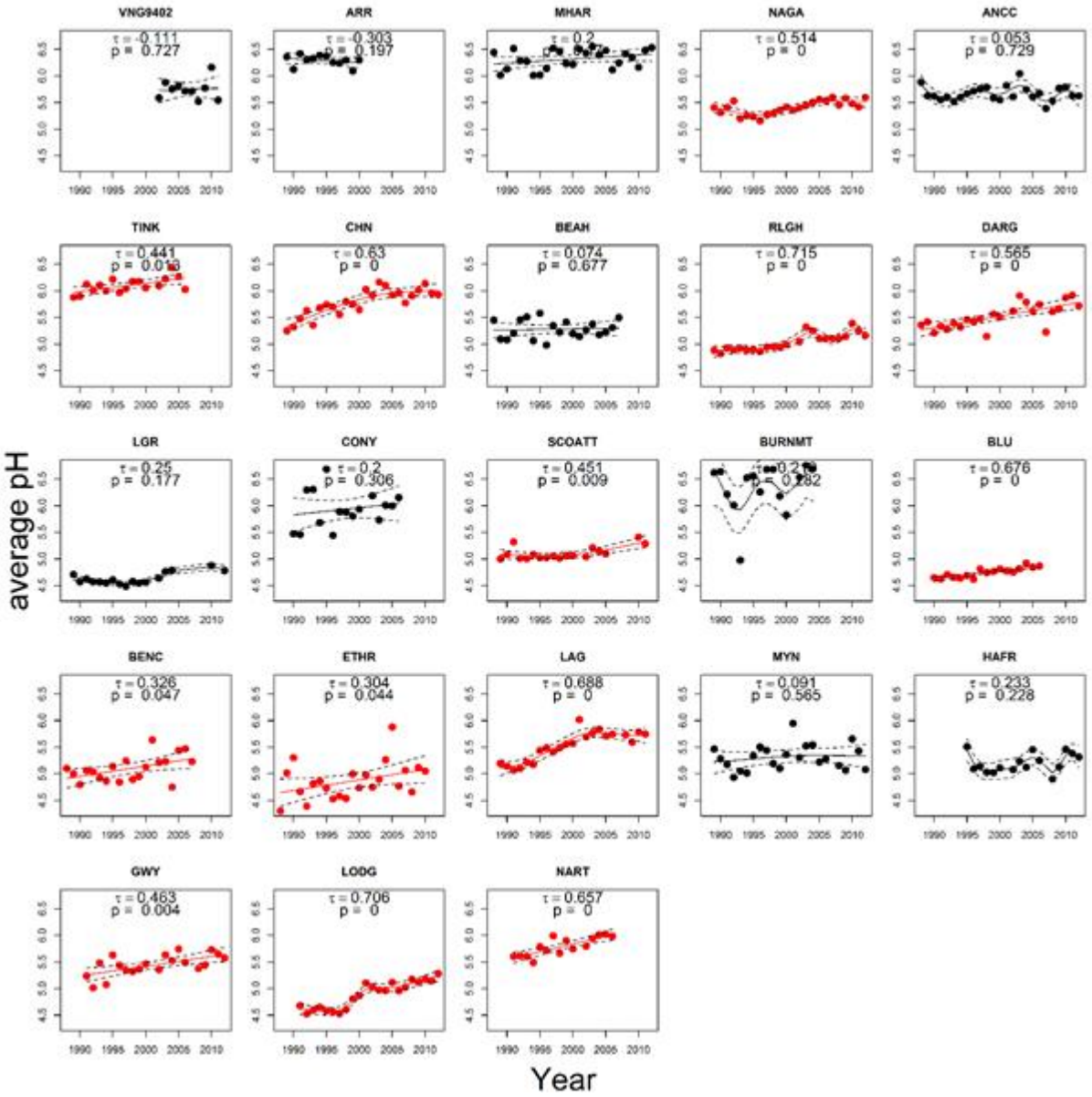



Figure C. Trends in average annual pH at each of the UWMN sites. Sites are arranged in order of their decreasing latitude, which can be used as a proxy for their initial acidified state, more acidified sites were generally in the south, while the least acidified sites were more northern.

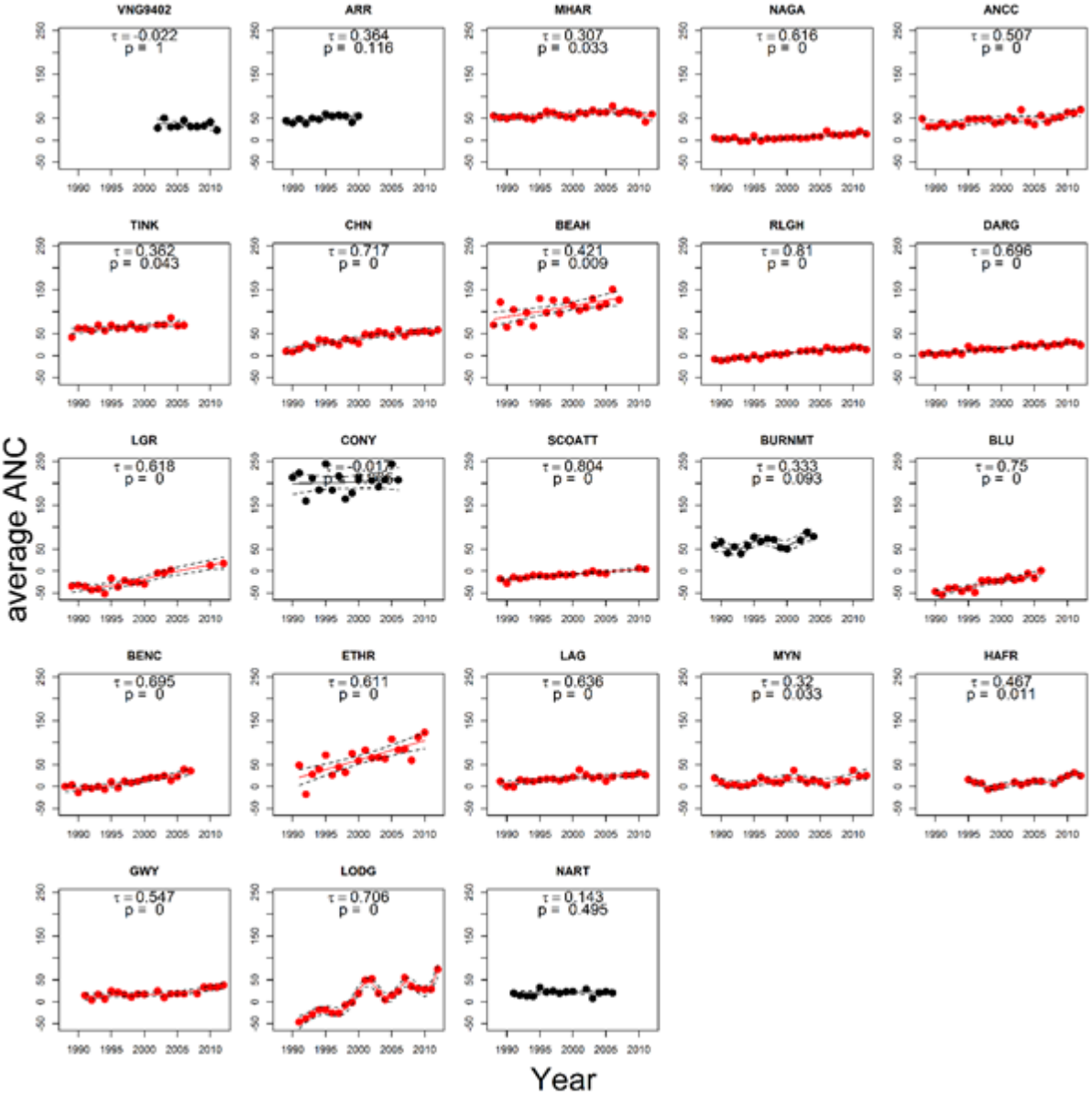



Figure D. Trends in average annual Acid Neutralising Capacity at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

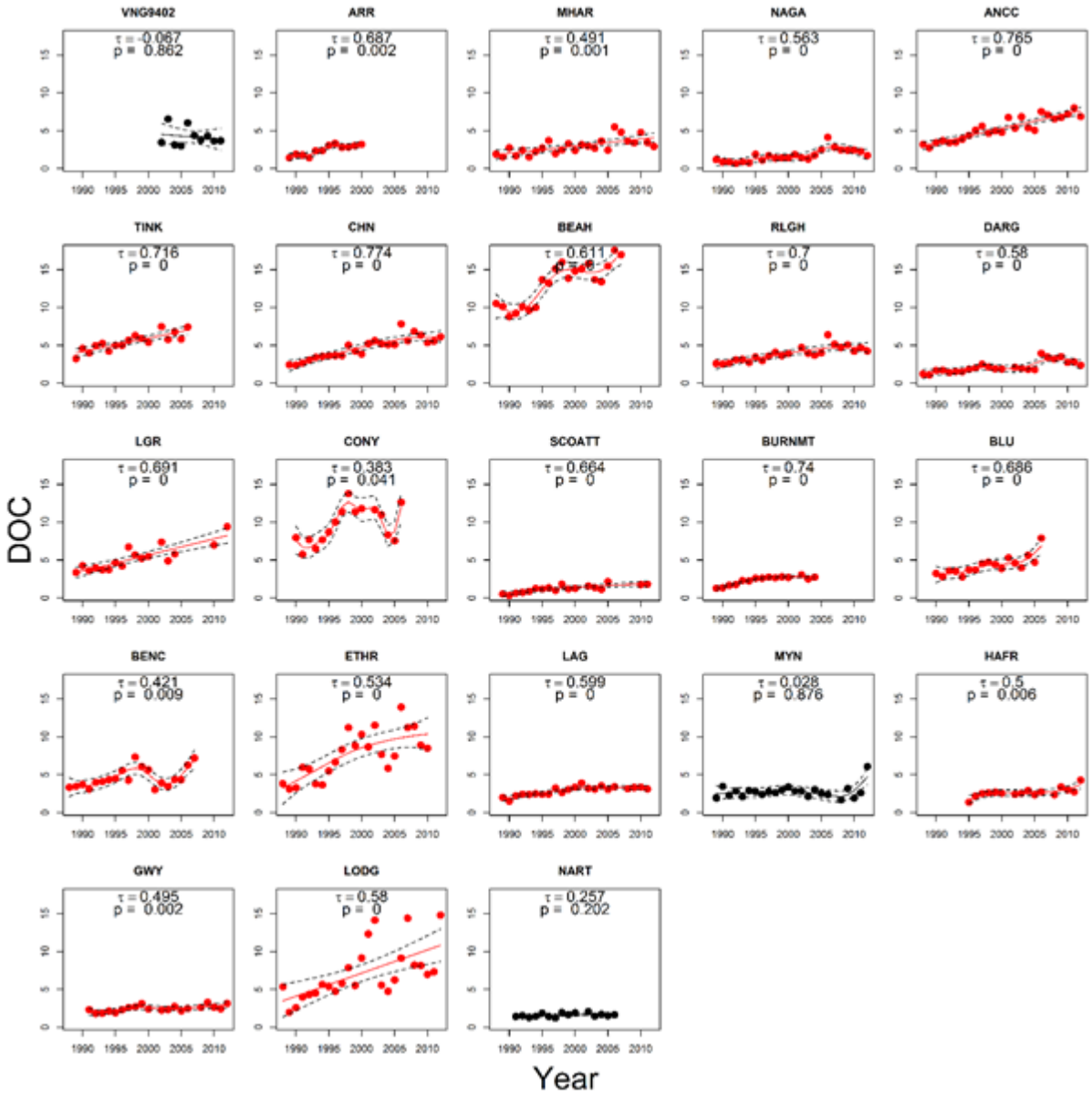



Figure E. Trends in average annual Dissolved Organic Carbon at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

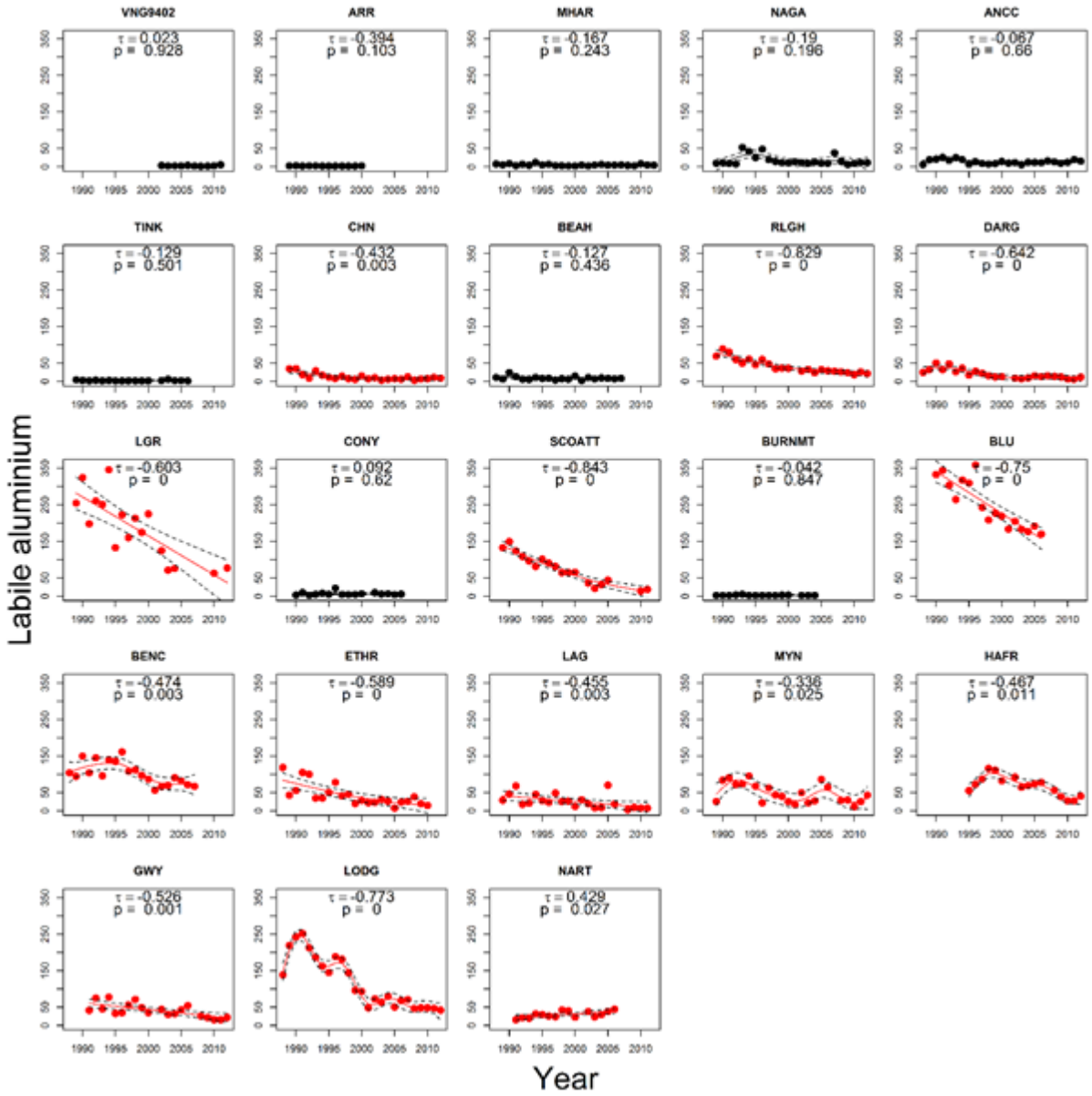



Figure F. Trends in average annual labile aluminium at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

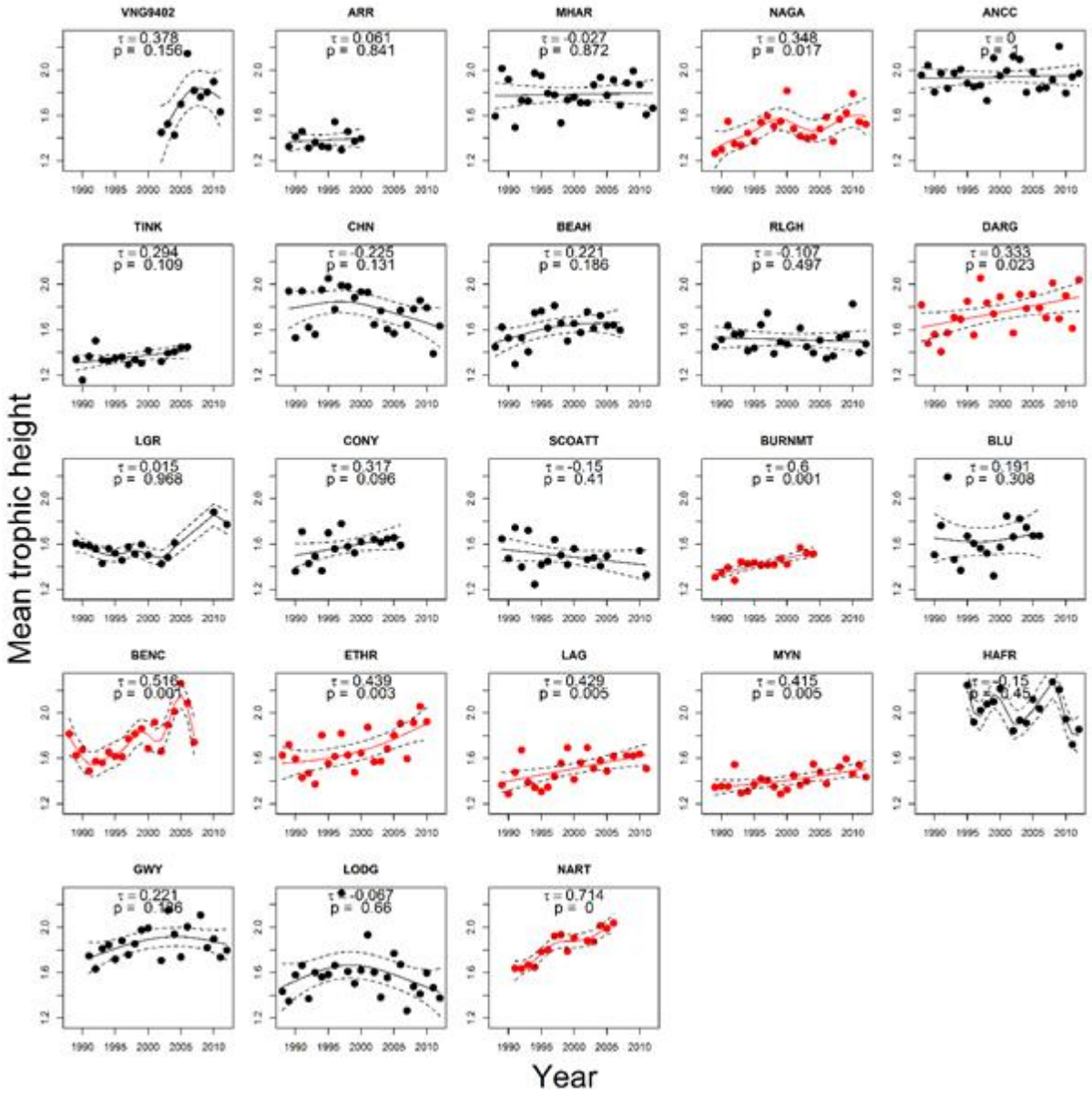



Figure G. Trends in mean trophic height at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

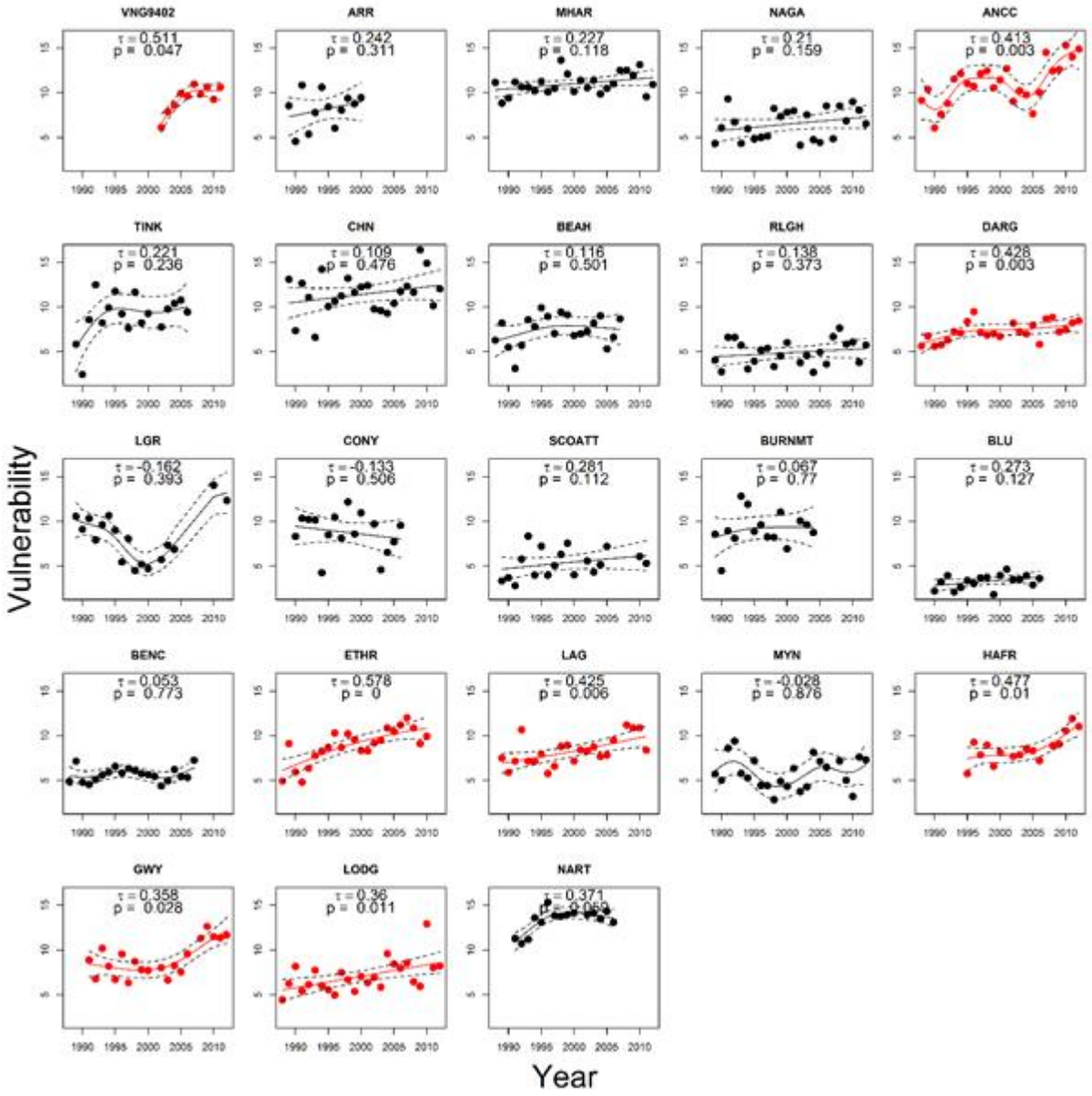



Figure H. Trends in average food web vulnerability at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

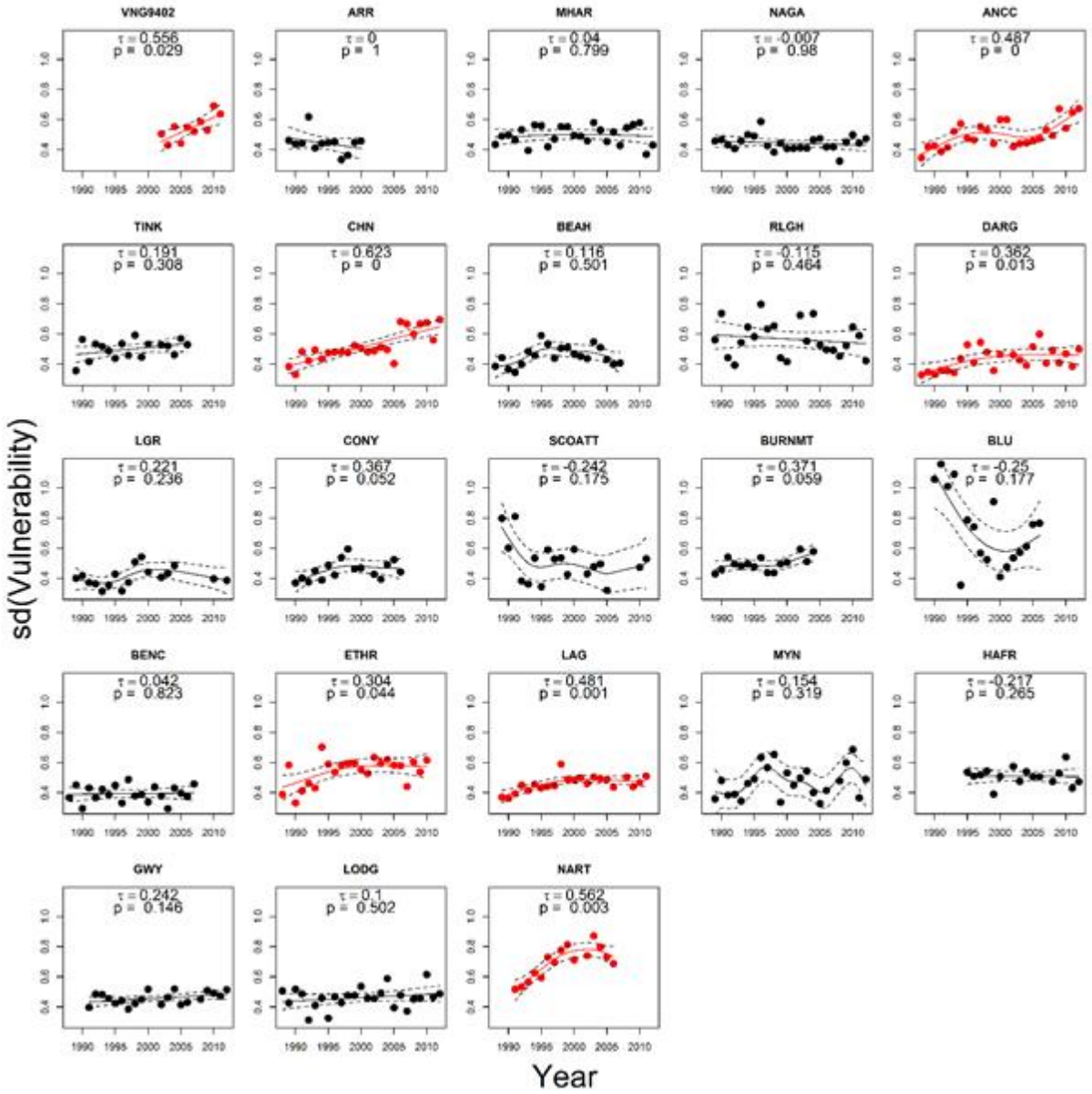



Figure I. Trends in the standard deviation of food web vulnerability at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

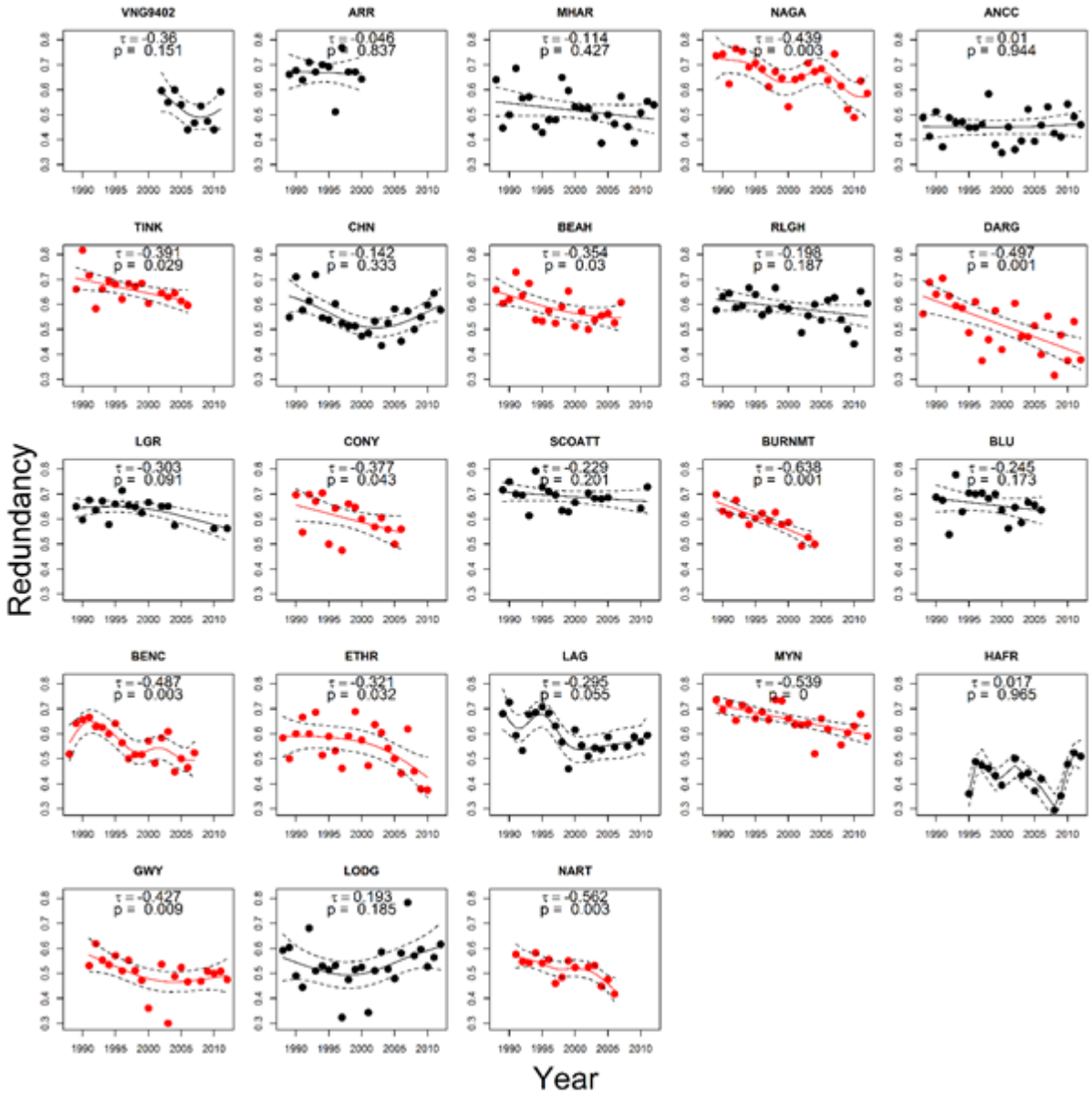



Figure J. Trends in food web redundancy at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

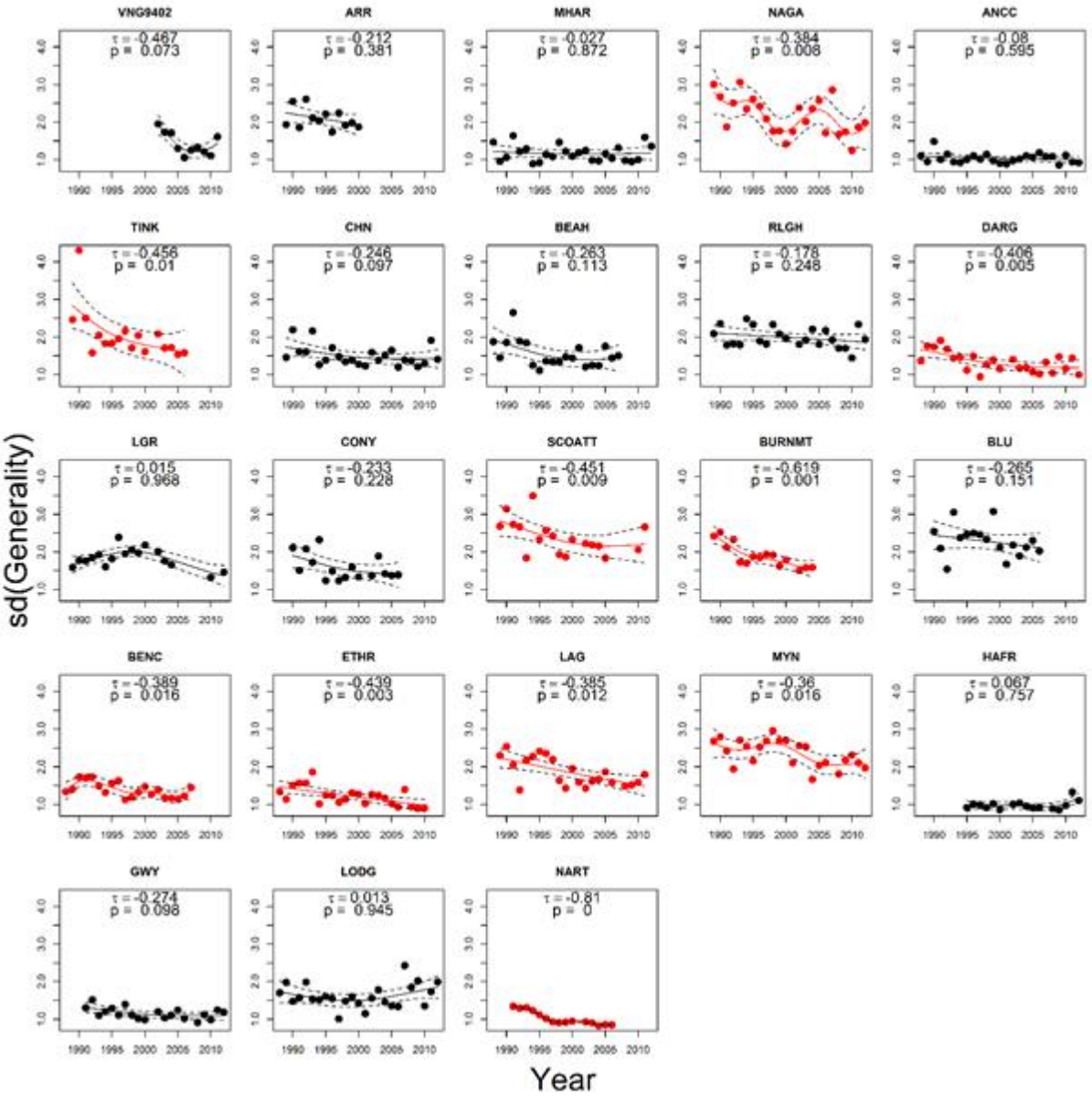



Figure K. Trends in the standard deviation of food web generality at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

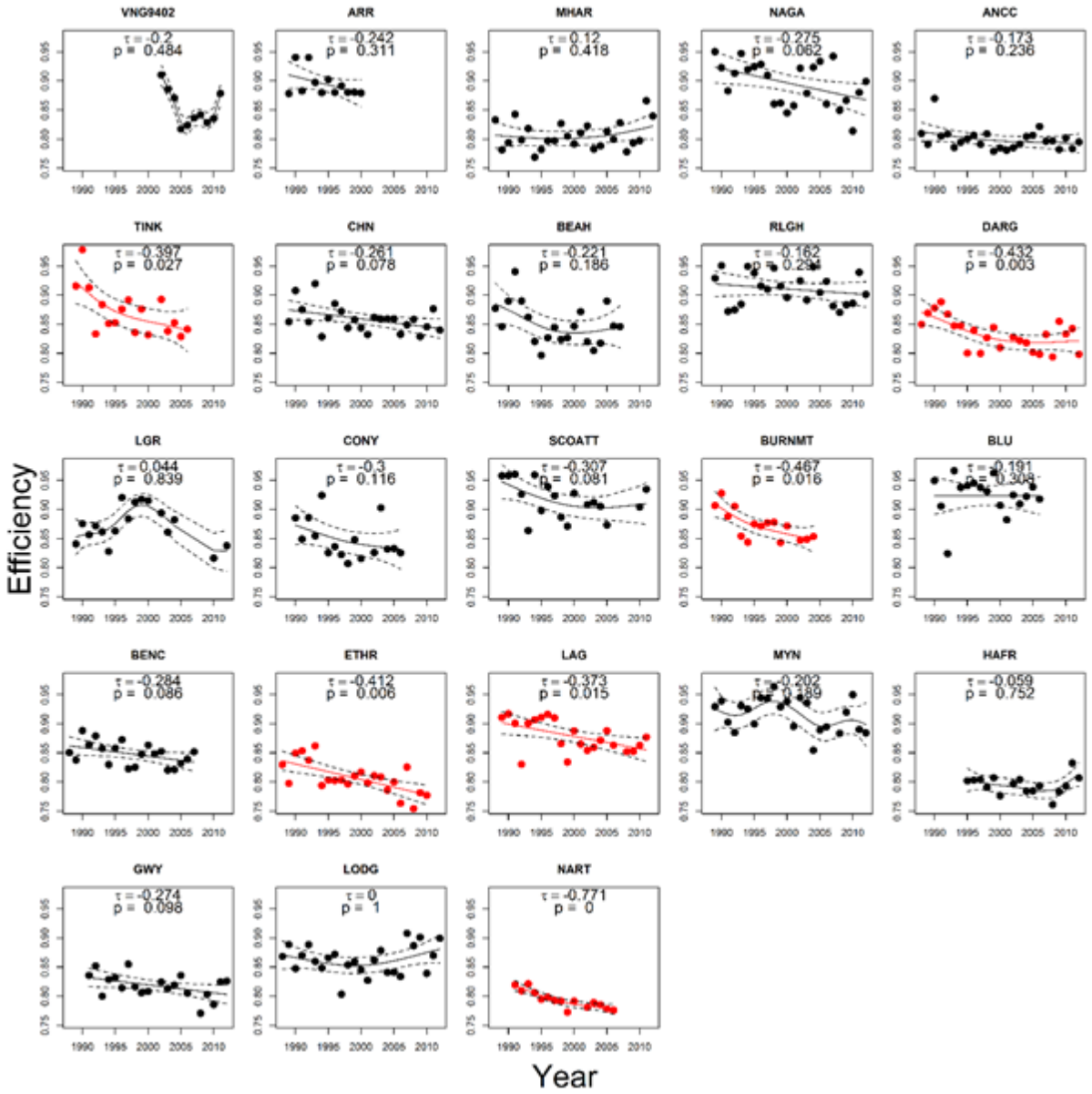



Figure L. Trends in food web efficiency at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

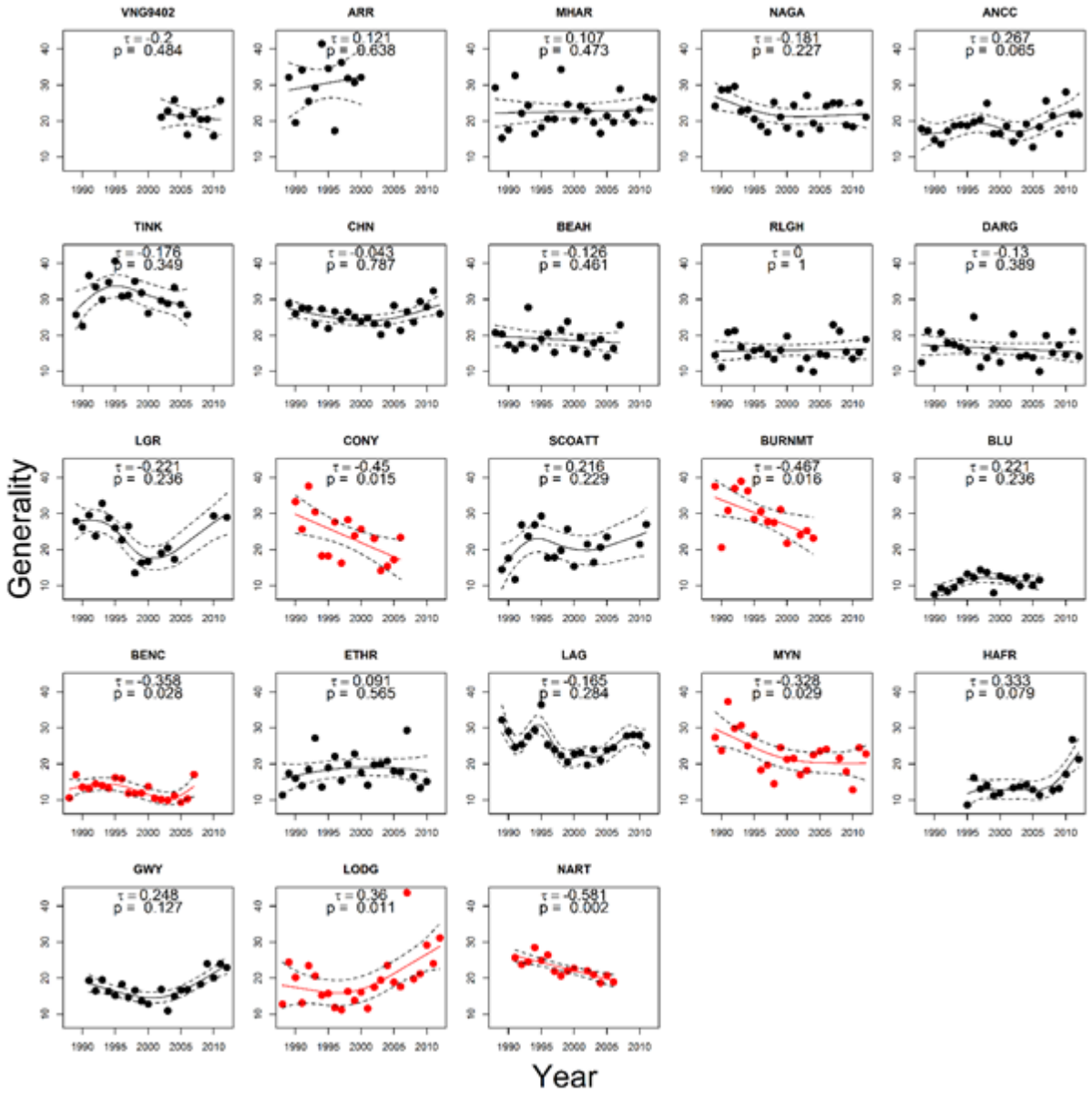



Figure M. Trends in food web generality at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

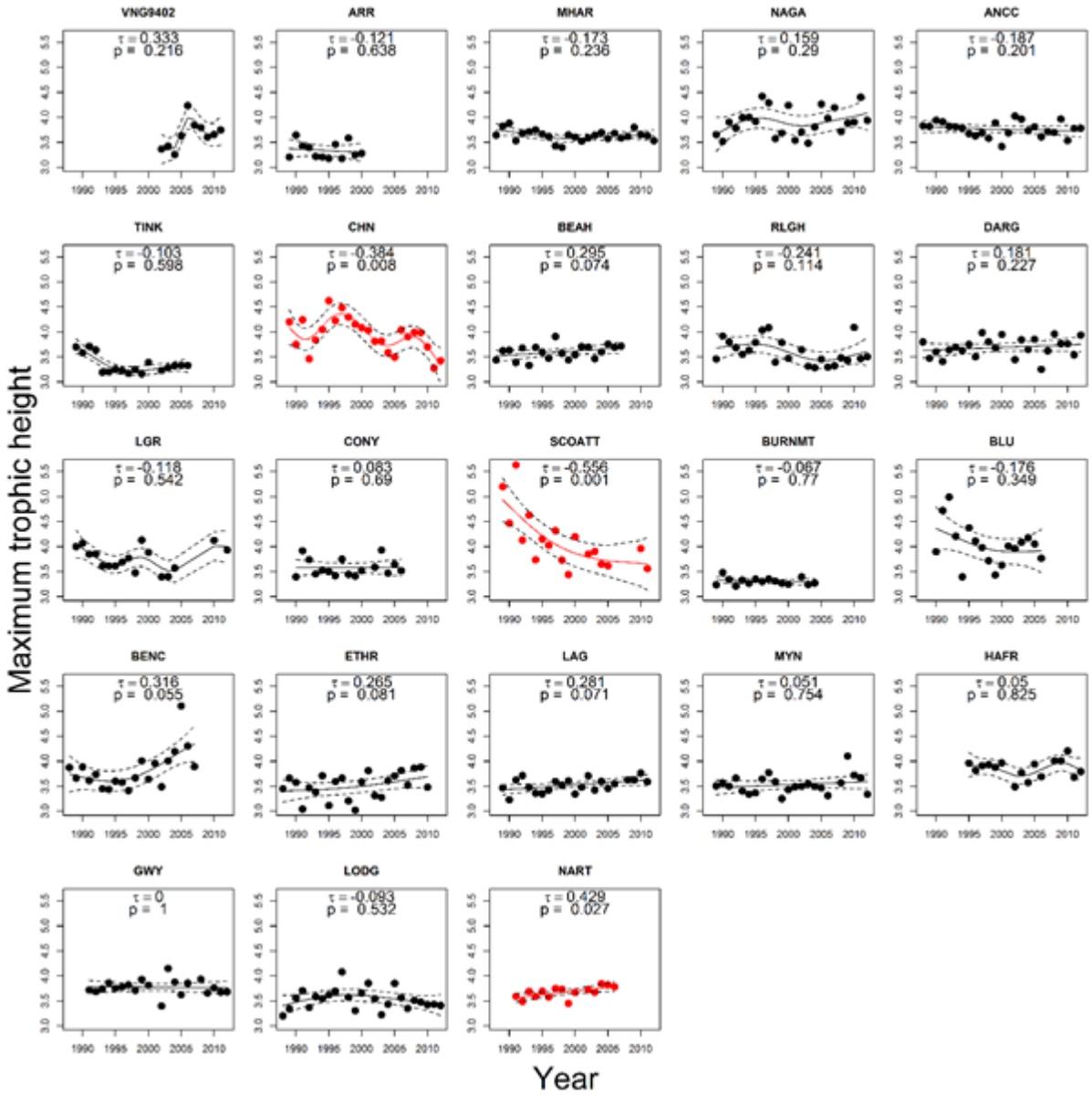



Figure N. Trends in maximum trophic height at each of the UWMN sites. Site ordering is explained in the legend of Figure C.

# Appendix F | Supplementary material for Chapter 4

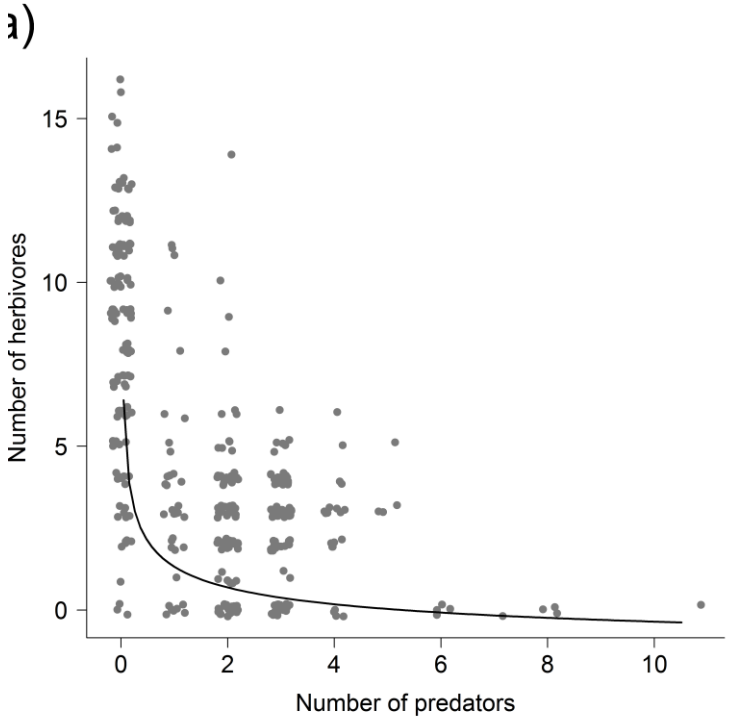



**Figure A**. The number of herbivores and predators in each food web. The line shows the fit of a general linear model (see main text for details).

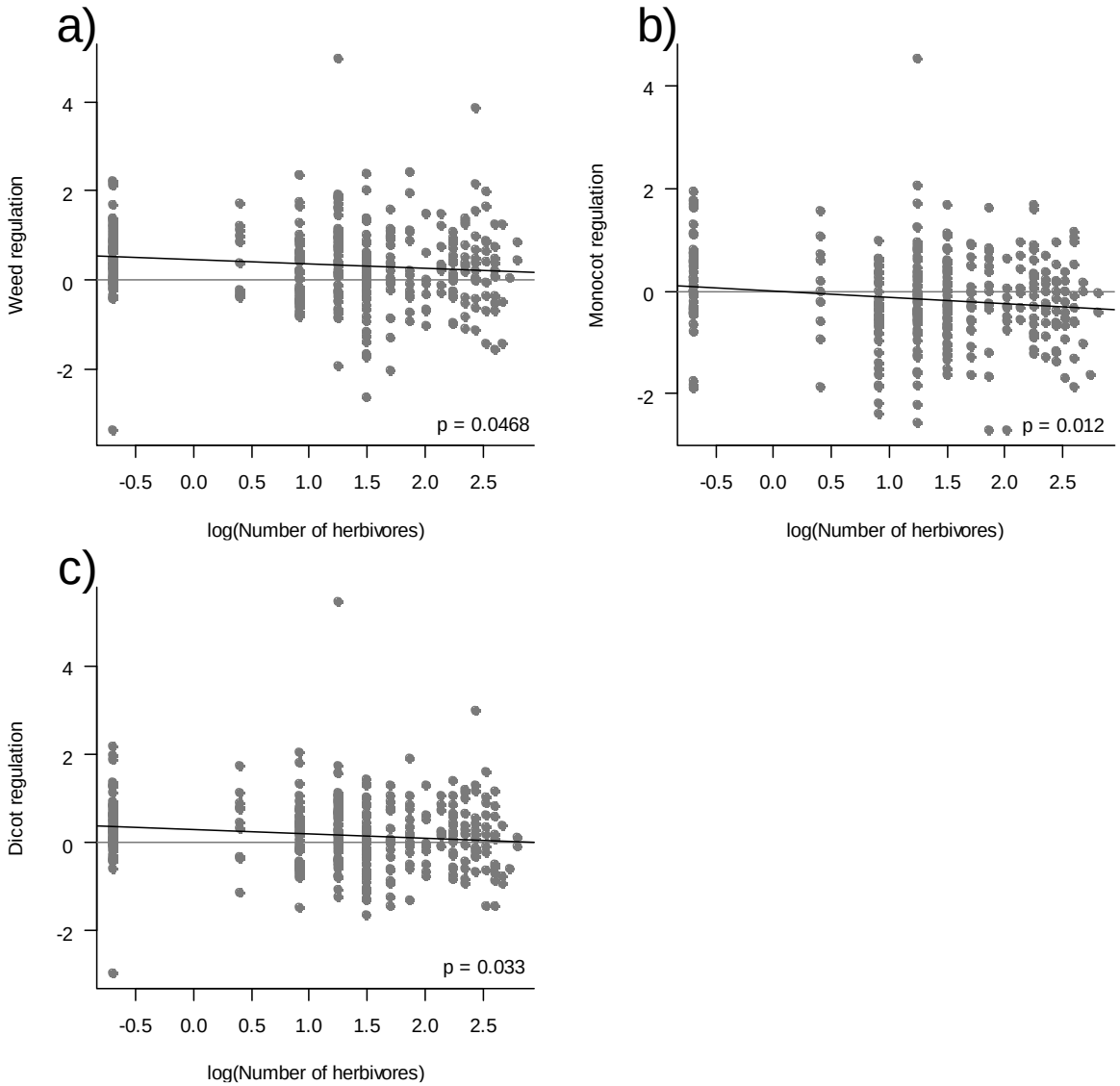



**Figure B**. The level of total weed regulation (a), monocot regulation (b) and dicot regulation (c) related to the number of herbivores in each network.

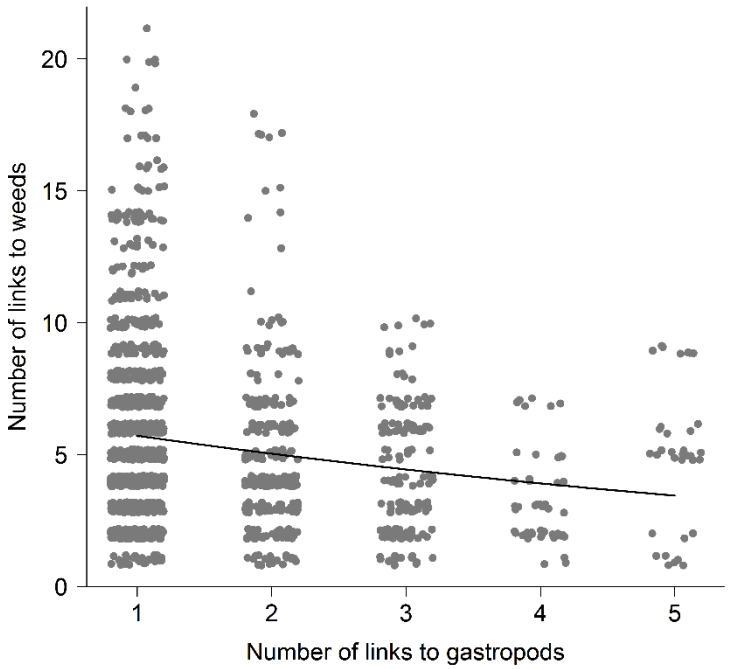



**Figure C**. The number of links to weeds and gastropods for omnivore nodes only within each food web.

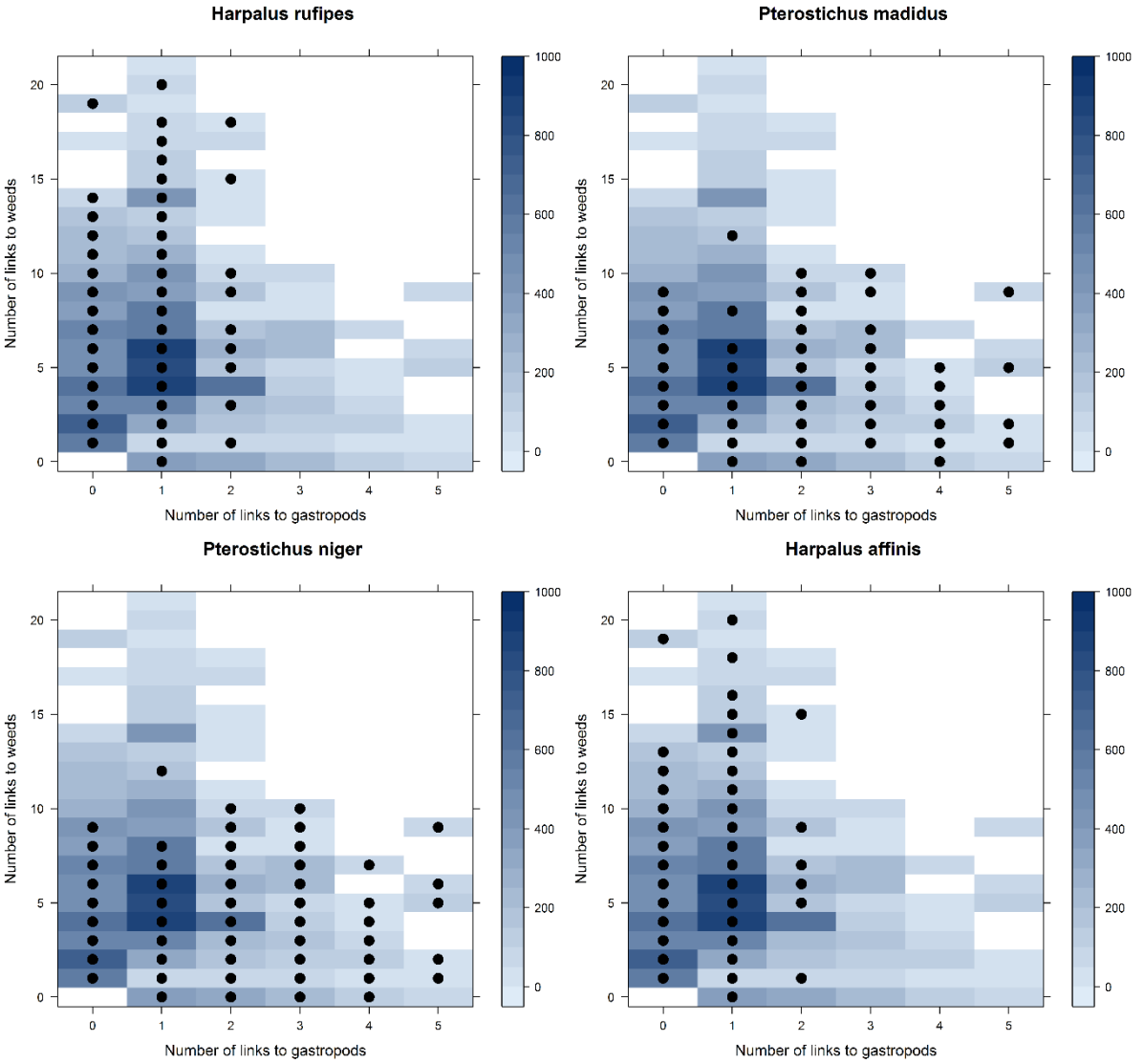



**Figure D**. A density plot showing the trade-off in herbivore links and predator links. Within each food web, for each carabid species, the number of links to weed and gastropod resources are plotted. Some carabids were pure herbivores or predators, but most were omnivores. Colour indicates the count of each particular weed-gastropod link combination. The occupation of this space of potential feeding interactions for the most common carabid species *Pterostichus melanarius* is shown in the main text, the next four most common carabid species are shown here in black.

**Table A**. statistics of fit for the multiple linear regressions and Generalised Linear Models discussed in the main text.

| **Linear regressions** | | | | |
|---|---|---|---|---|
| Response variable | Explanatory variable | | F statistic | p-value |
| Log(specialist herbivore interaction frequency) | Number of weed species | | $F_{1,365} = 11.883$ | 0.0006 |
| Log(specialist predator interaction frequency) | Number of gastropod species | | $F_{1,258} = 15.47$ | 0.0001 |
| Log(all herbivore interaction frequency) | Number of weed species | | $F_{1,365} = 5.41$ | 0.020 |
| Log(all predatory interaction frequency) | Number of gastropod species | | $F_{1,258} = 0.04$ | 0.84 |
| Log(omnivore predatory links) | Log(omnivore herbivory links) | | $F_{1,1583} = 66.981$ | <0.0001 |
| Total weed regulation | Log(number of herbivores) | | $F_{1,333} = 3.98$ | 0.047 |
| Monocot weed regulation | Log(number of herbivores) | | $F_{1,333} = 6.42$ | 0.012 |
| Dicot weed regulation | Log(number of herbivores) | | $F_{1,333} = 4.57$ | 0.033 |
| Total weed regulation | Log(specialist herbivore interaction frequency) | | $F_{1,333} = 5.16$ | 0.024 |
| Monocot weed regulation | Log(specialist herbivore interaction frequency) | | $F_{1,333} = 3.89$ | 0.049 |
| Dicot weed regulation | Log(specialist herbivore interaction frequency) | | $F_{1,333} = 5.76$ | 0.017 |
| Total weed regulation | Log(all carabid - weed interaction frequency) | | $F_{1,333} = 0.117$ | 0.907 |
| Monocot weed regulation | Log(all carabid - weed interaction frequency) | | $F_{1,333} = 0.097$ | 0.756 |
| Dicot weed regulation | Log(all carabid - weed interaction frequency) | | $F_{1,333} = 0.398$ | 0.528 |
| **Generalised linear models** | | | | |
| Response variable | Explanatory variable | Error distribution | F statistic | p-value |
| Number of herbivores | Log(number of predators + 0.05) | quasipoisson | $F_{1,372} = 339.5$ | <0.0001 |
| Number of omnivore links to weeds | Log(number of omnivore links to gastropods) | poisson | | <0.0001 |

# Appendix G | Supplementary material for Chapter 5

## List of identification keys used

Bass, J. (1998) *Last-Instar Larvae and Pupae of the Simuliidae of Britain and Ireland: a Key with Brief Ecological Notes.* Freshwater Biological Association Scientific Publication No. **55**. Titus Wilson and Son Ltd., Kendal, UK.

Brooks, S. (2004) *Field guide to the dragonflies and damselflies of Great Britain and Ireland*. British Wildlife Publishing, Gillingham, UK.

Edington, J.M. & Hildrew, A.G. (1995) *Caseless caddis larvae of the British Isles.* Freshwater Biological Association Scientific Publication No. **53**. Titus Wilson and Son Ltd., Kendal, UK.

Elliott, J.M., Humpesch, U.H. & Macan, T.T. (1988) Larvae of the British Ephemeroptera. Freshwater Biological Association Scientific Publication No. **49**. Titus Wilson and Son Ltd., Kendal, UK.

Friday, L.E. (1988) *A key to the adults of British water beetles*. Fields Studies, Vol 7 No.1, Dorset Press, Dorset, UK.

Harding, P. (Ed.) (2004) *Freshwater Fishes in Britain: the species and their distribution.* Harley Books, Colchester, UK.

Hynes, H.B.N. (1993) *Adults and nymphs of the British Stoneflies (Plecoptera).* Freshwater Biological Association Scientific Publication No. **17**. Titus Wilson and Son Ltd., Kendal, UK.

Krammer, K. & Lange-Bertalot, H. (1986) Bacillariophyceae, 1. Teil: Naviculaceae. In: *Süßwasserflora von Mitteleuropa* (Eds. H. Ettl, J. Gerloff, H. Heynig, D. Mollenhauer), Gustav Fischer Verlag, Stuttgart, Germany.

Krammer, K. & Lange-Bertalot, H. (1988) Bacillariophyceae, 2. Teil: Bacillariaceae, Epithemiaceae, Surirellaceae. In: *Süßwasserflora von Mitteleuropa* (Eds. H. Ettl, J. Gerloff, H. Heynig, D. Mollenhauer), Gustav Fischer Verlag, Stuttgart, Germany.

Krammer, K. & Lange-Bertalot, H. (1991a) Bacillariophyceae, 3. Teil: Centrales, Fragilariaceae, Eunotiaceae. In: *Süßwasserflora von Mitteleuropa* (Eds. H. Ettl, J. Gerloff, H. Heynig, D. Mollenhauer), Gustav Fischer Verlag, Stuttgart, Germany.

Krammer, K. & Lange-Bertalot, H. (1991b) Bacillariophyceae, 4. Teil: Achnanthaceae, Kritische Ergänzungen zu Navicula (Lineolatae) und Gomphonema. In: *Süßwasserflora von Mitteleuropa* (Eds. H. Ettl, J. Gerloff, H. Heynig, D. Mollenhauer), Gustav Fischer Verlag, Stuttgart, Germany.

Macan, T.T. (1960) *A key to the British fresh- and brackish-water Gastropods.* Freshwater Biological Association Scientific Publication No. **13**. Titus Wilson and Son Ltd., Kendal, UK.

Maitland, P.S. (2004) *Keys to the Freshwater Fish of Britain and Ireland, with notes on their*

*distribution and ecology.* Freshwater Biological Association Scientific Publication No. **62**. Titus Wilson and Son Ltd., Kendal, UK.

Wallace, I.D., Wallace, B. & Philipson, G.N. (2003) *Keys to the case-bearing caddis larvae of Britain and Ireland.* Freshwater Biological Association Scientific Publication No. **61**. Titus Wilson and Son Ltd., Kendal, UK.

Savage, A.A. (1989) *Adults of the British Aquatic Hemiptera Heteroptera: a key with ecological notes.* Freshwater Biological Association Scientific Publication No. **50**. Titus Wilson and Son Ltd., Kendal, UK.

Table A. Equations used to calculate macroinvertebrate individual dry mass (DM). HW: head-capsule width (mm); BL: total body length (mm); SL: shell length (mm). Morphologically similar taxa or higher taxonomic levels, shown in parentheses, were used where equations were unavailable for a given taxon.

| Taxon | y | X | Regression equation | $R^2$ | Reference |
|---|---|---|---|---|---|
| *Agapetus fuscipes* (*Glossosoma*) | ln(DM) | ln(HW) | y = -6.2 + 3.75x | 0.71 | Meyer (1989) |
| *Ancylus fluviatilis* | $\log_{10}$(DM) | $\log_{10}$(SL) | y = 1.913 + 3.3x | 0.99 | Calow (1975) |
| *Asellus aquaticus* | ln(DM) | Ln(BL) | $y = 1.2688 \times x^{3.326}$ | 0.69 | Baumgärtner & Rothhaupt (2003) |
| *Asellus meridianus* (*Asellus aquaticus*) | ln(DM) | ln(BL) | $y = 1.2688 \times x^{3.326}$ | 0.69 | Baumgärtner & Rothhaupt (2003) |
| *Athripsodes* (*Oecetis* spp.) | ln(DM) | ln(HW) | $y = 1.2688 \times x^{3.326}$ | 0.67 | Benke *et al.* (1999) |
| *Baetis* (*Baetis* spp.) | DM | HW | $y = 1.2688 \times x^{3.326}$ | 0.96 | Benke *et al.* (1999) |
| *Baetis rhodani* (*Baetis* spp.) | DM | HW | y = -4.13 + 1.12x | 0.96 | Benke *et al.* (1999) |
| *Baetis scambus* (*Baetis* spp.) | DM | HW | y = -0.91 + 3.35x | 0.96 | Benke *et al.* (1999) |
| *Baetis vernus* (*Baetis* spp.) | DM | HW | $y = 1.2688 \times x^{3.326}$ | 0.96 | Benke *et al.* (1999) |
| *Bezzia* sp. | ln(DM) | Ln(BL) | $y = 2.7842 \times x^{2.835}$ | 0.99 | Baumgärtner & Rothhaupt (2003) |
| *Caenis rivulorum* (*Caenis* spp.) | ln(DM) | lnHW | $y = 0.0089 \times x^{2.145}$ | 0.63 | Baumgärtner & Rothhaupt (2003) |
| *Centroptilum luteolum* (*Baetis* spp.) | DM | HW | y = -5.53 + 1.91x | 0.96 | Benke *et al.* (1999) |
| Chironomid (Chironomidae) | DM | HW | y = 0.4109 + 3.1678x | 0.9 | Benke *et al.* (1999) |
| *Dendrocoelum lacteum* (*Dugesia tigrina*) | DM | BL | y = -4.4518 + 2.4724 | 0.81 | Benke *et al.* (1999) |
| *Dicranota* sp. | ln(DM) | ln(BL) | y = -5.46 + 4.33x | 0.54 | Woodward & Hildrew |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | (2001) |
| *Drusus annulatus* (Limnephilidae) | ln(DM) | lnHW | y = -6.078 + 3.092x | 0.83 | Meyer (1989) |
| *Dysticidae* sp. (Coleoptera, larvae) | ln(DM) | ln(BL) | y = -6.21 + 2.52x | 0.57 | Meyer (1989) |
| *Elmis aenea* (Coleoptera, adults) | ln(DM) | ln(BL) | y = -3.20 + 2.22x | 0.78 | Burgherr & Meyer (1997) |
| *Elmis aenea* (Elmidae, larvae) | ln(DM) | ln(BL) | y = -4.95 + 2.83x | 0.83 | Towers *et al.* (1994) |
| *Eloeophila* sp. (Diptera) | ln(DM) | ln(BL) | y = -2.12 + 2x | 0.83 | Burgherr & Meyer (1997) |
| *Erpobdella octoculata* | ln(DM) | ln(BL) | y = -2.74 + 2.12x | 0.78 | Edwards et al. (2009) |
| *Gammarus pulex* (*Gammarus fossarum*) | ln(DM) | ln(BL) | y = -2.202 + 1.66 | 0.9 | Burgherr & Meyer (1997) |
| *Glossiphonia complanata* | ln(DM) | ln(BL) | $y = 1.265 \times x^{2.747}$ | 0.64 | Edwards *et al.* (2009) |
| *Helobdella stagnalis* | ln(DM) | ln(BL) | y = 1.30 + 3.62x | 0.62 | Edwards *et al.* (2009) |
| Hydracarina (Hydracarina spp.) | ln(DM) | ln(BL) | y = -4.4518 + 2.4724 | 0.48 | Baumgärtner & Rothhaupt (2003) |
| *Hydropsyche siltalai* (*Hydropsyche* spp.) | DM | HW | y = -4.4518 + 2.4724 | 0.87 | Benke et al. (1999) |
| Hydroptilidae (Trichoptera, cased) | ln(DM) | lnHW | y = 1.30 + 3.62x | 0.82 | Baumgärtner & Rothhaupt (2003) |
| *Hygrobia hermanni* (Coleoptera, larvae) | ln(DM) | ln(BL) | $y = 0.8496 \times x^{3.201}$ | 0.57 | Meyer (1989) |
| *Ilybius* sp. (Coleoptera, larvae) | ln(DM) | ln(BL) | y = 0.4109 + 3.1678x | 0.57 | Meyer (1989) |
| *Lepidostomata hirtum* (Trichoptera, cased) | ln(DM) | lnHW | y = -8.71 + 4.53x | 0.82 | Baumgärtner & Rothhaupt (2003) |
| *Leuctra* spp. (Leuctridae) | DM | HW | y = –4.95 + 2.83x | 0.9 | Benke *et al.* (1999) |
| *Limnephilus lunatus* (Limnephilidae) | ln(DM) | lnHW | y = 1.913 + 3.3x | 0.83 | Meyer (1989) |
| *Limnius volkmari* (*Limnius*, larvae) | ln(DM) | lnHW | $y = (\pi r^2 \times 1.05x)/4$ | 0.7 | Burgherr & Meyer (1997) |
| *Niphargus aquilex* (*Gammarus fossarum*) | ln(DM) | ln(BL) | $y = 0.0618 \times x^{2.502}$ | 0.9 | Burgherr & Meyer (1997) |
| *Oecetis* sp. (*Oecetis* spp.) | ln(DM) | lnHW | y = -8.71 + 4.53x | 0.67 | Benke *et al.* (1999) |

| | | | | | |
|---|---|---|---|---|---|
| Oligochaeta | DM(g) | | y = -6.21 + 2.52x | | Smock (1980) |
| *Oreodytes sanmarkii* (*Hydroporus*, Dytiscidae) | ln(DM) | ln(BL) | y = -0.83 + 4.25x | 0.71 | Benke *et al.* (1999) |
| *Oulimnius tuberculatus* (*Limnius,* larvae) | ln(DM) | lnHW | y = -2.69 + 2.11x | 0.7 | Burgherr & Meyer (1997) |
| *Oxycera* sp. (Diptera) | ln(DM) | ln(BL) | $y = 0.0163 \times x^{2.477}$ | 0.83 | Burgherr & Meyer (1997) |
| *Paraleptophlebia submarginata* (Leptophlebiidae) | ln(DM) | lnHW | y = 2.58 + 2.80x | 0.86 | Burgherr & Meyer (1997) |
| *Piscicola geometra* (Hirudinea spp.) | ln(DM) | ln(BL) | $y = 0.0089 \times x^{2.145}$ | 0.62 | Edwards *et al.* (2009) |
| *Pisidium* sp. | (DM) | SL | y = 0.4109 + 3.1678x | 0.87 | Benke *et al.* (1999) |
| *Plectrocnemia* (*Plectrocnemia conspersa*) | $\log_{\text{Meyer 19890}}$(µg) | $\log_{\text{Meyer 19890}}$HW | y = -6.21 + 2.52x | | Woodward & Hildrew (2001) |
| *Polycelis tenuis* (*Dugesia tigrina*) | (DM) | BL | y = 1.55 + 3.21x | 0.81 | Benke *et al.* (1999) |
| *Potamophylax latipennis* (Limnephilidae) | ln(DM) | lnHW | $y = 0.7255 \times x^{3.325}$ | 0.83 | Meyer (1989) |
| *Psychoda* sp. (Diptera) | ln(DM) | ln(BL) | y = 0.8613 + 3.576x | 0.83 | Burgherr & Meyer (1997) |
| *Rhyacophila dorsalis* | $\log_{\text{Meyer 19890}}$(µg) | $\log_{\text{Meyer 19890}}$HW | y = 0.20 + 3.32 x | 0.72 | Edwards *et al.* (2009) |
| *Serratella ignita* (*Serratella* sp.) | (DM) | HW | $y = 2.1694 \times x^{2.623}$ | 0.72 | Benke et al. (1999) |
| *Silo nigricornis* (Goeridae) | ln(DM) | lnHW | y = -5.30 + 2.36x | 0.75 | Meyer (1989) |
| *Simulium* sp. | ln(DM) | lnHW | y = -6.2 + 3.75x | 0.93 | Burgherr & Meyer (1997) |
| Tanypodinae | (DM) | HW | y = 1.913 + 3.3x | 0.85 | Benke *et al.* (1999) |
| *Tipula yamatotipula* (*Tipula abdominalis*) | ln(DM) | ln(BL) | y = -6.20 + 3.75x | 0.93 | Smock (1980) |

### D.8.1 References

Baumgärtner, D. & Rothhaupt, K.O. (2003). Predictive Length-Dry Mass Regressions for Freshwater Invertebrates in a Pre-Alpine Lake Littoral. *International Review of Hydrobiology*, **88**, 453–463.

Benke, A.C., Huryn, A.D., Smock, L. a & Wallace, J.B. (1999). Length-Mass Relationships for Freshwater Macroinvertebrates in North America with Particular Reference to the Southeastern United States Length-mass relationships for freshwater macroinvertebrates in North America with particular reference to the southeas. *Journal of the North American Benthological Society*, **18**, 308–343.

Burgherr, P. & Meyer, E.I. (1997). Regression analysis of linear body dimensions vs. dry mass in stream macroinvertebrates. *Archiv fur Hydrobiologie*, **139**, 101–112.

Calow, P. (1975). Length-dry weight relationships in snails: some explanatory models. *Journal of Molluscan Studies*, **41**, 357.

Edwards, F.K., Lauridsen, R.B., Fernandes, W.P.A., Beaumont, W.R.C., Ibbotson, A.T., Scott, L., Davies, C.E. & Jones, J.I. (2009). Re-introduction of Atlantic salmon, Salmo salar L., to the Tadnoll Brook, Dorset. *Proceedings of the Dorset Natural History and Archaeological Society*, **130**, 9–16.

Meyer, E. (1989). The relationship between body length parameters and dry mass in running water Invertebrates. *Archiv fur Hydrobiologie*, **117**, 191–203.

Smock, L.A. (1980). Relationships between body size and biomass of aquatic insects. *Freshwater Biology*, **10**, 375–383.

Towers, D.J., Henderson, I.M. & Veltman, C.J. (1994). Predicting dry weight of New Zealand aquatic macroinvertebrates from linear dimensions. *New Zealand Journal of Marine and Freshwater Research*, **28**, 159–166.

Woodward, G. & Hildrew, A.G. (2001). Invasion of a stream food web by a new top predator. *Journal of Animal Ecology*, **70**, 273–288.

**Table B**. The taxonomic resolution (i.e. generality) assigned to each node in the networks to create links between nodes.

| Node | Resolution |
|---|---|
| *Achnanthes clevei* | genus |
| *Achnanthes conspicua* | genus |
| *Achnanthes helvetica* | genus |
| *Achnanthes hungarica* | genus |
| *Achnanthes lanceolata* | genus |
| *Achnanthes lanceolata abbreviata* | genus |
| *Achnanthes lanceolata bimaculata* | genus |
| *Achnanthes lanceolata lanceolata* | genus |
| *Achnanthes lanceolata rostrata* | genus |
| *Achnanthes minutissima* | genus |
| *Achnanthidium minutissimum* | genus |
| *Agapetus fuscipes* | genus |
| *Alboglossiphonia heteroclita* | family |
| Amphipoda | exact |
| *Amphora aequalis* | genus |
| *Amphora inariensis* | genus |
| *Amphora ovalis* | genus |
| *Amphora pediculus* | genus |
| *Amphora veneta* | genus |
| *Ancylus fluviatilis* | family |
| *Asellus aquaticus* | family |
| *Athripsodes* sp. | family |
| *Baetis* sp. | genus |
| *Baetis rhodani* | genus |
| *Baetis scambus* | genus |
| *Baetis vernus* | genus |
| *Bezzia* sp. | family |
| *Caenis rivulorum* | genus |
| *Caenis robusta* | genus |
| *Centroptilum luteolum* | genus |
| Chironomidae | family |
| *Cloeon simile* | genus |
| *Cocconeis pediculus* | genus |
| *Cocconeis placentula* | genus |
| *Cocconeis pseudothumensis* | genus |
| Coleoptera | exact |
| *Cottus gobio* | genus |
| *Cyclotella* sp. | genus |
| *Cyclotella meneghiniana* | genus |
| *Cyclotella radiosa* | genus |
| *Cymatopleura elliptica* | genus |
| *Cymatopleura solea* | genus |
| *Cymbella* sp. | genus |
| *Cymbella cistula* | genus |
| *Cymbella proxima* | genus |
| Dystiscidae | family |
| *Dendrocoelum lacteum* | family |
| *Diatoma vulgaris* | genus |
| *Dicranota* sp. | genus |
| *Diploneis oblongella* | genus |
| *Diploneis parma* | genus |
| Diptera | exact |
| *Drusus annulatus* | genus |
| *Elmis aenea* | genus |
| *Eloeophila* sp. | family |
| *Encyonema silesiacum* | genus |
| Ephemeroptera | exact |
| *Erpobdella octoculata* | genus |
| *Fragilaria* sp. | genus |
| *Fragilaria bidens* | genus |
| *Fragilaria capucina* | genus |
| *Fragilaria capucina gracilis* | genus |
| *Fragilaria capucina radians* | genus |
| *Fragilaria capucina rumpens* | genus |
| *Fragilaria construens venter* | genus |
| *Fragilaria elliptica* | genus |
| *Fragilaria leptostauron* | genus |
| *Fragilaria nitzschioides* | genus |
| *Fragilaria ulna* | genus |
| *Fragilaria vaucheriae* | genus |
| *Fragilariforma virescens* | genus |
| *Gammarus pulex* | family |
| *Gasterosteus aculeatus* | genus |
| *Glossiphonia complanata* | family |
| *Gomphonema* sp. | genus |
| *Gomphonema angustum* | genus |
| *Gomphonema augur* | genus |
| *Gomphonema clavatum* | genus |
| *Gomphonema olivaceum* | genus |
| *Gomphonema parvulum* | genus |
| *Gyrosigma acuminata* | genus |
| *Gyrosigma attenuatum* | genus |

| | |
|---|---|
| *Hantzschia amphioxys* | genus |
| *Helobdella stagnalis* | family |
| *Hemerodromia* sp. | family |
| Hydracarina | family |
| Hydraenidae | genus |
| *Hydropsyche siltalai* | genus |
| *Hydroptila* sp. | genus |
| Hydroptilidae | family |
| *Hygrobia hermanni* | genus |
| *Ilybius* sp. | genus |
| *Lampetra planeri* | genus |
| *Lepidostoma hirtum* | genus |
| *Leuctra* sp. | genus |
| *Leuctra hippopus* | genus |
| *Leuctra inermis* | genus |
| Limnephilidae | family |
| *Limnephilus lunatus* | genus |
| *Limnius* sp. | genus |
| | |
| *Melosira varians* | genus |
| *Meridion circulare* | genus |
| *Navicula* sp. | genus |
| *Navicula atomus* | genus |
| *Navicula bacillum* | genus |
| *Navicula cincta* | genus |
| *Navicula cryptonella* | genus |
| *Navicula exilis* | genus |
| *Navicula ignota* | genus |
| *Navicula lanceolata* | genus |
| *Navicula margalithii* | genus |
| *Navicula minima* | genus |
| *Navicula seminulum* | genus |
| *Navicula slesvicensis* | genus |
| *Neidium dubium* | genus |
| *Niphargus aquilex* | family |
| *Nitzschia* sp. | genus |
| *Nitzschia amphibia* | genus |
| *Nitzschia capitellata* | genus |
| *Nitzschia dissipata* | genus |
| *Nitzschia fonticola* | genus |
| *Nitzschia frustulum* | genus |
| *Nitzschia heufleriana* | genus |
| *Nitzschia linearis* | genus |
| *Nitzschia palea* | genus |
| *Nitzschia recta* | genus |
| *Nitzschia sigmoidea* | genus |
| *Nitzschia sublinearis* | genus |
| *Oecetis* sp. | family |
| Oligochaeta | genus |
| *Oreodytes sanmarkii* | genus |
| *Oulimnius tuberculatus* | genus |
| *Oxycera* sp. | family |
| *Paraleptophlebia submarginata* | genus |
| *Phoxinus phoxinus* | genus |
| *Pinnularia* sp. | genus |
| *Piscicola geometra* | family |
| *Pisidium* sp. | genus |
| *Planaria torva* | family |
| *Planorbis* sp. | family |
| *Polycelis tenuis* | family |
| *Potamophylax latipennis* | genus |
| *Proasellus meridianus* | family |
| *Procloeon pennulatum* | family |
| *Psammodictyon constrictum* | genus |
| *Pseudostaurosira brevistriata* | genus |
| *Psychoda* sp. | family |
| *Pungitius pungitius* | genus |
| *Rhoicosphenia abbreviata* | genus |
| *Rhyacophila dorsalis* | genus |
| *Salmo trutta* | genus |
| Scirtidae | family |
| *Serratella ignita* | genus |
| *Silo nigricornis* | genus |
| *Simulium* sp. | genus |
| *Simulium vernum* | genus |
| *Stauroneis* sp. | genus |
| *Stauroneis smithii* | genus |
| *Staurosira construens* | genus |
| *Staurosira elliptica* | genus |
| *Staurosira pinnata* | genus |
| *Staurosirella lapponica* | genus |
| *Staurosirella leptostauron* | genus |
| *Staurosirella pinnata* | genus |
| *Surirella brebissonii* | genus |
| *Surirella capronii* | genus |
| *Synedra* sp. | genus |
| *Synedra parasitica* | genus |
| *Synedra ulna ulna* | genus |
| Tanypodinae | family |

| | |
|---|---|
| *Thymallus thymallus* | family |
| *Tipula* sp. | genus |
| Trichoptera | exact |
| Undifferentiated centric diatom | exact |
| CPOM | exact |
| FPOM | exact |