The University of Birmingham

School of Geography, Earth and Environmental Sciences

M.Sc. Dissertation

River Environments and their Management

Hugh Graham

Student Number: 1385432

An investigation into the impacts of a non-native Gammarid, *Dikerogammarus haemobaphes*, on the benthic macroinvertebrate community and ecosystem function of the River Cherwell.

September 2014

Acknowledgements

I would like to thank several people for the advice and support they have given me throughout the dissertation process.

Firstly, I would like to thank Tim Flood, from the Environment Agency in Wallingford, who gave me invaluable advice in the early stages of the project. In particular, I am hugely grateful for the lessons in the identification of *D. haemobaphes*.

Many thanks to Jon Sadler, my dissertation advisor, for his fast and helpful advice throughout my dissertation.

I am very grateful to Andy Moss for all his help during my laboratory work and with the identification of molluscs in particular.

Thanks must go to Melanie Bickerton for helping me with getting to grips with Caddisfly larvae identification and making sure I had all the necessary field equipment.

I am really grateful to all of the landowners who allowed me to access the river from their land. In particular I'd like to thank Mr Pearson, Mr Wadland and Mr Allen for the extra help and information they provided me with.

Many thanks to Mr Brian Clarke (Banbury and District Fishing Club), who helped me to get in touch with some of the landowners.

Finally I would like to thank all my friends and family who have helped me for the duration of the project. In particular I'd like to thank Hannah Willard, Siobhan Graham and Rod Graham for listening to my endless "shrimp ramblings" and helping during field work.

Abstract

Dikerogammarus haemobaphes is a species of amphipod shrimp, native to the Ponto-Caspian region. In 2012, it was discovered in the UK, in the River Severn. The species has spread across much of eastern and western Europe, establishing large self-sustaining populations in many freshwater rivers and lakes. The invasion of *D. haemobaphes* in Europe has, more often than not, been accompanied by other invasive amphipod species including the 'killer shrimp', *Dikerogammarus villosus*. As a result, the specific threats that *D. haemobaphes* present to British freshwaters is uncertain.

D. haemobaphes were first discovered in the River Cherwell, a tributary of the River Thames, in 2012. The invasive Gammaridae made their way into the Cherwell from the Oxford Canal and have since spread throughout the Thames catchment. The increasing gradient of *D. haemobaphes* density along the upper reaches of the Cherwell allowed for the study of their impacts in low and high abundances.

A survey of benthic macroinvertebrates revealed that the impacts of *D. haemobaphes* were significant at family, functional feeding group and community levels. Nine families were found to respond negatively to the presence of *D. haemobaphes* and only one (Ephemeridae) responded positively. Through the creation of a general linear mixed effects model, predator and shredder feeding groups were found to have a significant negative response. The community structure of the different sites was shown to be significantly different with the largest differences seen at sites where very large populations of *D. haemobaphes* were present.

An experimental in-situ enclosure study was carried out to investigate the impacts of *D. haemobaphes* on leaf litter decomposition. The study showed that with an increasing proportion of invasive Gammaridae, leaf decomposition decreased significantly with a large effect size.

The findings of both the invertebrate survey and experimental study suggest that *D. haemobaphes* pose a considerable threat to British fresh water environments. The invasion of *D. haemobaphes* is likely to: (i) cause significant alterations to invertebrate community structure through predation and competition for resources and (ii) have a significant effect on ecosystem function through the alteration of organic matter decomposition rates.

Contents

List of Figures

List of Tables

1.0. Introduction

Dikerogammarus haemobaphes is an invasive species of freshwater Gammaridae (Amphipoda) originating from the Ponto-Caspian region (Aldridge, 2013). It was first discovered in the United Kingdom in 2012, in the River Severn, and has since formed extensive self-sustaining populations in other freshwaters in the Midlands and South East of England. Following the initial discovery in the Severn, *D. haemobaphes* was discovered in numerous locations in the Trent, the Thames and their associated canals (*Figure 1)* (Environment Agency (EA), 2012). Since 2012, *D. haemobaphes* have expanded their range within the Thames catchment. Self-sustaining populations can now be found along the Thames' navigable extent, between Lechlade and London. Personal communications with Tim Flood (EA, Wallingford) suggested that the most likely origin of *D. haemobaphes* in the Thames is from the River Cherwell. Between Cropredy and Oxford, the Cherwell flows alongside the Oxford Canal which is connected to the Midlands Canal network. In numerous locations along the Cherwell, between Cropredy and Somerton, there are discharge channels from the Canal to the river and, near Adderbury (grid reference SP 49405 33784), the river mixes completely with the canal. The transfer of water between the canal and the river is the most likely cause of the transmission of the species.

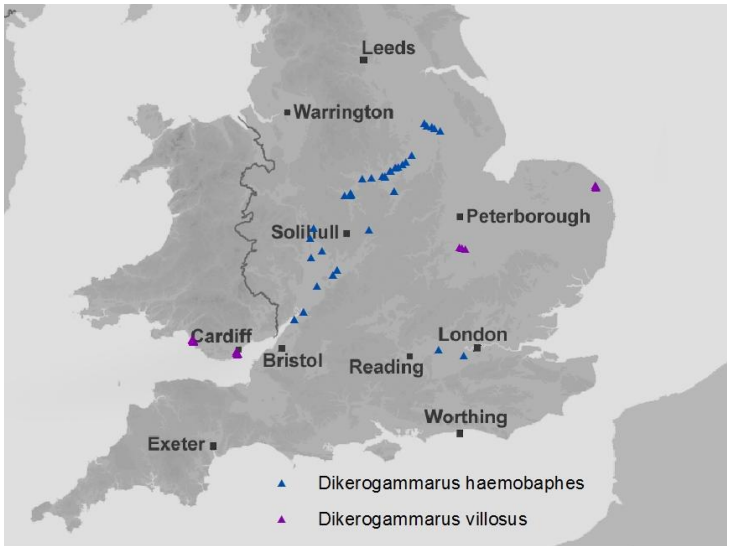

Figure 1: Distribution of D. haemobaphes and D. villosus across the UK (12/11/2012). From EA (2012)

D. haemobaphes have been found in rivers across eastern and western Europe, but their potential impact is not clear as their introduction commonly occurs alongside other Ponto-Caspian amphipods such as *D. villosus*, "the killer shrimp" (Aldridge, 2013). *D. haemobaphes* shares many life history traits with *D. villosus* (Grabowski, et al. 2007) and is therefore likely to exert similar impacts, which include predation on numerous benthic invertebrates and the possible replacement of native

amphipods such as *Gammarus pulex* (Kinzler, et al. 2009; MacNeil and Platvoet, 2005). Amphipods, and Gammaridae in particular, play a key role in ecosystem energy flow through the breakdown of detritus by shredding organic material. Introductions of *D. villosus* have been shown to affect the rates of leaf litter breakdown (MacNeil, et al. 2011), which could make significant alterations to ecosystem function. A recent laboratory study by Bovy, et al. (2014) compared the predation impact of *D. haemobaphes*, *D. villosus* and *G. pulex* on native midge larvae (*Chironomus sp.)* and the invasive amphipod, *Chelicorophium curvispinum*. The study indicated that D. *haemobaphes* exerted a stronger predation effect on *C. curvispinum* than *G. pulex,* but the predation effect on *Chironomus* was not significantly distinguishable.

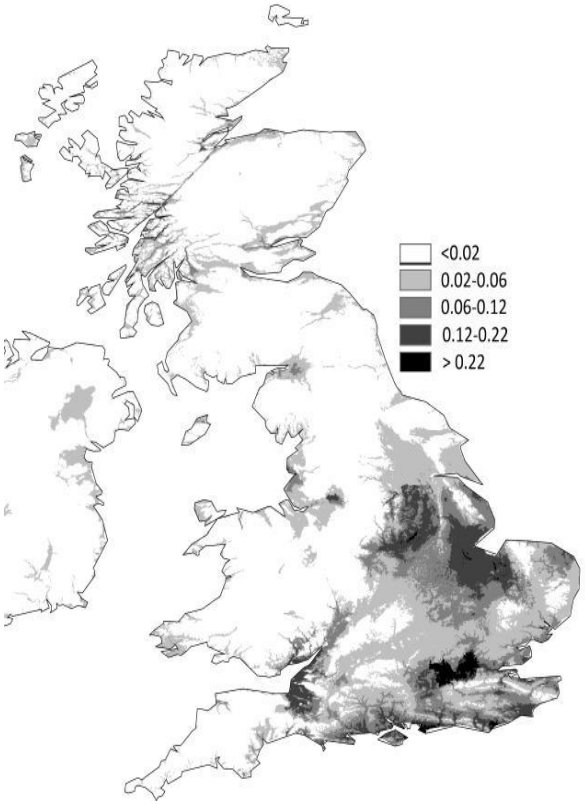

Figure 2: A habitat suitability map for D. haemobaphes, based on current European distributions. The model is based on climatic conditions, water chemistry and altitude. From Gallardo and Aldridge (2013)

There are only a limited number of studies that focus on the specific impacts of *D. haemobaphes,* none of which consider empirically based field measurements (Aldridge, 2013). The majority of research on invasive Ponto-Caspian amphipods has focused on *D. villosus* (Pockl, 2009, Dick et al. 2002). In the UK, the spread of *D.villosus* is considered highly probable (Gallardo, et al. 2012) and it is likely that *D. villosus* presents a more acute risk to ecological function and ecosystem structure than *D. haemobaphes* (Jazdzenski, et al. 2004, Bovy, et al. 2014). However, the distribution of *D. haemobaphes* across the UK is considerably greater and the water bodies which they inhabit are interconnected by canal networks. *Figure 2* shows the regions of the

UK which have suitable habitats to sustain *D. haemobaphes* based on current European distributions (Gallardo and Aldridge, 2013). *D. haemobaphes* therefore presents a potentially significant risk to freshwater ecosystems across the UK. This study aims to determine how this recent invasion may impact freshwater, and more specifically, riverine ecosystems in the UK.

1.1 The River Cherwell

The River Cherwell is a tributary of the Thames. It rises in Northamptonshire and flows largely through north Oxfordshire, where it eventually meets with the Thames in Oxford. The catchment size is approximately 943km^2 and the predominant land use is agricultural. The Cherwell is a lowland river with low relief (approximately 171 m). Contained within the catchment are two large towns, Banbury and Bicester. The river flows through the centre of Banbury. The Oxford canal is positioned alongside the river between Cropredy and Oxford and was constructed in 1778 (Neal, et al. 2006). *Figure 5* shows the catchment map and study area.

The catchment geology is dominated by clay, and therefore discharge is largely associated with direct runoff. This results in a relatively 'flashy' flood hydrograph response. Base flow is maintained during the summer months by aquifer sources, particularly in the upper reaches of the river, where this study is based. The base flow index of the river is approximately 0.4 between Cropredy and Somerton, rising to 0.65 further downstream in the catchment. The nature of the catchment geology causes the river to be susceptible to flooding which increases the chance of water transmission between the river and the canal. There have historically been significant pollution issues within the Cherwell, particularly downstream of Banbury (Neal, et al. 2006), but this has been improving in recent years and the chemical quality of the river is considered by the EA to be "good". The EA class the biological quality of the river as "moderate" upstream of Cropredy, "poor" between cropredy and Nellbridge and "moderate" between Nellbridge and the confluence with the River Ray at Islip, near Oxford (EA, 2014).

1.2 Aims and Objectives

- To determine the impact of *D. haemobaphes* on the benthic invertebrate community of the River Cherwell.
- To investigate the effect that *D. haemobaphes* has on ecosystem function and energy flow through the alteration of leaf litter decomposition dynamics.

2.0 Literature Review

This section summarises some of the literature relevant to this study. It will include: (i) a review of the invasion ecology theory which underpins the study and (ii) information on the life history traits and potential impacts of *Dikerogammarus*.

2.1 Invasion Ecology

There are a broad range of definitions with multiple meanings that can be used to describe the introduction of non-indigenous species into novel systems (Lockwood, et al. 2007). For the purpose of this dissertation, the term invasive species will be used to describe those fauna which, facilitated by human activity, have moved from their native range into a system which had not experienced that particular species up until the time of invasion.

The study of invasive species has developed significantly over recent years for a number of reasons. Firstly, the impacts of invasive species have increased significantly over recent years, much of them now seeming unavoidable (Lockwood, et al. 2007). The numbers of species being transferred out of their native environment and subsequently establishing successful populations elsewhere has also been continually increasing since the mid-1800s. Recently, it would appear that there have been an increase in the number of problematic invasions. It is now almost impossible to conduct any type of ecological study without encountering some form of invasive species, be that one who has had a significant impact on an ecosystem or not. Increasing understanding of how climate change is likely to affect ecosystems across the globe (Stocker, et al. 2013) suggests that the rate and success of future biological invasions will be altered (Lockwood, et al. 2007).

The reason for studying invasive species is twofold. Firstly, many invasive species have undeniably detrimental impacts on native ecosystems, which in turn can also affect the economic value of that ecosystem (Pimentel, et al. 2000). The specific impact of invaders on native fauna and flora is highly variable. They can for example, compete with, prey on and hybridize with natives. This is almost always to the detriment of the natives (Lockwood, et al. 2007). Invasive species may directly impact humans by: blocking waterways and navigation routes, causing the death of livestock and fish (commonly due to disease) and direct damage to man-made structures and homes (Mooney, et al. 2005). Secondly, invasive species give scientists a unique

insight into the ecology of both the origin and destination ecosystems, as well as providing information about the evolutionary history of these animals (Sax, et al. 2005).

2.2 Transport Vectors

A vector, in invasive ecology, can be considered as the means by which a species is transported along a pathway. In an ever developing and globalizing world, increasing transport within and between continents has become more frequent and faster. Human-mediated transport vectors can cause the inadvertent release of non-native species and consequently, humans have had an undeniable effect on the rate of invasions. Although biological invasions do take place naturally, the frequency of establishment is very small in comparison. For example, the National Research Council (2002) showed that, prior to human colonisation, successful natural establishment of plant species in the Hawaiian Islands took place once every 100,000 years. Following the arrival of Polynesian people, this rate rose to one new species every 50 years and with the settlement of Europeans, this figure increased further to one every 22 years. Many human-mediated invasions have been caused deliberately, either for ornamental or game purposes (Lockwood, et al. 2007), but since the introduction of *D. haemobaphes* into the UK was almost certainly an unintentional act, this section will focus on the unintentional transfers.

Vectors of particular importance to this study are those that affect aquatic life. Perhaps the most well documented and largest of these vectors is ship ballast. Ship ballast is required to maintain ship stability and handling. Traditionally, ballast consisted of rubble or any other easily sourced heavy material (Mack, 2004). This was responsible for the transportation of numerous species both faunal and floral, aquatic and terrestrial. In more modern ships, ballast usually consists of water. This water is taken on-board at the port of origin and discharged at the destination following unloading (Lockwood, et al. 2007). Carlton (1999), suggested that approximately 10,000 different species may be transported every day via ballast water. Although many of these species will not survive the journey nor the new environment, some species will survive and develop self-sustaining populations in the destination environment. Today, it is common practice to carry out a ballast water exchange mid-way through any ocean crossing. The intention of this is to reduce the number of animals released into the destination port. This has been shown to significantly reduce the risk of successful invasion (Wonham, et al. 2001), but it does not rule out the chance of species establishment.

At a smaller spatial scale, it is possible for some aquatic fauna to attach themselves to boats or become trapped in recreational equipment such as fishing waders and nets. Therefore, the movement of these items can also facilitate the range expansion of a species (Aldridge, 2013).

2.3 Propagule Pressure

The term “propagule” is used to describe the community of an invasive species (Lockwood, et al. 2007). A larger propagule “size” contains a greater number of individuals, whereas the propagule “number” describes the number of release events of that species (Lockwood, et al. 2009). The “condition” of the propagule is an important consideration, as a healthier propagule will have more chance of success in its new environment. Together, the propagule size, number and condition can be described as the “propagule pressure” (Lockwood, et al. 2005). This propagule pressure is strongly related to the success of the invasion, although any establishment is also dependent on the presence of suitable conditions at the location of release (Lockwood, et al. 2007).

2.4 The Role of Disturbance

Disturbance undoubtedly plays an important role in ecological invasions. However, the circumstances of the invading species and the receiving ecosystem play an equally important role in controlling the outcome of an invasion (Lockwood, et al. 2007). There are countless examples of invasions that appear to have been facilitated by disturbance (e.g. Minchinton, 2002; Brooks, et al. 2004; Lozon and MacIsaac, 1997). Many of these examples attribute successful invasions to the opening of invasion windows, or gaps in the ecosystem. There are opposing examples where disturbance has prevented invasions. For example, Smith and Knapp (1999) found that the invasibility of C4 grassland was directly related to community structure, irrespective of disturbance. It should be emphasised that any weakening of the ecosystem due to discrete disturbance events may actually benefit native species equally. It should also be noted that there is likely to be a significant bias in the literature because most studies of invasive species focus on those that have successfully established (Lockwood, 2007).

It is important to consider that the impact of natural disturbances, such as fire (Brooks, et al. 2004), will differ to anthropogenic disturbances such as river flow regime

alteration (Schreiber, et al. 2003). In particular, where human-facilitated invasions are concerned, species are often already adapted to human disturbance (Lockwood, et al 2007) and therefore may respond more positively to anthropogenic disturbance than natives.

Invasive species may even be the cause (driver) of a disturbance, although this is often hard to determine without detailed information about the effected system pre and post invasion. MacDougall and Turkington (2005) tested this "driver" or "passenger" theory by conducting an experiment on fragmented and fire-supressed savannah where two non-native grasses had been introduced. The overriding model in this example showed that invasive species were passengers of disturbance. However, this does not rule out the possibility of invasive species causing disturbances elsewhere. In some cases, invasive-driven disturbance can lead to a positive feedback system where the presence of invasive-disturbance allows for the establishment of new invasive species. Simberloff and Von Holle (1999) refer to this as an invasive meltdown. This concept has been identified as a potential driver of *D. haemobaphes* success by Bovy, et al. (2014), a theory which will be discussed in more detail later in this dissertation.

2.5 Dikerogammarus – Life History Traits

There are three known species of *Dikerogammarus: D. villosus*, *D. haemobaphes* and *D. bispinosus*, all of which are native to the Ponto-Caspian region (Muller and Schramm, 2001, Muller, et al. 2002). All three species have had varying degrees of success as invaders across eastern and western Europe, although *D. bispinosus* has not yet been discovered in the UK (Van der Velde, 2000; Jazdzenski, et al. 2004; Grabowski, 2007; Aldridge, 2013). The *Dikerogammarus* genus share a number of highly favourable life history traits which have enabled their rapid range expansion and facilitated the deletion of many native amphipod species (particularly *Gammarus*) throughout Europe (Grabowski, 2007; Bacela, et al. 2009). These traits include:

(i) Large brood size, with recorded *D. haemobaphes* egg clutches ranging between 37 (Bacela, et al. 2009) and more than 100 (Kley and Maier, 2006).
(ii) Higher partial fecundity (brood size / female size) than native Gammaridae (Pockl, 2009, Grabowski, 2007).
(iii) A lower maturity index (minimum/mean breading size) than native gammarids (Bacela, et al. 2009).
(iv) Short embryonic development time and short maturation time (Pockl, 2009)

(v) A high number of generations per year – both *D. haemobaphes* and *D. villosus* produce three generations a year compared to only one for *G. pulex* (Grabowski, 2007).

(vi) Timing of larval development coincides with the time of year when food availability is at its maximum (Pockl, et al. 2009).

(vii) Shorter life span than most other European amphipods (Bacela, 2009).

(viii) Tolerance of a wide range of environmental conditions such as salinity, dissolved oxygen, human degradation and temperature; temperature tolerances of D. haemobaphes have been found to be as high as 30°C (Aldridge, 2013).

2.6 *Dikerogammarus* – Predation

Traditionally, Gammaridae were considered only as shredders. However, as knowledge developed it became clear that many species of *Gammarus* and *Dikerogammarus* were extremely effective predators (MacNeil, et al. 1997). All species of *Dikerogammarus* are known to be very strong intra-guild predators (IGP). In particular, *D. villosus* and *D. haemobaphes* have been found to exert a similar IGP pressure on each other and on other amphipods such as the native *G. pulex* (Kinzler, et al. 2009). Further lab based studies have revealed that *D. villosus* preys on taxa from a wide range of trophic and habitat niches to an extent that far exceeds that of native Gammaridae (Dick, et al. 2002; Dodd, et al. 2014). Equally, there is evidence of cannibalism within many amphipod species, particularly *D. haemobaphes* (Kinzler, et al. 2009). This cannibalistic behaviour, along with slightly less effective predation on some species such as *Corophium curvispinum* (Bovy, et al. 2014), may explain why *D. haemobaphes* populations have been replaced by *D. villosus* in numerous locations across Europe (Kinzler, et al. 2009; Jazdzenski, et al. 2004). These behaviours have shown that *D. villosus* have a strongly negative impact on a variety of different benthic invertebrate taxa. However, there is limited evidence to show with any certainty the specific impacts of *D. haemobaphes* on freshwater systems (Aldridge, 2013).

2.7 Dikerogammarus – Leaf litter Shredding

Shredding invertebrates are responsible for converting coarse particulate organic matter (CPOM) into fine particulate organic matter (FPOM) (Wallace and Webster, 1996). In many streams the dominant input of CPOM consists of leaf litter from the riparian zone (Hladyz, et al. 2010). This leaf processing by shredders allows other invertebrates to filter this FPOM from the water column or gather it from sediments (Graca, 2001). An increase in FPOM will also lead to an increase in the surface area of

organic material, facilitating more microbial decomposition. Therefore, any changes to the rate of leaf litter decomposition could have major direct and indirect impacts at multiple trophic levels within an ecosystem (Graca and Conhoto, 2006).

Gammarid species, although shown to feed at multiple trophic levels (MacNeil, et al. 1997), are one of the most dominant and effective shredders in British rivers and therefore exert strong controls on ecosystem energy flow (Navel, et al. 2010). With the invasion of non-native Gammaridae threatening to replace native *G. pulex* in UK freshwater systems, it is possible that there might be significant alterations to leaf litter processing which could fundamentally change ecosystem function and structure (MacNeil, et al. 2011).

Two laboratory based studies compared the leaf litter shredding capabilities of a selection of native and non-invasive Gammaridae with *D. villosus* under un-stressed environmental conditions (Piscart, et al. 2011; MacNeil et al. 2011). These studies showed that the leaf litter processing ability of *D. villosus* was far inferior to the native Gammarus species. In contrast, Truhlar, et al. (2014) conducted a laboratory study comparing leaf shredding rates of *D. villosus* and *G. pulex* under elevated temperatures and conductivities, and discovered that the rate of leaf shredding by *D. villosus* at 25°C was considerably higher than *G. pulex*. They reported that rates at lower temperatures were not significantly different and that conductivity differences had no species-specific effect. This study suggests that, under stressed conditions (particularly high temperatures) invasive *D. villosus* are more effective at shredding leaf litter. However, leaf breakdown rates are highly dependent on the type of leaf and the time that they have been submerged. Many shredding species, including *G. pulex*, prefer a degree of microbial decomposition, or conditioning, before consuming leaf litter (Graca, 2001). To my knowledge, there is currently no published literature observing the leaf litter shredding behaviour of *D. haemobaphes*.

2.8 D. haemobaphes – Range Expansion

The Ponto-Caspian region is something of an “invasive species hotspot”, particularly with regards to aquatic invertebrates (Von Vaupel Klein, 2000). There have been numerous invasions from the Ponto-Caspian range into Europe, including 19 different species of Crustacea arriving since 1800 (Von Vaupel Klein, 2000; Bij de Vaate, et al. 2002). *Figure 3* shows a map depicting the transport pathways of *D. villosus and D. haemobaphes* from their Ponto-Caspian range to Western Europe. D. The most likely

form of transport was via ‘hitchhiking’ on boats along canal networks across Europe (Jazdzenski, et al. 2004).

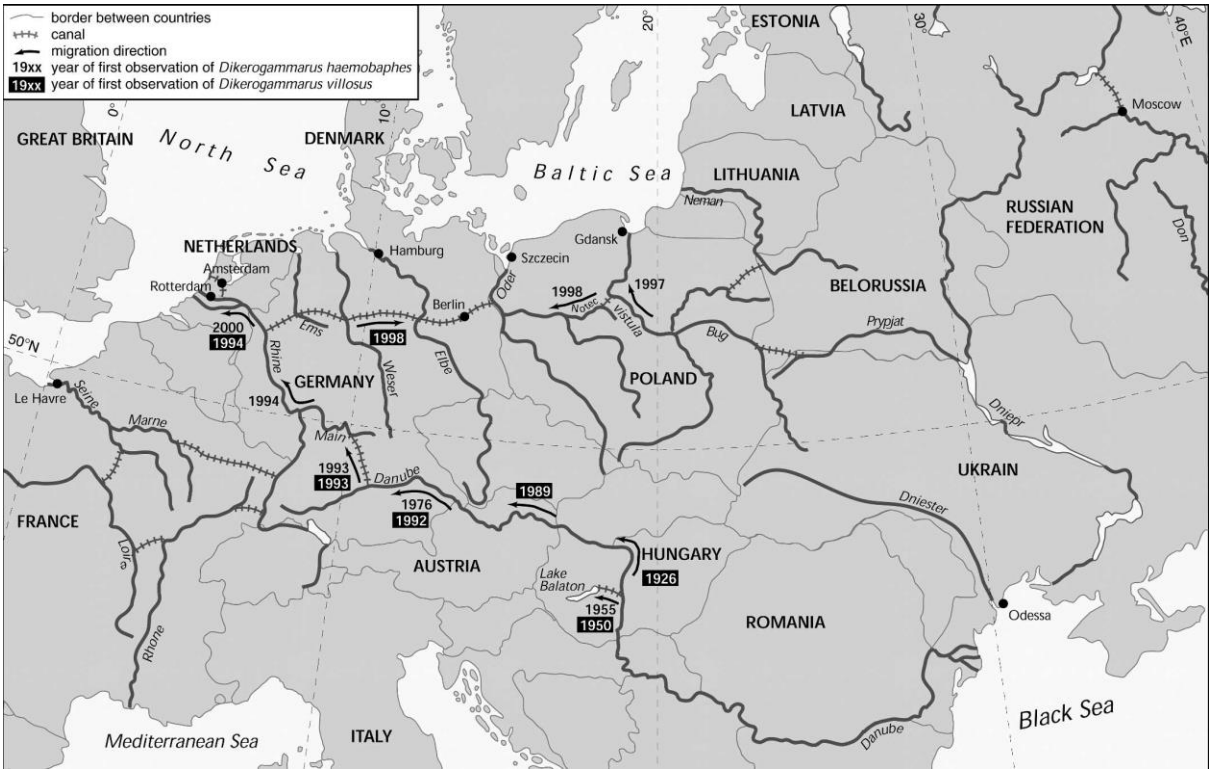

Figure 3: Main Dispersal routes of D. haemobaphes and *D. Villosus* from their Ponto-Caspian range to Western Europe. From Bij de Vaate, et al. (2002).

D. haemobaphes was first found outside its natural range in Hungary in 1955. Travelling up the Danube, the species was first found in the upper reaches of the Danube in 1976. From here it travelled to the Rhine via canal networks that make up the “Southern corridor” (*Figure 4*). *D. haemobaphes* was recorded in the main Danube canal in 1993 and eventually reached the North Sea basin via the Rhine. Movement through the “Central corridor” (*Figure 4*) also took place with *D. haemobaphes* first being discovered in Poland, in the Vistula River in 1997. By 1998, *D. haemobaphes* was the dominant amphipod species in many parts of the Vistula and its tributaries (Bij de Vaate, et al. 2002).

Following arrival in the Netherlands, it is likely that large populations became established in freshwater ports, such as Rotterdam, where salinity levels are lower than the species’ maximum salinity tolerance of 8‰ (Bin de Vaate, et al. 2002; Pockl, et al. 2009; MacNeil and Platvoet, 2005). From these large ports, the transport of ballast water from the Netherlands to the UK is the most likely means of transmission. 7.6% of

traded goods between the EU and the UK is transported via Rotterdam (Aldridge, 2013). Consequently, *D. haemobaphes* was first discovered in the UK in 2012 (EA, 2012) and now occupies a relatively large spatial area, as shown in *Figure 1*.

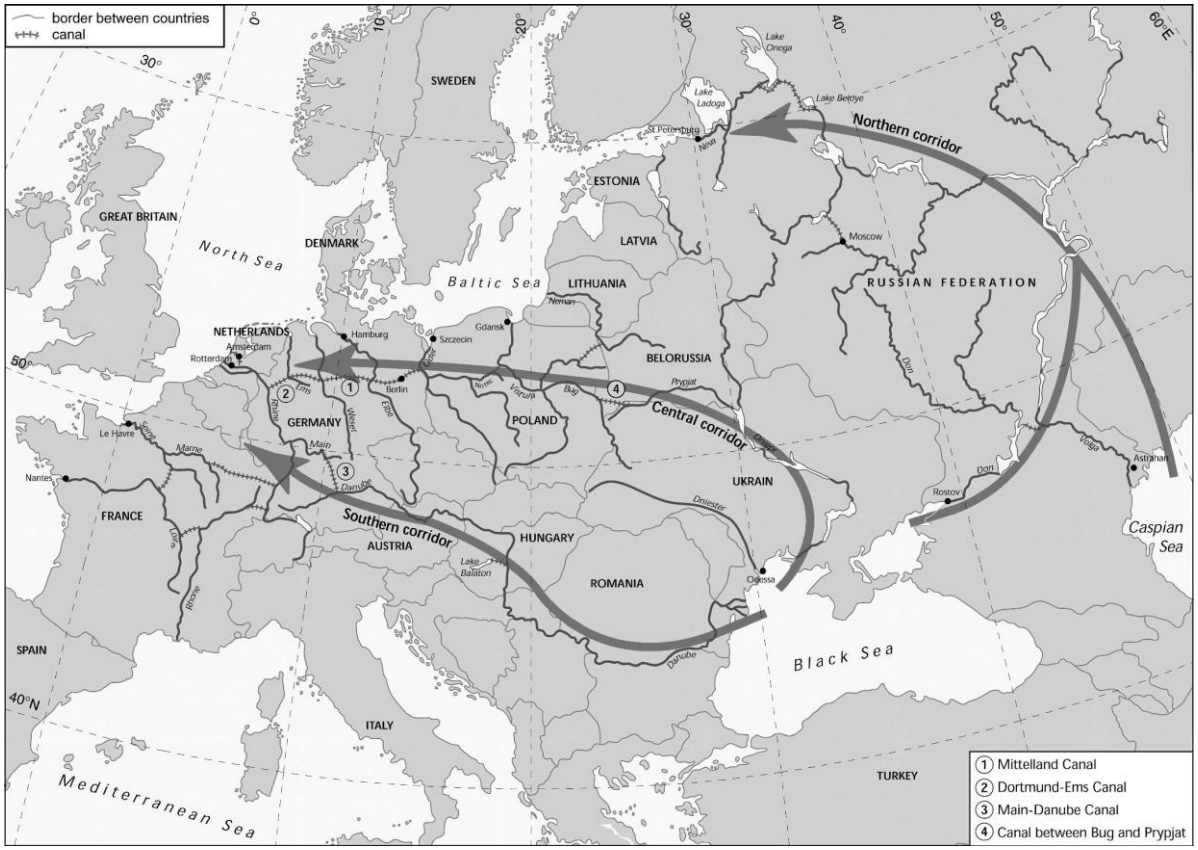

Figure 4: The main transport pathways of Ponto Caspian aquatic invaders. Central and southern corridors were utilized by D. haemobaphes. From Bij de Vaate, et al. (2002).

3.0 Methods

This section will provide details of the different field, laboratory, experimental and statistical techniques used during the study. Before undertaking any of them, a preliminary pilot survey was undertaken in order to refine the techniques used in the main study. The details of this can be found in *Appendix 1*.

3.1 Macroinvertebrate Survey Design

Seven reaches of the River Cherwell, of varying size, were selected between Cropredy and Somerton (Oxfordshire), as shown in *Figure 5*. These sites were selected based on their numbers of *D. haemobaphes*, ease of access and physical condition. Within each site, samples were taken from suitable locations based on the physical habitat of the river. At some sites, there were fewer suitable survey locations and therefore a lower number of samples were taken from them. The sample number per site ranged between 2 (*Site F*) and 10 (*Site B*). The selection of sites and sampling within, was designed in such a way to measure the impact of an increasing abundance of *D. haemobaphes*.

In order to allow comparison between samples, all survey locations were, as far as possible, located within similar physical environments. All surber samples were collected on riffle sections with mobile substrate (approximately 0.5 – 10cm diameter). Locations with abundant macrophyte growth were avoided where possible.

Macroinvertebrate samples were recorded using a surber sampler, with a quadrat size of 30x30 cm and a mesh size of 500 µm. The substrate within the quadrat area was disturbed for 30 seconds by hand. Large cobbles present in the quadrat were lifted and disturbed to ensure the entire area had been agitated. After sampling, the contents of the net were placed into bags and preserved with 70% Industrial Methylated Spirit (IMS).

Samples were returned to the laboratory in Birmingham University, where they were cleaned with water in a 500 µm sieve and all invertebrates were counted and identified to family level using a low powered microscope. The “Guide to British Freshwater Macroinvertebrates for Biotic Assessment” (FBA, 2011) was used to identify invertebrates. Gammarid species were identified to species level.

3.2 Site descriptions and locations

Figure 5 shows the Cherwell catchment and the site locations where surveys were conducted.

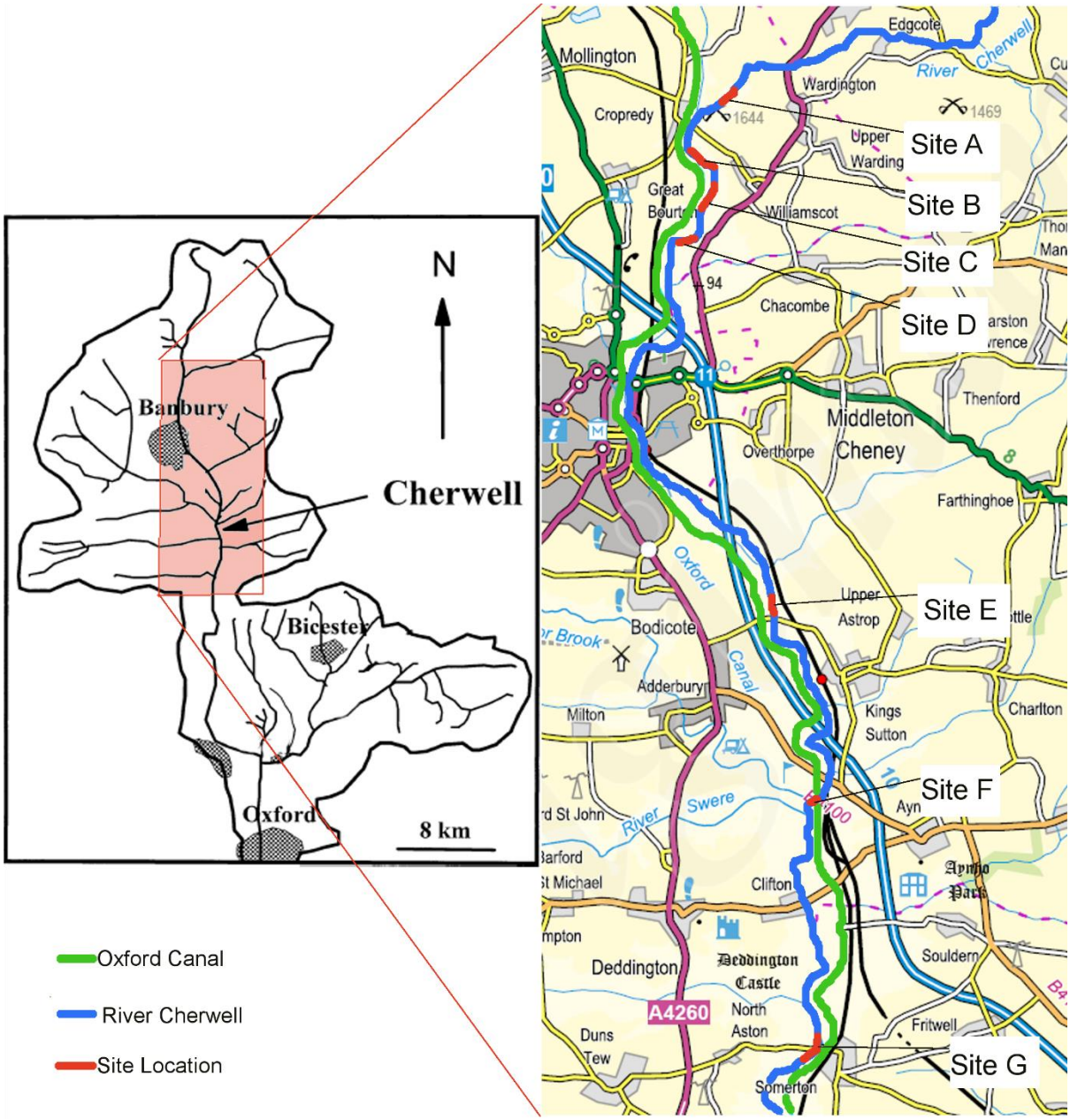

Figure 5: River Cherwell catchment overview and the section studied for this dissertation. Ordinance survey map indicates site locations, and the position of the river and the Oxford Canal. Catchment overview modified from Neal, et al. (2006), Ordinance Survey map taken from Digimap (2014).

Site A: Cropredy Manor, SP 47922 47324

Upstream from Cropredy Manor, a small footbridge crosses the river. At this point there is a large gravel riffle that extends 10m upstream and 5m downstream of the bridge.

The dominant substrate type is mobile gravel (approx. 2cm in diameter) with some sections having larger, less mobile cobbles. There are a further two short riffles downstream, formed by the root balls of the aligning trees. Upstream of the bridge there is extensive macrophyte growth, however, patches of bare gravel are present. Compared to the downstream sites, riparian cover and tree growth is greater at this location. In total, six samples were taken from this reach. Photos in *Appendix 2*.

Site B: Cropredy Farm, SP 46923 46185

The reach in question is in the field adjacent to the Cropredy playing fields. At the start of the reach (given in the grid reference), the river splits in two with a deep canalised section, with a small portion of the flow, diverted to a mill. The rest of the flow passes over a large weir (approx. 3m). There are 5 artificially created riffles on this section, each of which was surveyed twice at a different point on the riffle. The gravel substrate of the riffles has a diameter of approximately 1-2cm. There is limited macrophyte growth within the riffles. Both banks have riparian vegetation consisting of arable grassland, nettles and several species of large trees. The field on the east side of the river is used for cattle grazing, but there is no access to the river for the cattle. Photos in *Appendix 3*.

Site C: Williamscot Farm, SP 47530 45474

The floodplain to the east and west of the river is utilised for cattle grazing. A riparian buffer (approx. 2m) is present on both banks. Four artificial riffles are present in this reach – a single sample was taken from each riffle. Instream habitats are comparable to *Site B*. A fenced drinking access point is located towards the downstream end of the reach. Photos in *Appendix 4*.

Site D: Red Lunch Barn field, SP 47145 44356

Agricultural and environmental conditions at this site are very similar to site 3. There are two artificial riffles present at this site, two samples were taken from the upstream riffle and four from the larger downstream riffle. There is a cattle drinking access point downstream of both the riffle sections. Photos in *Appendix 5*.

Site E: Cherwell Valley Silos, Kings Sutton, SP 48670 37351

An industrial plant is located on the east bank whilst agricultural fields align the west bank. The banks of the river are reinforced with large immobile boulders and concrete. Despite this, extensive reed and macrophyte growth is present on the margins of the river. The substrate is largely immobile with finer (0.5cm diameter) gravels present in

some sections. The river is channelized along this reach with relatively high flow velocities. Four samples were taken from this site, approximately 5m apart from each other. Photos in *Appendix 6*.

Site F: Nellbridge Farm, SP 49377 33787

At Nellbridge Farm, the River Cherwell mixes with the Oxford Canal before flowing over a weir (approx. 2m high). Immediately downstream of the weir is a fast flowing, turbulent rapid/riffle section. Two samples were taken from gravelly (1cm diameter) patches in this location. No suitable sites were located downstream. The site of sampling was heavily shaded by riparian willow tree growth. *Figure 6* shows the point at which the canal and river mix. Further photos in *Appendix 7*.

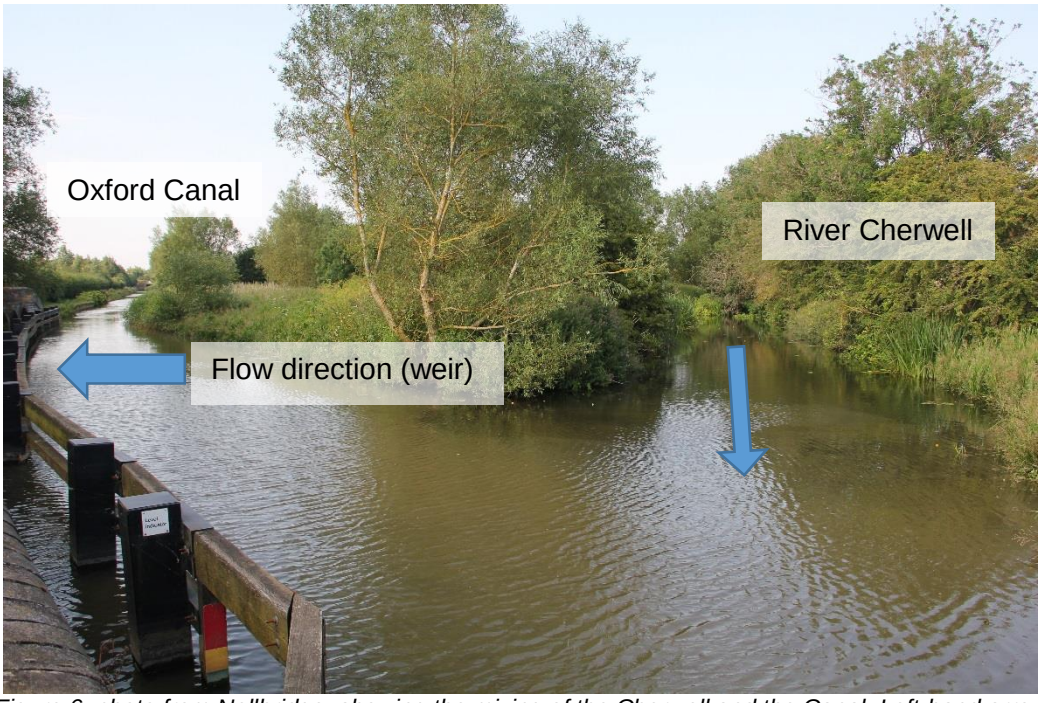

Figure 6: photo from Nellbridge, showing the mixing of the Cherwell and the Canal. Left-hand arrow indicates where the flow discharges over a weir.

Site G: Somerton bridge, SP 49529 29076

Upstream of the Somerton road bridge is a large riffle section extending approximately 20m. This riffle section consists of two channels separated by a sequence of mid channel bars which, apart from the most upstream of the bars, are vegetated by tall reeds. The right hand bank (looking downstream) is lined by a wall which supports the road. The left bank is unfenced and used for cattle grazing. There is limited riparian vegetation other than some reeds and bushes aligning the left bank. The substrate consists mostly of mobile gravels (approx. 2-3cm diameter) with numerous small

macrophytes present across the riffle. Four samples were taken from this riffle, upstream of the bridge. This locality was also used for the experimental study.

Immediately downstream of the bridge there is a small broken weir which has created an artificial riffle. Several large trees on the right bank shade the river from the west. The substrate is larger here with a mixture of medium sized (approx. 3-4cm) gravels and larger angular cobbles (approx. 20cm diameter) from the dismantled weir. No macrophyte growth was present. Two samples were taken from this section.

25m downstream is a natural riffle, similar in character to the riffle upstream of the bridge, although much shorter (5m in length). Two more samples were taken at this location. Photos in *Appendix 8.*

3.3 Leaf Litter Shredding Experiment Design

Experimental encolsures were constructed from white, square guttering with dimensions of 20x6.5x6.5cm. The base area of the enclosure was therefore 130cm^2. 3g of dried Alder leaves were weighed and placed inside each of the enclosures. The ends of the enclosures were sealed with 500 µm mesh to prevent the loss of leaf material, the loss of Gammaridae and the entrance of other invertebrates.

G. pulex were collected from *Site A* (Cropredy Manor) and *D. haemobaphes* were collected from *Site G* (Somerton Bridge) by kick sampling with a pond net. The contents of the net were placed in a white tray and the required species was removed with a spoon to prevent injury. Similarly sized Gammaridae were selected by eye and those who appeared injured or were an unsuitable size were returned to the river. Based on the pilot study (*Appendix 2*) a mean density of *D. haemobaphes* at *Site G* was calculated as approximately 30 per 130cm^2.Therefore a total number of 30 Gammaridae were placed in five of the six treatments at different ratios. Each of the treatments was replicated four times. The treatments were as follows - ratios are given as number of *D. haemobaphes:G. pulex*: (i) 0:0 (control treatment) (ii) 0:30 (iii) 7:23 (iv) 15:15 (v) 23:7 (vi) 30:0.

The enclosures were placed in the stream at *Site G*, upstream of the bridge. Their position was slightly sheltered from the main flow by a small macrophyte. The water depth at the start of the experiment was approximately 20cm deep. The enclosures were attached to bricks in pairs and tied to a tree to prevent being moved downstream by the flow. The enclosures were positioned at right angles to the flow in order to limit mechanical degradation of the leaf litter by flow.

The enclosures were removed from the river after 22 days. This time would allow for microbial conditioning to occur (Graca, 2001) and therefore both species of gammarid to feed at their optimum rate. The enclosures were returned to the laboratory where all leaf material was collected, dried overnight at 60°C and weighed. The leaf material from each treatment was then placed in a ceramic crucible and placed in a furnace at 500°C where it was kept for 2 hours. The remaining Ash Free Dry Mass was measured and taken away from the post experiment dry mass to calculate the overall consumption.

3.4 Statistical Methods

All statistical analysis was carried out using R studio, version 3.1.1 (R Development Core Team, 2014). A briefer description of the purpose of each of these analyses is given in the results section.

3.4.1 Survey Analyses

Family Level Response

The responses of different families to the increasing presence of *D. haemobaphes* were analysed using one way logistic regression models. Initially, both linear modelling and (quasi)poisson general linear modelling (GLM) were attempted. However, due to the heterogeneity of the data and the absence of many families in different samples (presence of zero values), the model fits were heteroscedastic and could not be validated.

Logistic regression was used in two scenarios: (i) to determine the probability of presence or absence of different families under an increasing number of *D. haemobaphes* and (ii) for families, such as *Gammarus*, that were present in every sample, the above method would not be suitable. Therefore, the presence or absence of *D. haemobaphes* was modelled against the abundance the family to determine if the presence of *D. haemobaphes* affected the abundances of the relevant family.

Using logistic regression removed errors relating to over or under dispersion as binomial GLMs cannot be over/under dispersed (Crawley, 2005). Before accepting any model, cook's distance values were calculated to measure the impact of influential points. Those with values greater than 1.0 were removed and the model was run again. This was only necessary for Baetidae and Ecnomidae models.

The effect size of the model was then calculated using the method outlined by Chinn (2000). This involved calculating the Odds Raios (OR) and dividing by 1.81. Calculation of Effect size is important because Null Hypothesis testing Statistics (NHTS) simply test whether or not to accept or reject the null hypothesis (Nakagawa and Cuthill, 2007). It is therefore important to calculate an effect size that can be compared between studies.

In light of the study by Bovy, et al. 2014, further investigation was required for the interaction between *D. haemobaphes* and Corphiidae, but because of the absence of Corophiidae in all but the last two sites, a logistic regression model was not possible. Therefore, a simple comparison of mean abundance per site was created in the form of a box plot. The R-script for this section of the analysis can be found in *Appendix 9*.

Comparison of communities

Sorensen similarity coefficients were used to investigate the 'distances' (relative similarity) between the different sites A-G and sites where *D. haemobaphes* was present and absent. Using the Vegan package (Oksanon, et al. 2013) in R, a Binary Bray-Curtis index (equivalent to Sorensen) was applied to the survey data. The calculated distances are then subset into the required group (either presence/absence or site) and the model is created.

Analysis of Variance (ANOVA) tests were run for the models in order to determine if there was a significant difference between groups. The distances were then plotted in an ordination style graph to show the relative differences between the groups.

The model was run a third time using abundance as a means of grouping the data. The models can only be plotted in their ordination form where the group is a factor. Therefore, for this model, the distances were extracted and used to create a logistic regression model. The logistic regression model was required as the Sorensen model produces proportional data – binary models are capable of processing this type of data (Crawley, 2000). The R-script for this section of the analysis can be found in *Appendix 10*.

Analysis of Impacts on Functional Feeding Groups

In order to identify which areas of the food web *D. haemobaphes* could be affecting, a general linear mixed effects model (GLMM) was developed to investigate the impact on different functional feeding groups. Using a modelling procedure derived from Jamil, et al. (2013) (who investigated how environmental gradients could affect different species that were grouped by specified common traits) it was possible to identify which functional feeding groups responded to an increasing abundance of *D. haemobaphes*. The GLMM method was chosen because it can be used where pseudo-replication and heteroscedatic variance prevent the use of other techniques (Jamil, et al. 2013, Zuur, et al. 2009). This is achieved by adding family and site to the model as random factors, thus randomising both the intercept and slope.

A pseudo R^2 was calculated (Byrnes, 2008) which, although not providing an effect size which is comparable between studies, is useful nevertheless in estimating the effect size of the model and in determining how much of the variance can be explained by the model. The R-script for this section of the analysis can be found in *Appendix 11*.

3.4.2 Experimental Analysis

The overall leaf litter consumption was analysed using a one-way factorial ANOVA, where each treatment represented a different factor. Homogeneity of variance was confirmed using ‘fitted vs. residuals’ plots and normality was tested using ‘QQplots’ (*Appendix 12*). An eta-squared value was calculated to show the effect size of the response (Blumstein, 2006, Nakagawa and Cuthill, 2007). The R-script for this section of the analysis can be found in *Appendix 13*.

4.0 Results

In this section, the key findings from the field survey and the experimental study will be displayed and interpreted.

4.1 Field Survey – Family Level Response

A total of 40 different families were recorded from all seven sites and two species of Gammaridae were recorded. *D. haemobaphes* and *G. pulex,* were considered separately in the analysis. Chironomidae and Simuliidae were identified to family level, however other members of the Diptera Order were not identified beyond order level due to time constraints. For full results see *Appendix 14*.

Of the 40 families, the responses of the 20 most abundant were compared to either (i) the abundance of *D. haemobaphes* or (ii) the presence/absence of *D. haemobaphes* in a series of Logistic regression models. Of the 20 models, 10 families recorded a statistically significant ($p<0.1$) response. The results of these 10 models are displayed below.

4.1.1 Logistic Regression Models where Family data is binary and *D. haemobaphes* abundance is continuous.

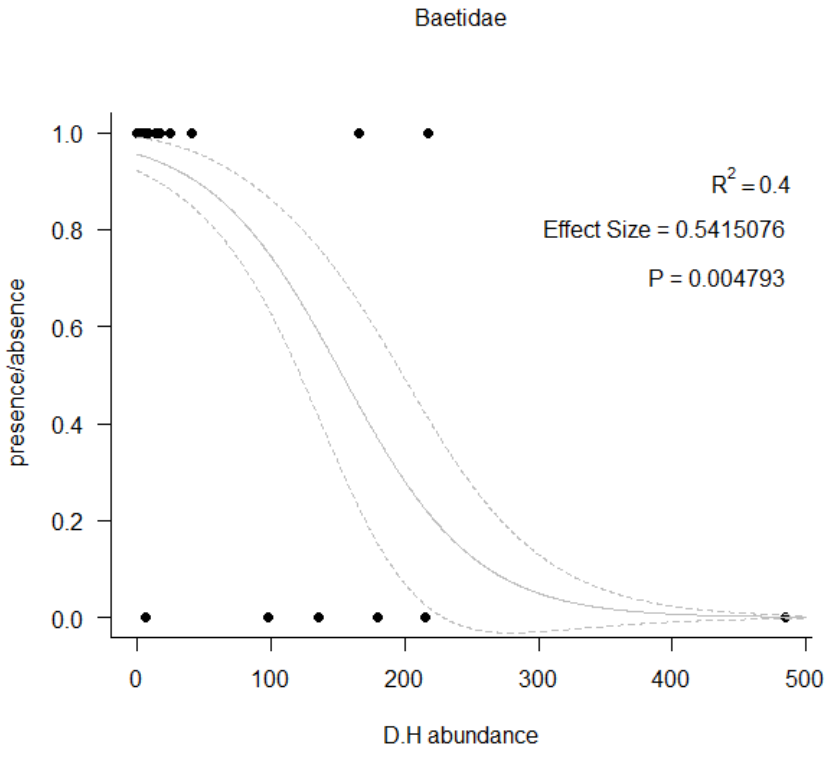

Figure 7: Logistic regression model showing the probability of Baetidae presence/absence under an increasing population of D. haemobaphes.

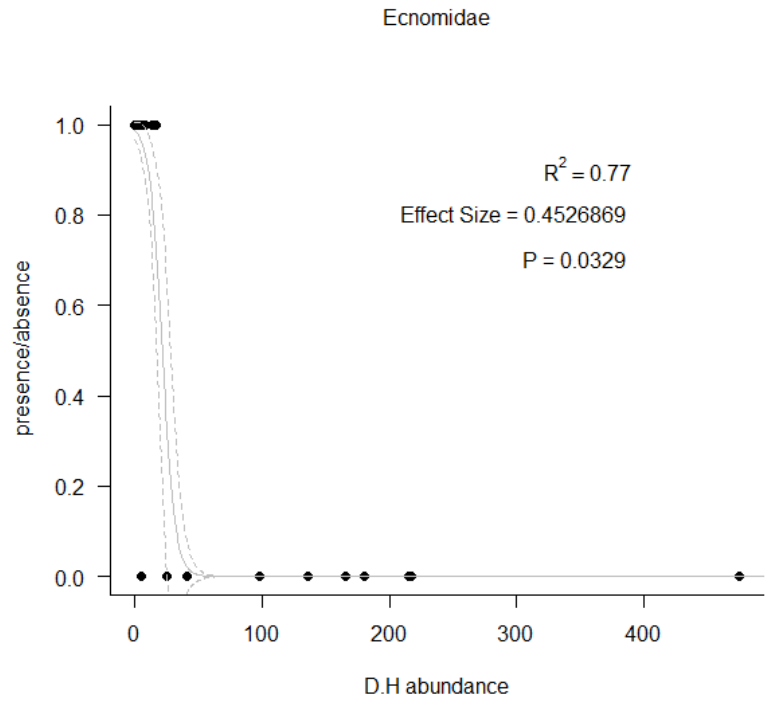

Figure 8: Logistic regression model showing the probability of Ecnomidae presence/absence under an increasing population of D. haemobaphes.

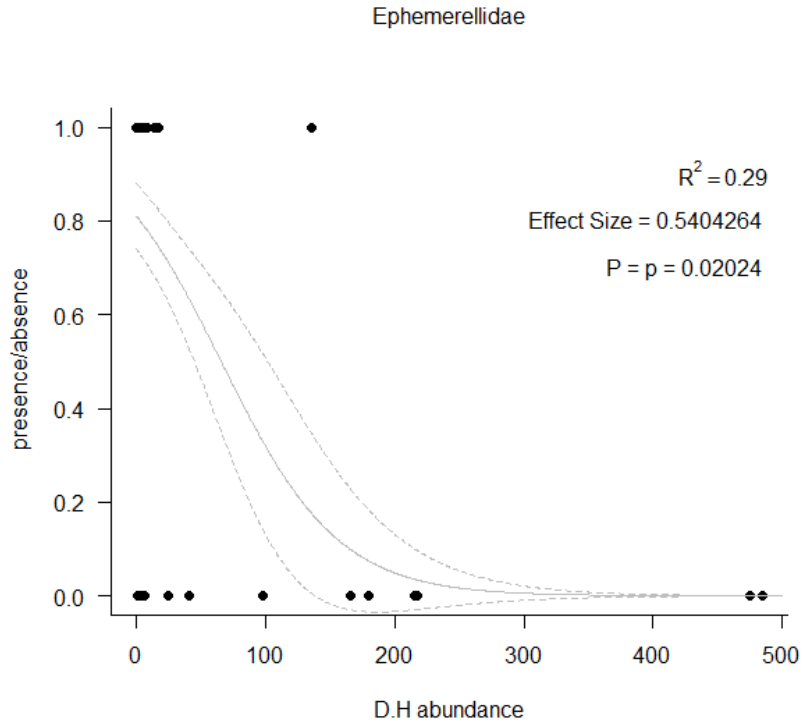

Figure 9: Logistic regression model showing the probability of Ephemerellidae presence/absence under an increasing population of D. haemobaphes.

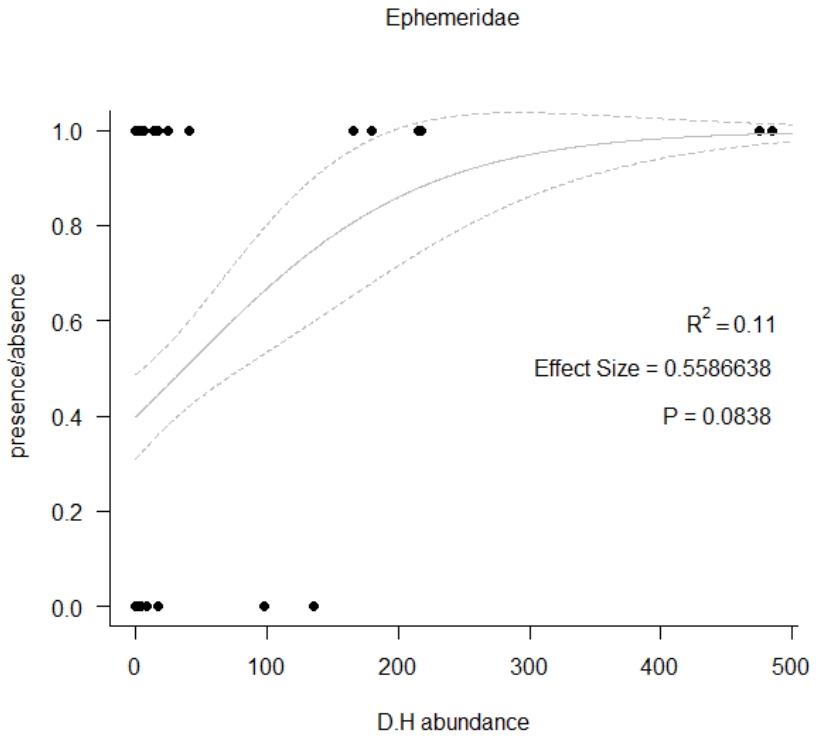

Figure 10: Logistic regression model showing the probability of Ephemeridae presence/absence under an increasing population of D. haemobaphes.

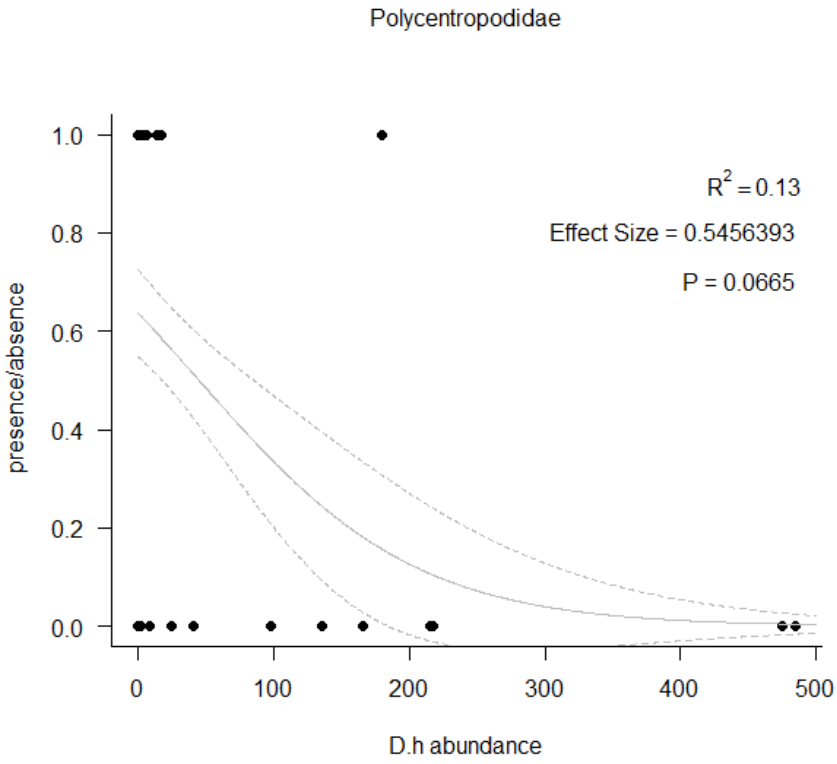

Figure 11: Logistic regression model showing the probability of Polycentropodidae presence/absence under an increasing population of D. haemobaphes.

Family	P value	R^2	Effect Size	95% Confidence Intervals
Baetidae	0.0048	0.407	0.541	0.96 – 0.99
Ecnomidae	0.0329	0.785	0.453	0.63 – 0.94
Ephemerellidae	0.02024	0.29	0.540	0.95 – 0.99
Ephemeridae	0.0838	0.11	0.559	1.00 – 1.02
Polycentropodidae	0.0665	0.13	0.546	0.97 – 0.99

Table 1: A list of the key numerical outputs from the logistic regression models, family response is binary.

Figures 7, 8, 9 and *11* all show that an increasing abundance of *D. haemobaphes* had a negative impact on the probability of occurrence of the families. *Figure 10* shows that an increasing number of *D. haemobaphes* has a positive impact on the presence of Ephemeridae, however, the significance value is lower than the other models and only 11% of the variance is explained by *D. haemobaphes* abundance. Baetidae and Ecnomidae presence/absence both show a particularly strong response to increasing *D. haemobaphes* abundance. Effect sizes for all models are intermediate.

4.1.2 Logistic Regression Models where Family data is continuous and *D. haemobaphes* abundance is Binary.

In this section the logistic regression models show how the presence/absence of *D. haemobaphes* controls the abundance of the specified family. In these models the response variable (family abundance) is shown on the x axis and the control variable (*D. haemobaphes* presence/absence) is on the y axis.

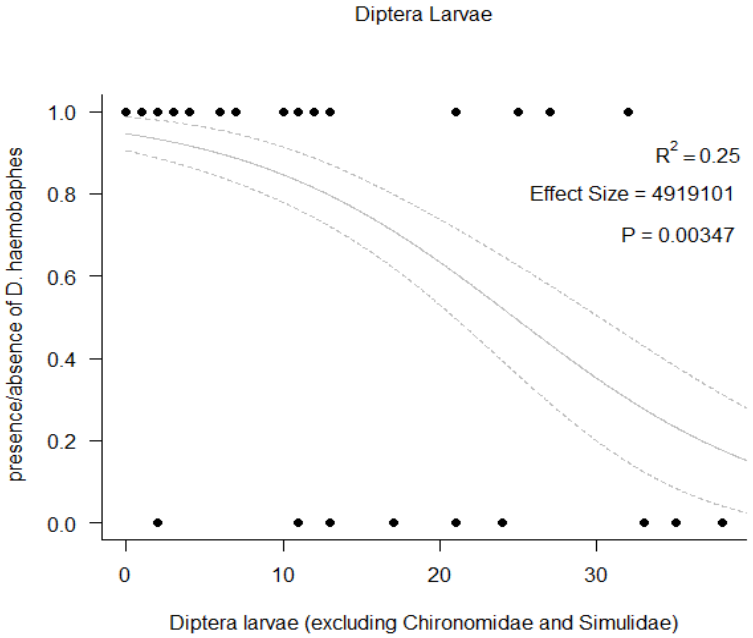

Figure 12: Logistic regression model showing the change in abundance of Diptera larva with the presence/absence of D. haemobaphes.

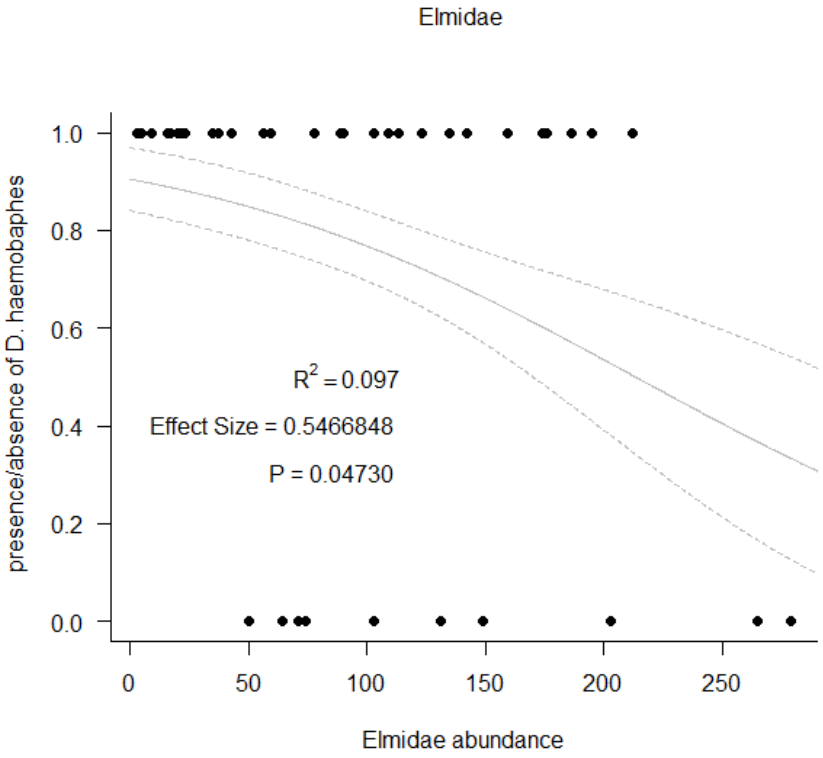

Figure 13: Logistic regression model showing the change in abundance of Elmidae with the presence/absence of D. haemobaphes.

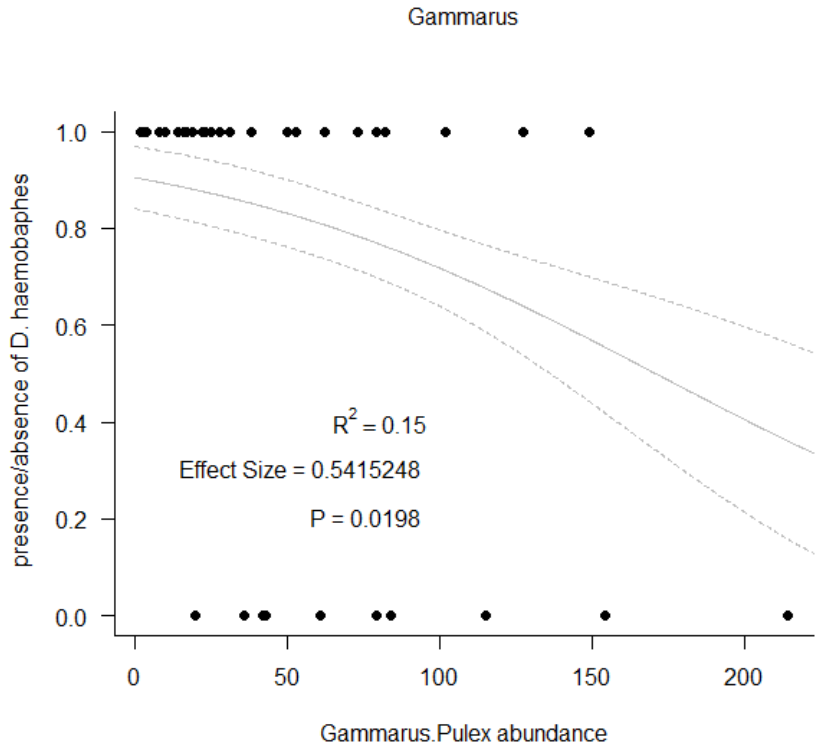

Figure 14: Logistic regression model showing the change in abundance of Gammarus Pulex with the presence/absence of D. haemobaphes.

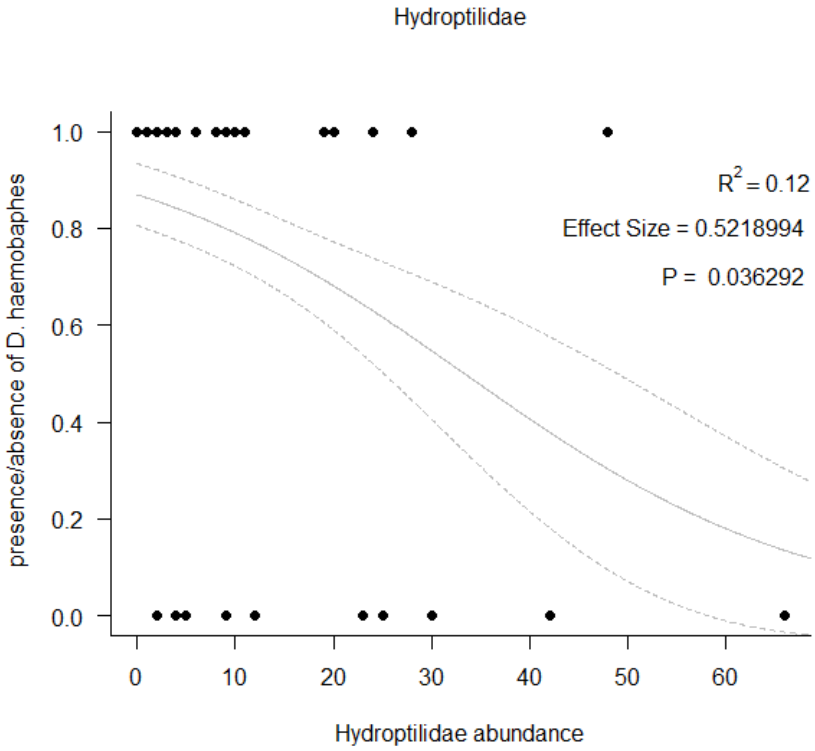

Figure 15: Logistic regression model showing the change in abundance of Hydroptilidae with the presence/absence of D. haemobaphes.

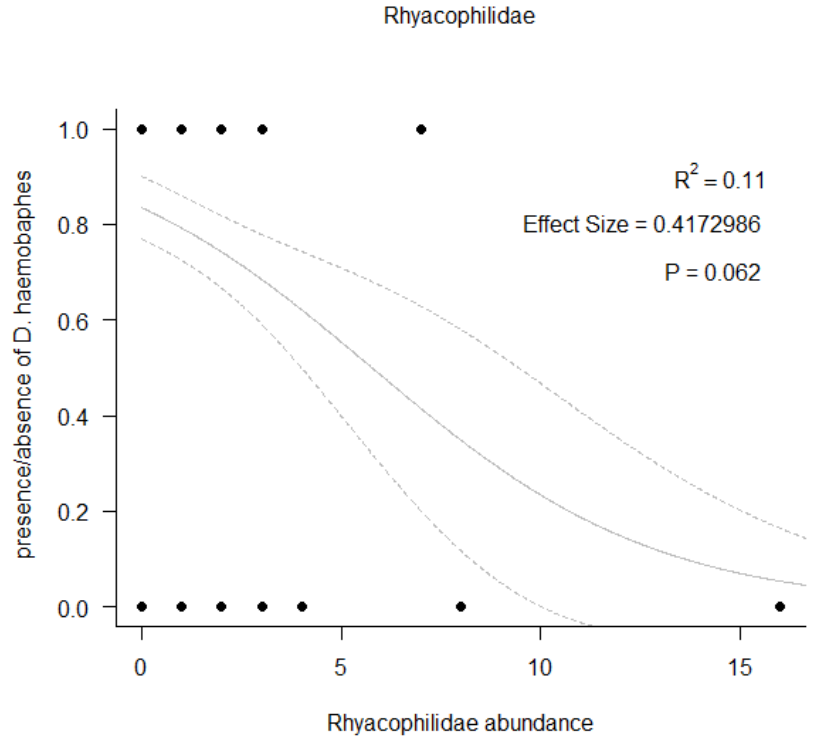

Figure 16: Logistic regression model showing the change in abundance of Rhyacophilidae with the presence/absence of D. haemobaphes.

Family	P value	R^2	Effect Size	95% Confidence Intervals
Diptera larva *(excluding Chironomidae and Simuliidae)*	0.00347	0.248	0.492	0.81 – 0.96
Elmidae	0.0473	0.098	0.546	0.98 – 0.99
Gammarus	0.0198	0.15	0.542	0.96 – 0.99
Hydroptilidae	0.036	0.12	0.522	0.89 – 0.99
Rhyacophilidae	0.062	0.11	0.41	0.53 – 0.97

Table 2: A list of the key numerical outputs from the logistic regression models, where family response is continuous.

Figures 12 – 16 all show that the presence of *D. haemobaphes* has a significant ($p<0.1$) negative impact on the abundances of the respective families. As in *Section 4.1.1*, all effect sizes are intermediate.

4.1.3 Comparison of *D. haemobaphes* and Corophiidae Site Abundances

Based on the findings of Bovy, et al. (2014), a visual comparison of the site abundances of *D. haemobaphes* and Corophiidae was created.

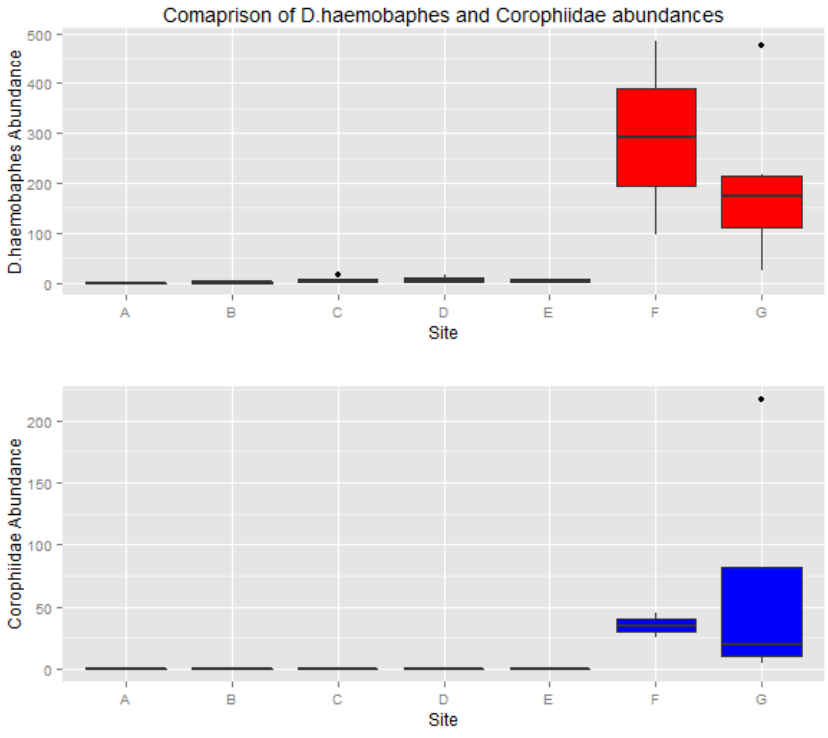

Figure 17: Site abundances of D. haemobaphes (red) and Corophiidae (blue)

Figure 17 shows that where there is a large and dominant population of *D. haemobaphes*, there is also a thriving population of Corophiidae (another invasive ponto-caspian amphipod). The absence of Corophiidae in sites other than F and G suggest that successful invasions into the river have only taken place downstream of the river-canal confluence.

4.2 Field Survey – Site Comparisons

Inter-site comparisons were made using the Sorensen similarity coefficient. Two models were created, one to compare each site and the second to compare samples where *D. haemobaphes* were present and absent. This presence/absence model is displayed both as an ordination plot and in a logistic regression model.

4.2.1 Comparison of Sorensen distances between each site

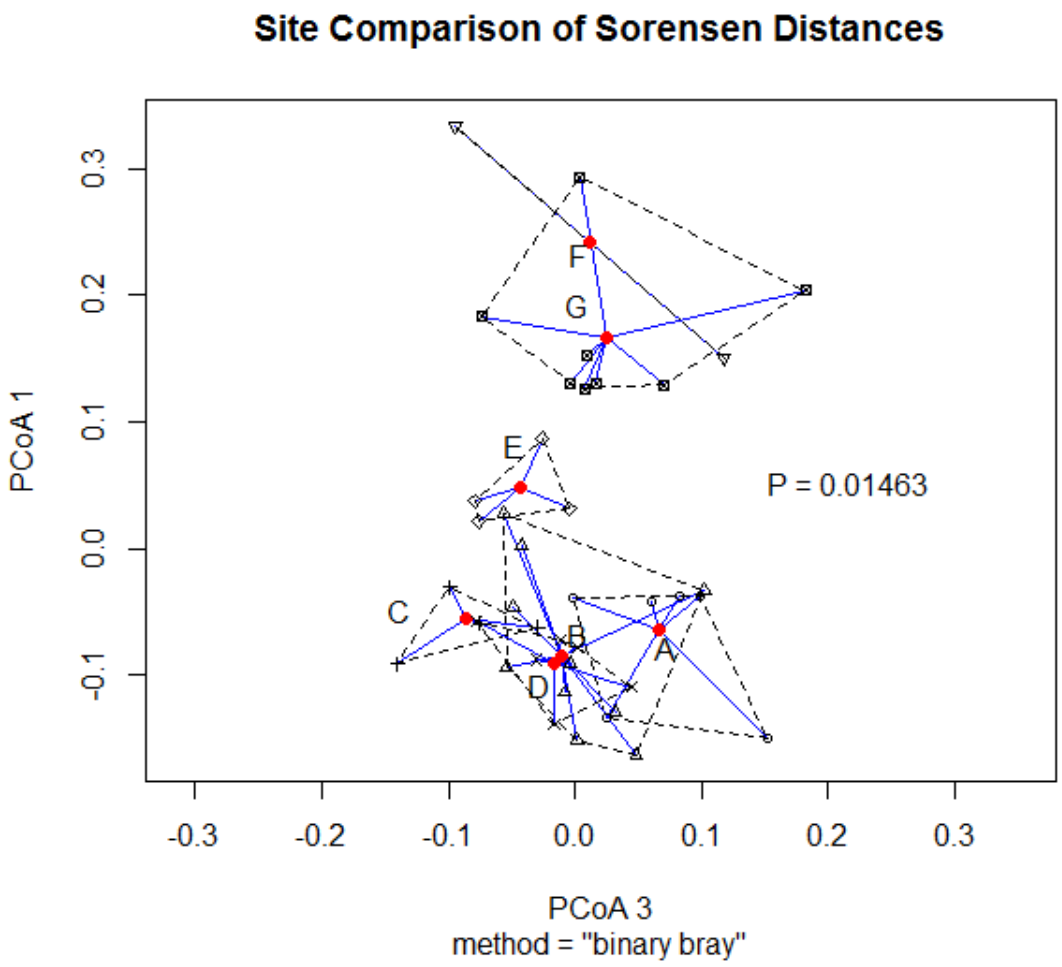

Figure 18: Sorensen distances compared between sites. Red points show site 'centroids'.

Figure 18 shows that the most downstream sites, E, F and G, which have the most dominant populations of *D. haemobaphes* (*Figure 22*) have a different community structure to those sites upstream, under less pressure from *D. haemobaphes*. ANOVA of the model shows a significant difference (P=0.015) between the sites.

4.2.2 Comparison between samples where *D. haemobaphes* are present and absent

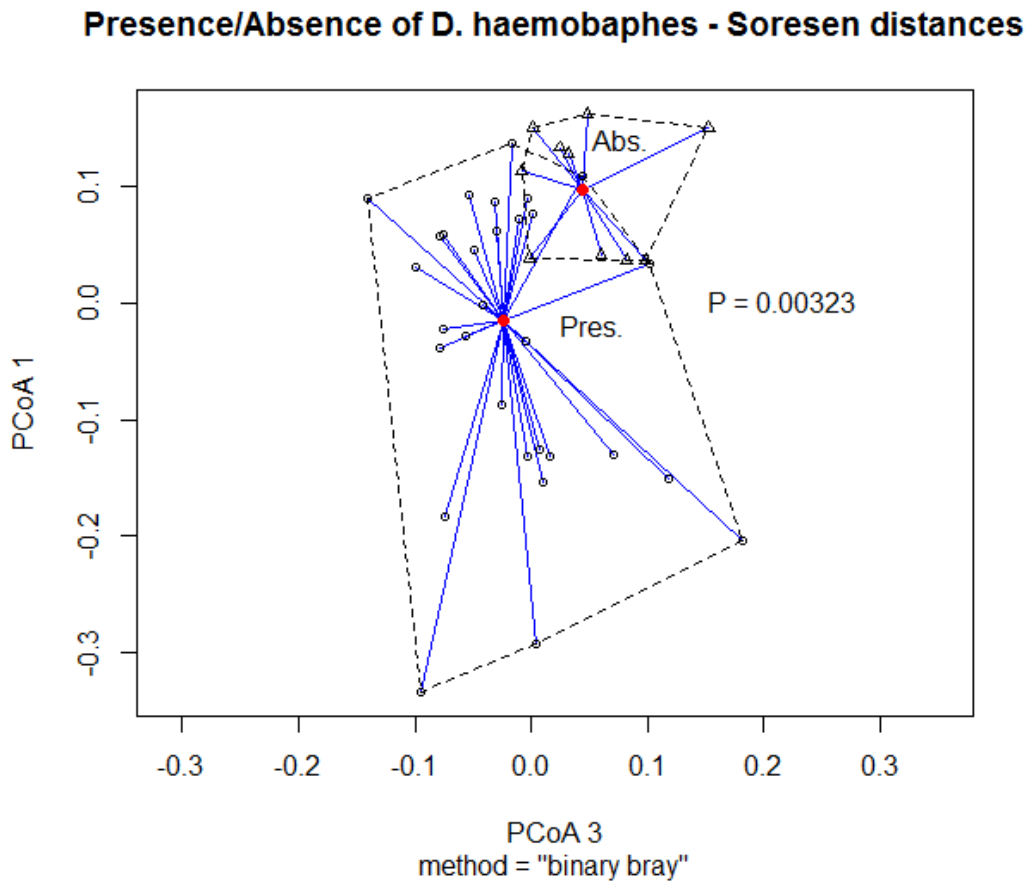

Figure 19: Sorensen distances compared between samples where D. haemobaphes were present and absent. Red points show group 'centroids'.

An ANOVA test run on the model, displayed in *Figure 19*, shows that there is a highly significant difference (P=0.00323) between the communities where *D. haemobaphes* are present or absent. There is a slight overlap of the two groups in *Figure 19*, although in general the plot shows a strong separation between the two groups.

Figure 20, represents the same information as in *Figure 19*. However, the data is presented in the form of a logistic regression. The model shows how the Sorensen

distances between pairs increases with the reduced probability of the presence of *D. haemobaphes*. This model shows that there is a significant (P=0.034) difference between the community structure of samples taken from locations where *D. haemobaphes* are present and absent.

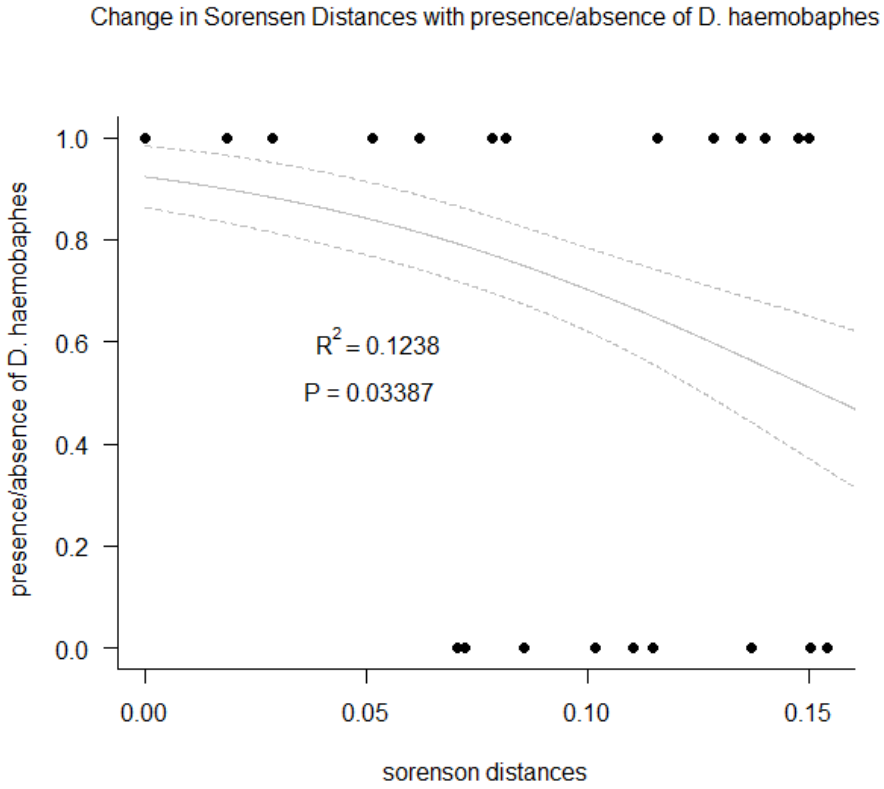

Figure 20: Logistic regression of Sorensen distances between pairs for samples where D. haemobaphes were absent and where they were present. Once again, D. haemobaphes, although on the y axis, is the control variable.

4.3 Field Survey – Functional Feeding Group Response

A general linear mixed effects model (GLMM) was created to investigate the impact of *D. haemobaphes* on the different functional feeding groups in the River Cherwell.

Fixed effects:					
	Estimate Std.	Error	z value	Pr(>|z|)	
(Intercept)	1.032532	0.909201	1.136	0.2561	
D..haemobaphes	-0.002638	0.001503	-1.755	0.07921	.
Gatherer/Collector	0.024194	1.073571	0.022	0.98202	
Predator	-3.481148	1.307137	-2.663	0.00774	**
Scraper	-2.106319	1.196957	-1.76	0.07845	.
Shredder	-3.471517	1.655965	-2.096	0.03605	*
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
Formula: y ~ D..haemobaphes + Trophic.Niche + (1 + D..haemobaphes | sp) + (1 | site)					

Table 3: Fixed effects summary of the GLMM. GLMM formula shown in the lowest cell. Significance values and codes indicated in the two right hand cells. For a full summary table including: random effects summary, AIC and BIC scores, see Appendix 15.

Table 3 shows that the two functional feeding groups which are significantly affected (P<0.05) by the abundance of *D. haemobaphes* are predators and shredders. *Figure 21* visualises the results from *Table 3*, indicating that all functional feeding groups respond negatively to *D. haemobaphes* abundance. However, in order to consider responses as significant in GLMMs, significance levels must be less than 0.05 (Zuur, 2009).The pseudo R^2 value recorded for this model was 0.62.

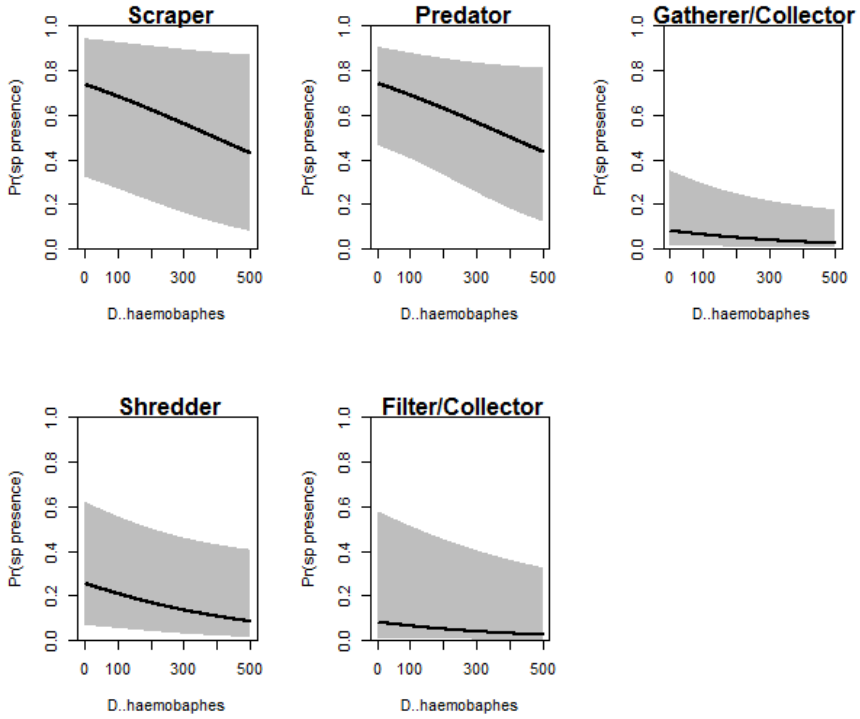

Figure 21: General linear mixed effects model showing the response of five different functional feeding groups to an increasing abundance of D. haemobaphes. Pseudo R^2 = 0.62

4.4 Field Survey – A comparison of Gammaridae community structure

In order to understand the specific impacts that *D. haemobaphes* have on the native Gammaridae community a simple comparison of abundances was undertaken. The proportion of *D. haemobaphes* in the gammarid community was calculated (as a percentage), which is a helpful indicator of community change, irrespective of the influence of habitat.

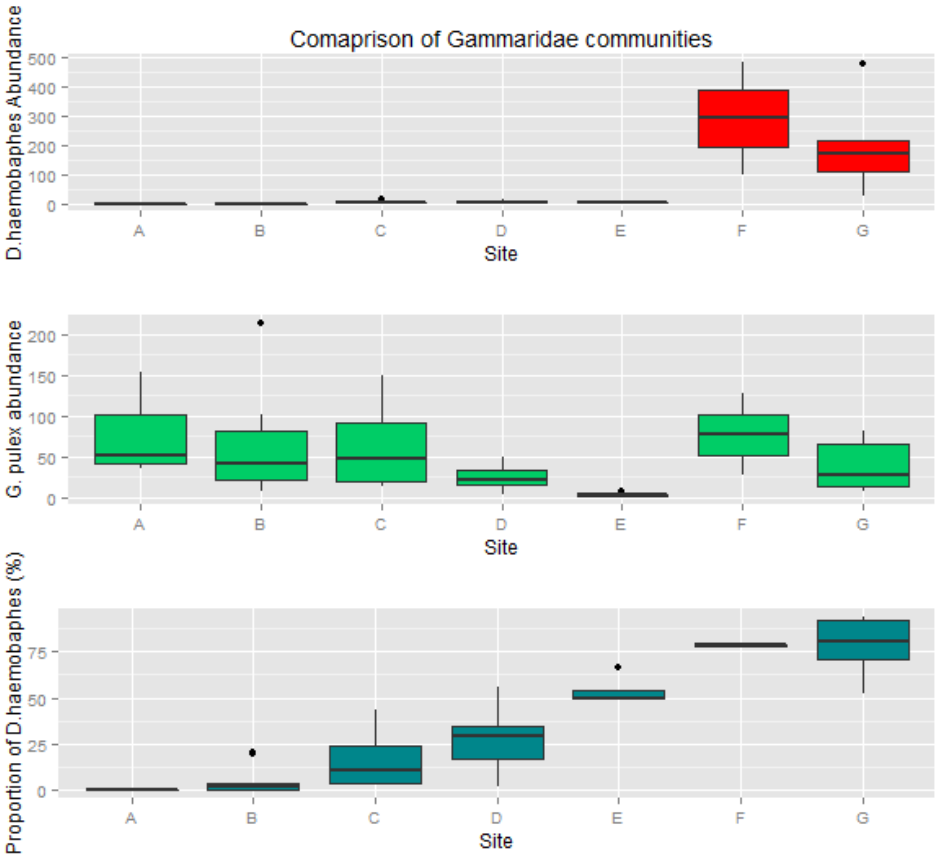

Figure 22: Sequence of boxplots showing the site abundances of D. haemobaphes (red) G. pulex (green) and the Percent of D. haemobaphes in the gammarid community (blue).

Figure 22 helps to illustrate the extent to which *D. haemobaphes* become the dominant gammarid species at each site. By simply looking at the abundances of both species, their interaction is not immediately clear. *Figure 22* shows that with increasing distance downstream, *D. haemobaphes* make up an increasing proportion of the Gammaridae community. Further statistical analysis of the *D. haemobaphes* proportion could not be carried out due to the proportional nature of the data (Crawley, 2005).

4.5 Experimental Study – Leaf Litter Decomposition

A one-way factorial ANOVA was conducted to investigate how a changing ratio of *D. haemobaphes:G. pulex* would affect the decomposition of leaf litter over 22 days. *Figure 23* displays all treatments including the control and shows that, in all treatments where Gammaridae were present, consumption rates were greater than the control. *Figure 24* shows that with an increasing proportion of *D. haemobaphes* there is a highly

significant (P=0.0017) decrease in leaf litter consumption. An Eta-squared value of 0.66 suggests that the effect size of the response is relatively large. Validation plots are presented in *Appendix 13*.

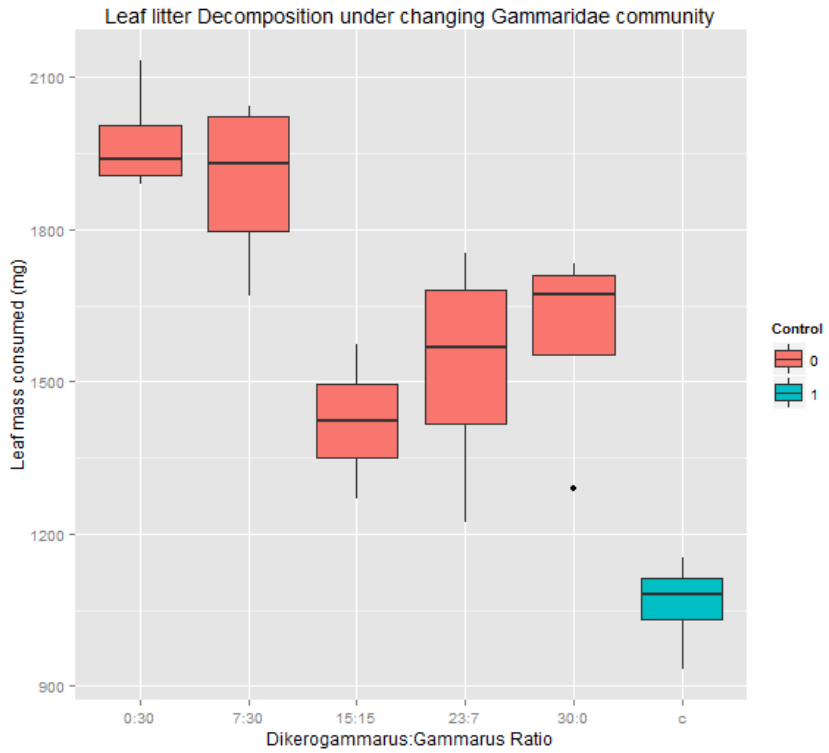

Figure 23: A boxplot displaying the overall results of the leaf litter decomposition experiment. Treatments where Gammaridae were present are shown in red, the control treatment, containing no Gammaridae, is shown in blue.

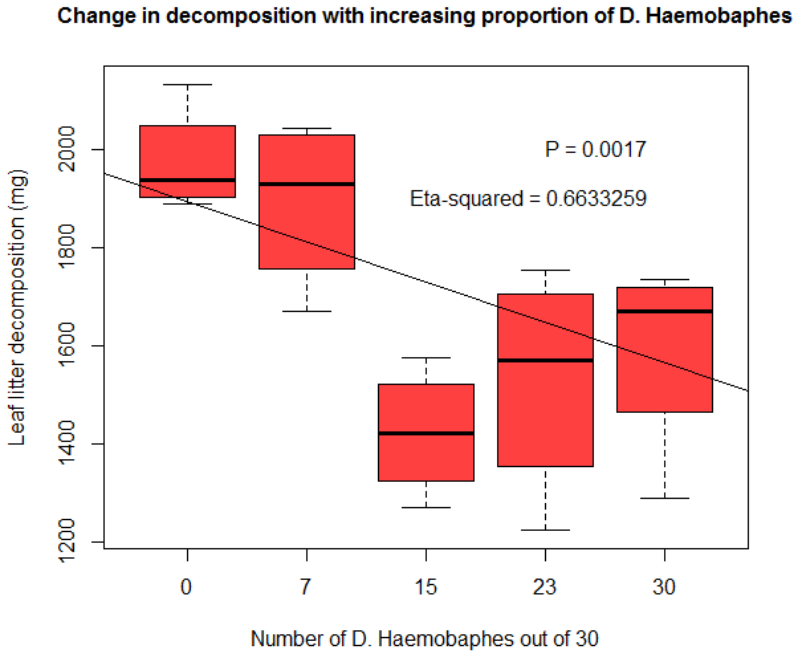

Figure 24: Treatments containing Gammaridae, presented with linear model fit (P=0.0017) and effect size (0.66).

5.0 Discussion

This study has investigated how the invasive gammarid, *D. haemobaphes*, has affected the benthic macroinvertebrate community of the River Cherwell. A survey was conducted along a section of the river that is host to a large *D. haemobaphes* abundance gradient (*Figure 22*), from no individuals present at *Site A*, to dominant populations (>400 per 0.09m^2), making up more than 70% of the gammarid community, downstream of Banbury at Somerton. In addition to the survey, an in-stream enclosure study was undertaken to investigate how changes in Gammaridae community may impact leaf litter decomposition rates.

5.1 The Distribution of *D. haemobaphes* in the River Cherwell

The increasing downstream abundance of *D. haemobaphes* in the River Cherwell was an important factor in allowing the survey to be conducted as it was. This distribution alone is an interesting discussion point as it highlights some important issues about ecological invasions, and more specifically the invasive behaviour of *D. haemobaphes*. In Cropredy, the river and the canal come within 15m of each other and several discharge channels allow for the release of excess canal water into the river. During flooding, it is highly likely that mixing of this water takes place. Upstream of Cropredy, the canal and river are far apart and there can be no interaction. Given the presence of *D. haemobaphes* less than 1km downstream of Cropredy, it would seem that *D. haemobaphes* are not capable of migrating upstream and their downstream dispersal must be controlled by drift. This suggests that any upstream migration, seen elsewhere in Europe (Bij de Vaate, 2002), must be facilitated by a vector such as a boat.

Figure 22 shows the abundances and percentages of *D. haemobaphes* present at each site. There is a clear difference between those sites up and downstream of the canal-river confluence at Nellbridge Farm. Even at *Site E*, located approximately 5km upstream of the confluence, densities of *D. haemobaphes* were less than 10 per 0.09m^2, considerably lower than at *Site G* where densities were more than 20 times higher. This suggests that something is limiting the ability of *D. haemobaphes* to establish the high density communities seen downstream of the confluence. The most likely cause of this varying density is a difference in propagule pressure above and below the river-canal confluence. Upstream of the confluence, introductions of *D.*

haemobaphes are likely to take place in small numbers and infrequently (most likely during high flows). The transmission of the *Dikerogammarus* between the canal and river may also be likely to cause injury, reducing the propagule condition. Downstream of the confluence, there is a continuous supply of large numbers of *D. haemobaphes* all year round and therefore these individuals are likely to be in good condition. This explains the rapid rate at which *D. haemobaphes* have established dominant populations downstream of the confluence. From *Figure 17*, it is clear that Corophiidae must be entering the river from the canal at the confluence. This suggests that the propagule pressure of Corophiidae is not significant enough to establish communities upstream. Feeding rates on Corophiidae by *D. haemobaphes* have been shown to be greater than on some native taxa (Bovy, et al. 2014). Therefore, Corophiidae may also play a role in sustaining the large populations of *D. haemobaphes*, both in the canal and downstream in the river. Further investigation is required to understand this interaction in greater detail.

The nature of this distribution is also important with regards to the alteration of benthic community structure. *Figure 18* shows that sites F and G are significantly different from upstream sites, near Cropredy, which although contain small numbers of *D. haemobaphes*, do not have community assemblages that differ significantly from the control site (*Site A*). This may indicate that where propagule pressures are small, the direct impacts on other taxa may be limited. However, the upper reaches of the Cherwell may simply be indicative of the early stages of invasion, as Grabowski and Bracela (2005) observed in the Vistula valley, where small numbers of *D. haemobaphes* were found to co-occur with native Gammarus species.

Figures 19 and *20* show that overall, the presence of *D. haemobaphes* does have a significant impact on the community structure. The effect size is likely to be greater where abundances are very high (below the canal-river confluence) as shown by the difference in sites in *Figure 18*. However, *Figure 22* illustrates that, despite the smaller abundances in sites B-E, they can make up a sizeable proportion of the gammarid community in these reaches. This suggests that, in small abundances the community impact may be limited, but the impact on gammarid community may be proportionally significant. These alterations to the gammarid community could have knock on impacts for ecosystem function, particularly with regard to organic litter processing.

5.2 The Impacts of *D. haemobaphes* on Individual Families

The impacts on individual families was analysed using a series of logistic regression models. The models were applied to the 20 taxa that were present in sufficiently high numbers to achieve a statistically reliable model. Of the 20 models run, 10 statistically significant responses were identified.

Figures 7 – 11 investigated how an increasing abundance of *D. haemobaphes* would affect the probability of the family's presence or absence. The three most abundant families of mayfly present in the Cherwell were Baetidae (*Figure 7*), Ephemerellidae (*Figure 9*) and Ephemeridae (*Figure 10*). Both Baetidae and Ephemerellidae responded negatively to increases in *D. haemobaphes* abundance, suggesting zero probability of occurrence at densities in excess of 400 per 0.09m^2. Ephemeridae however, was the only family in the study which appeared to respond positively to the abundance of shrimp. It is possible that the negative impact of *D. haemobaphes* on other predators (*Figure 21*) removes other competition, allowing for their success or simply that the size and mobility of Ephemeridae makes them unsuitable prey for *D. haemobaphes*.

Both Ecnomidae (*Figure 8*) and Polycentropodidae (*Figure 11*), filtering / collecting caseless caddis larvae, responded negatively to *D. haemobaphes*. Ecnomidae, in particular, was predicted to have zero probability of occurrence where *D. haemobaphes* densities were greater than 70. This represents a potentially significant alteration to community structure as the mean Ecnomidae density at *Site A* was 27 per 0.09 m^2.

Figures 12-16 show how the presence or absence of *D. haemobaphes* affected the abundances of specific families. The analyses were undertaken in this way due to the families' presence in all (or the significant majority of) samples. All models showed a negative response showing that with the presence of *D. haemobaphes* there were smaller densities of these families. The response of *Gammarus* is in agreement with many other studies on *D. villosus* (Jazdzenski, et al. 2004; Dick, et al. 2002; MacNeil, et al. 2005), which suggests that *G. pulex* are preyed upon by *D. haemobaphes*. It is also likely that competition plays a significant role in the reduced densities of *G. pulex* in the presence of *D. haemobaphes*.

The results from the family response analysis show that the negative impacts of *D. haemobaphes* are not solely limited to amphipod communities and that the diverse predatory habitats observed in *D. villosus* (Dick, et al. 2002; Dodd, et al. 2014) could also be true for *D. haemobaphes*. It is important to note that, as this study focused on a single habitat type, it cannot account for the possible redistribution of taxa into other habitats as seen by Grabowski and Bacela (2005). Although, this is considered to be unlikely because *D. haemobaphes* were found in a variety of habitats during the pilot study, and *D. haemobaphes* are known to survive under a wide range of environmental conditions and in different habitat types (Pockl, 2009; Bij de Vaate, 2002). The survey does not provide information on the mechanism of the negative family responses although it is probable that a combination of predation and competition are the main causes. The effect sizes for all models (*Figures 7 – 16*) were shown to be intermediate. In combination, these effects could equate to a large and significant community effect.

5.3 The Impacts of *D. haemobaphes* on different Functional Feeding Groups

The impact of increasing *D. haemobaphes* density on different functional feeding groups was analysed using a GLMM (*Figure 21*). The results of the model showed that the two groups which were significantly ($P<0.05$) affected were predators and shredders (*Figure 21 and Table 3*). These trophic niches are both utilised by *D. haemobaphes* and therefore both predators and shredders are more likely to come into contact with *D. haemobaphes*, putting them at greater risk from predation and competition. A pseudo R^2 of 0.62 suggests that the variance in the model is explained reasonably well by the density of *D. haemobaphes* which indicates that this community level response is potentially of great concern. The fact that there is both a response at a family and trophic group level shows that changes in ecosystem function are highly likely. The long term implications of these changes are significant unknowns which require further investigation.

5.4 How do *D. haemobaphes* affect leaf litter decomposition?

This study and others (MacNeil and Platvoet, 2005; Dick and Platvoet, 2000; Kinzler, et al. 2009) have clearly shown that the presence of *Dikerogammarus* can lead to the

replacement of native gammarid species. The implications of this are compounded by the possibility that the rates of CPOM breakdown will be significantly affected. Increasing proportions *of D. villosus* have been shown to have a negative impact on leaf litter decomposition rates (Piscart, et al. 2011; MacNeil, et al. 2011). Conversely, leaf litter decomposition by *D. villosus* has been shown to be greater than *G. pulex* under elevated temperatures. The experiment undertaken in this study was conducted in the River Cherwell, in a location where *D. haemobaphes* have established dominant populations (*Figure 23*). The experiment revealed that leaf litter decomposition was significantly reduced (P=0.0017) by an increasing proportion of *D. haemobaphes*. The eta-squared value of the response was calculated as 0.66 suggesting that the effect size was relatively large. The results show that where there is a mixture of native and non-native Gammaridae the consumption rate is slightly lower than predicted. This may be caused by a re-prioritisation of activities – *Gammarus* may consume less due to the evasion of predators and *D. haemobaphes* may focus their energy on the capture of *G. pulex*, a more energy-rich food source. Further investigation would be needed to validate these assumptions.

This study is in agreement with MacNeil, et al (2011) and Piscart, et al. (2011), finding that *Dikerogammarus* have a negative impact on the rate of leaf litter processing in rivers. The implications of these findings are highly important in terms of ecosystem function, because it seems that even small densities of *D. haemobaphes* may be capable of significantly altering breakdown rates of leaf litter (and CPOM). This could have many impacts, both direct and indirect, across multiple trophic levels.

5.5 Limitations

(i) As the study was only conducted on the River Cherwell, it can only provide information specific to that water body. The findings are relevant elsewhere with consideration to local environmental, hydrological and geomorphological factors.

(ii) In order to study the specific impacts of *D. haemobaphes* and allow comparison between samples, a single habitat type was selected. It is possible that this may have led to the reporting of bias responses in some families. However, this is considered unlikely due to the presence of *D. haemobaphes* across multiple habitat types.

(iii) Due to time constraints it was not possible to identify macroinvertberates beyond family level. Species level identification would have provided more detail.

(iv) Other factors that could have affected the macroinvertebrate assemblage of the river, such as pollution were not measured due to time constraints. However, the presence of high scoring BMWP taxa at all sites suggests that there is not a significant difference in organic pollution between sites.

(v) The experiment was carried out in small enclosures and the results are therefore not directly comparable with natural conditions. Enclosure environments can alter the behaviour and interactions of species depending on the size and design of the enclosures, and the duration of the experiment (Vance-Chalcraft, et al. 2004). However, the study provides valuable information on the nature of the response which is highly likely to be representative of the natural system.

5.6 Further Research

This study is one of relatively few that have looked at the specific implications of *D. haemobaphes* invasion. The species is now abundant in numerous water bodies across the UK and clearly presents some significant threats to the assemblage of benthic macroinvertebrate fauna and ecosystem function. As a result the species warrants further study, partuicularly with regards to its potential impacts in British waters. The following are suggested areas that may help to build understanding of the potential impacts that *D. haemobaphes* present to freshwater environments:

(i) Bovy, et al. (2014) suggest that the presence of the invasive amphipod *C. curvispinum* may help to facilitate the expansion of *D. haemobaphes*. Due to the limited sites at which *C. curvispinum* were found it was not possible to determine the interaction between the two species. A more extensive survey at locations where both species are present may help to determine if the establishment of high densities of *D. haemobaphes* is reliant on prey that have evolved alongside them.

(ii) So far, much of the literature on *Dikerogammarus* has focused on *D. villosus*. A more detailed understanding of the type of fauna that *D. haemobaphes* prey on would provide valuable information on the causes of some of the responses presented in this study. This could be achieved through conducting mesocosm studies that investigate the type of prey that *D. haemobaphes* preferentially feed upon (A design similar to the study on *D.villosus* by Dick et al. (2002) could prove extremely useful).

(iii) Kelleher, et al. 1998 found that the presence of *D. haemobaphes* caused fish, particularly Percidae, Gobiidae and Anguillidae families, to focus their feeding on *D.*

haemobaphes, suggesting that increases in *D. haemobaphes* may be beneficial for fish. However, Casellato, et al. (2007) observed *D. villosus* predating on small fish. If *D. haemobaphes* are also capable of predating on fish then the species may be detrimental to freshwater fish populations. Further investigation in to the interaction of *D. haemobaphes* and different British fish species is required to determine if *D. haemobaphes* could significantly alter fish communities.

(iv) In the early stages of *D. haemobaphes* invasion, the species is likely to cohabit with native *G. pulex*. The results from this study suggest that the interactions between the species may result in complex changes in leaf litter decomposition. An investigation into the interaction between *G. pulex* and *D. haemobaphes* could help to identify if there is a change in behaviour in the presence of Gammaridae from a different species.

(v) This study was based on data taken from a single water body. Similar studies on larger and smaller rivers, canals and still waters would provide invaluable information about the species' potential impacts in different environments.

(vi) This study has identified that one of the biggest threats presented by the invasion of *D. haemobaphes* is an alteration to ecosystem function. The long term impacts of these changes could have serious impacts on British freshwater ecosystems. Long term monitoring of organic matter processing and macroinvertbrate assemblages, in affected streams, would provide some indication of the changes that may take place in the future.

6.0 Conclusions

A survey of macroinvertebrates along the River Cherwell, between Cropredy and Somerton, and an in-situ experimental study have been used to identify a number of potential impacts that the invasive gammarid, *Dikerogammarus haemobaphes*, presents to freshwater environments in the UK. The main conclusions of this study are as follows:

(i) The densities *D. haemobaphes* and the proportion of *D. haemobaphes* to *G. pulex* increased downstream. No *D. haemobaphes* were found upstream of Cropredy; approximately 75% of the gammarid community at Somerton consisted of *D. haemobaphes*, with maximum densities in excess of 400 per 0.09m^2 (at sites F and G).

(ii) A comparison of invertebrate communities revealed that there was a significant difference between sites (*Figure 18*). In particular, community assemblages at sites with large densities of *D. haemobaphes* (sites F and G) were considerably different to those without or with low densities of *D. haemobaphes*.

(iii) Community assemblages of samples, where *D. haemobaphes* were present and absent, were compared (*Figure 19*), showing that there was a significant difference ($P = 0.034$) between community structure where the invasive gammarid was present and absent (*Figure 20*).

(iv) Although changes to the community as a whole are limited where *D. haemobaphes* densities are low, the proportional changes to gammarid community (*Figure 22*) can be considerable.

(v) Family level impacts were analysed using a series of logistic regression models (*Figures 7:16*). Half of the families that were analysed showed a significant ($P<0.1$) response to *D. haemobaphes*. 9/10 of these responses were negative, with Ephemeridae being the only family whose likelihood of occurrence increased with increasing *D. haemobaphes* density.

(vi) Increases in *D. haemobaphes* led to the replacement of native *G. pulex* (*Figure 14 and 22*). There were also negative impacts to a wide variety of families from a range of trophic groups (*Figures 7:16*). Predation and competition are considered to be the most likely causes of the observed negative responses.

(vii) Intermediate effect sizes were recorded for all family level responses. It is possible that the cumulative effects could be much greater at a community level.

(viii) A GLMM was used to investigate which functional feeding groups were most affected by *D. haemobaphes* abundance. The results showed that all groups were negatively affected but only predators and shredders were significantly ($P < 0.05$) affected. These are the trophic areas that are most commonly utilised by *D. haemobaphes* and therefore, the fauna in these groups, are more likely to interact with the invasive gammarid. This has the potential to cause significant changes to ecosystem function.

(ix) An experimental in-stream enclosure study was created to investigate the impact of an increasing proportion of *D. haemobaphes* (within the gammarid community) on the decomposition of leaf litter. The results showed that with an increasing proportion of *D. haemobaphes* there was a significant ($P = 0.0017$) decrease in the amount of leaf litter decomposition. The effect size of the response was calculated as 0.66 indicating that the effect of the response was relatively large. Therefore, even small numbers of *D. haemobaphes* have the ability to alter the rates of CPOM decomposition in freshwater environments. This has major implications for ecosystem function and is likely to directly and indirectly effect fauna from multiple trophic levels.

7.0 References

Aldridge, D.C. (2013). GB Non-native Organism Rapid Risk Assessment for *Dikerogammarus haemobaphes* (Eichwald, 1841). www.nonnativespecies.org

Bacela, K., Konopacka, A., & Grabowski, M. (2009). Reproductive biology of Dikerogammarus haemobaphes: an invasive gammarid (Crustacea: Amphipoda) colonizing running waters in Central Europe. *Biological invasions*,*11*(9), 2055-2066.

Bij de Vaate, A., Jazdzewski, K., Ketelaars, H. A., Gollasch, S., & Van der Velde, G. (2002). Geographical patterns in range extension of Ponto-Caspian macroinvertebrate species in Europe. *Canadian Journal of Fisheries and Aquatic Sciences*, *59*(7), 1159-1174.

Blumstein, D. T. (2006). Developing an evolutionary ecology of fear: how life history and natural history traits affect disturbance tolerance in birds. *Animal behaviour*, *71*(2), 389-399.

Bovy, H. C., Barrios-O'Neill, D., Emmerson, M. C., Aldridge, D. C., & Dick, J. T. (2014) Predicting the predatory impacts of the "demon shrimp" Dikerogammarus haemobaphes, on native and previously introduced species. *Biological Invasions*, 1-11.

Brooks, M. L., D'antonio, C. M., Richardson, D. M., Grace, J. B., Keeley, J. E., DiTOMASO, J. M.,& Pyke, D. (2004). Effects of invasive alien plants on fire regimes. *BioScience*, *54*(7), 677-688.

Byrnes, J.. (2008). *Coefficient of determination (R^2) when using lme().* Available: http://thread.gmane.org/gmane.comp.lang.r.lme4.devel/684. Last accessed 10th Aug 2014.

Carlton, J. T. (1999). *Invasive species and biodiversity management*. Dordecht: Kluwer Academic. 195-212.

Casellato, S., Visentin, A., & La Piana, G. (2007). The predatory impact of Dikerogammarus villosus on fish. In *Biological invaders in inland waters: Profiles, distribution, and threats* (pp. 495-506). Springer Netherlands.

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in medicine*, *19*(22), 3127-3131.

Crawley, M. J. (2005). *Statistics:An Introduction Using R*. Chichester: John Wiley and Sons ltd.

Dick, J. T., Platvoet, D., & Kelly, D. W. (2002). Predatory impact of the freshwater invader Dikerogammarus villosus (Crustacea: Amphipoda).*Canadian Journal of Fisheries and Aquatic Sciences*, *59*(6), 1078-1084.

Dodd, J. A., Dick, J. T., Alexander, M. E., MacNeil, C., Dunn, A. M., & Aldridge, D. C. (2014). Predicting the ecological impacts of a new freshwater invader: functional responses and prey selectivity of the 'killer shrimp', Dikerogammarus villosus, compared to the native Gammarus pulex. *Freshwater Biology*, *59*(2), 337-352.

Environment Agency. (2012). *Invasive shrimp (D. haemobaphes) Interim Briefing Note.*Available:

file:///C:/Users/Hugh/Downloads/Haemobaphes_Interim_National_Briefing_121112%20(1).pdf. Last accessed 10th Aug 2014.

Environment Agency. (2013). *What's in my Back Yard.* Available: http://maps.environmentagency.gov.uk/wiyby/wiybyController?topic=wfd_rivers&layerGroups=default&lang=_e&ep=map&scale=1&x=357682.99999999994&y=355133.99999999994. Last accessed 10th Aug 2014.

Gallardo, B., & Aldridge, D. C. (2013). The 'dirty dozen': socio-economic factors amplify the invasion potential of 12 high-risk aquatic invasive species in Great Britain and Ireland. *Journal of Applied Ecology*, *50*(3), 757-766.

Gallardo, B., Errea, M. P., & Aldridge, D. C. (2012). Application of bioclimatic models coupled with network analysis for risk assessment of the killer shrimp, Dikerogammarus villosus, in Great Britain. *Biological Invasions*, *14*(6), 1265-1278.

Grabowski, M., Bacela, K., & Konopacka, A. (2007). How to be an invasive gammarid (Amphipoda: Gammaroidea)–comparison of life history traits.*Hydrobiologia*, *590*(1), 75-84.

Graca, M. A. (2001). The role of invertebrates on leaf litter decomposition in streams–a review. *International Review of Hydrobiology*, *86*(4-5), 383-393.

Hladyz, S., Tiegs, S. D., Gessner, M. O., Giller, P. S., RÎŞNOVEANU, G., Preda, E., & Woodward, G. (2010). Leaf-litter breakdown in pasture and deciduous woodland streams: a comparison among three European regions.*Freshwater Biology*, *55*(9), 1916-1929.

Jamil, T., Ozinga, W. A., Kleyer, M., & ter Braak, C. J. (2013). Selecting traits that explain species–environment relationships: a generalized linear mixed model approach. *Journal of Vegetation Science*, *24*(6), 988-1000.

Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M.

Jazdzewski, K., Konopacka, A., & Grabowski, M. (2004). Recent drastic changes in the gammarid fauna (Crustacea, Amphipoda) of the Vistula River deltaic system in Poland caused by alien invaders. *Diversity and Distributions*,*10*(2), 81-87.

Kelleher, B., Bergers, P. J. M., Van den Brink, F. W. B., Giller, P. S., Van der Velde, G., & Bij de Vaate, A. (1998). Effects of exotic amphipod invasions on fish diet in the Lower Rhine. *Archiv für Hydrobiologie*, *143*(3), 363-382.

Kinzler, W., Kley, A., Mayer, G., Waloszek, D., & Maier, G. (2009). Mutual predation between and cannibalism within several freshwater gammarids: Dikerogammarus villosus versus one native and three invasives. *Aquatic Ecology*, *43*(2), 457-464.

Kley, A., & Maier, G. (2006). Reproductive characteristics of invasive gammarids in the Rhine-Main-Danube catchment, South Germany.*Limnologica-Ecology and Management of Inland Waters*, *36*(2), 79-90.

Lockwood, et al. (2007). *Invasion Ecology*. Oxford: Blackwell Publishing.

Lockwood, J. L., Cassey, P., & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution*, *20*(5), 223-228.

Lockwood, J. L., Cassey, P., & Blackburn, T. M. (2009). The more you introduce the more you get: the role of colonization pressure and propagule pressure in invasion ecology. *Diversity and Distributions*, *15*(5), 904-910.

Lozon, J. D., & MacIsaac, H. J. (1997). Biological invasions: are they dependent on disturbance?. *Environmental Reviews*, *5*(2), 131-144.

MacDougall, A. S., & Turkington, R. (2005). Are invasive species the drivers or passengers of change in degraded ecosystems?. *Ecology*, *86*(1), 42-55.

Mack, R. N. (2004). *Invasive species: vectors and management strategies*. Washington, DC: Island Press. 3-30.

MacNeil, C., & Platvoet, D. (2005). The predatory impact of the freshwater invader Dikerogammarus villosus on native Gammarus pulex (Crustacea: Amphipoda); influences of differential microdistribution and food resources.*Journal of Zoology*, *267*(1), 31-38.

MacNeil, C., Dick, J. T., & Elwood, R. W. (1997). The trophic ecology of freshwater Gammarus spp.(Crustacea: Amphipoda): problems and perspectives concerning the functional feeding group concept. *Biological Reviews of the Cambridge Philosophical Society*, *72*(03), 349-364.

MacNeil, C., Dick, J. T., Platvoet, D., & Briffa, M. (2011). Direct and indirect effects of species displacements: an invading freshwater amphipod can disrupt leaf-litter processing and shredder efficiency. *Journal of the North American Benthological Society*, *30*(1), 38-48.

MacNeil, C., Dick, J. T., Platvoet, D., & Briffa, M. (2011). Direct and indirect effects of species displacements: an invading freshwater amphipod can disrupt leaf-litter processing and shredder efficiency. *Journal of the North American Benthological Society*, *30*(1), 38-48.

Minchinton, T. E. (2002). Disturbance by wrack facilitates spread of Phragmites australis in a coastal marsh. Journal of Experimental Marine Biology and Ecology, 281(1), 89-107.

MiniScale [NTF geospatial data], Scale 1:, Tile(s): , Updated: Jul 2014, Ordnance Survey (GB), Using: EDINA Digimap Ordnance Survey Service, <http://edina.ac.uk/digimap>, Downloaded: Aug 2014

Mooney, H. A. (2005). Invasive alien species*: a new synthesis*. Washington, DC: Island Press.

Müller, J. C., Schramm, S., & Seitz, A. (2002). Genetic and morphological differentiation of Dikerogammarus invaders and their invasion history in Central Europe. *Freshwater Biology*, *47*(11), 2039-2048.

Müller, J., & Schramm, S. (2001). A third Dikerogammarus invader is located in front of Vienna. *Lauterbornia*, *41*, 49-52.

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. Biological Reviews, 82(4), 591-605.

National research council (United States) (2002). Predicting invasions of nonindigenous plants and plant pests. Washington, DC: National Acadamy Press.

Neal, C., Neal, M., Hill, L., & Wickham, H. (2006). River water quality of the River Cherwell: an agricultural clay-dominated catchment in the upper Thames Basin, southeastern England. *Science of the total environment*, *360*(1), 272-289.

Pimentel, D., Lach, L., Zuniga, R., & Morrison, D. (2000). Environmental and economic costs of nonindigenous species in the United States. *BioScience*,*50*(1), 53-65.

Piscart, C., Mermillod-Blondin, F., Maazouzi, C., Merigoux, S., & Marmonier, P. (2011). Potential impact of invasive amphipods on leaf litter recycling in aquatic ecosystems. *Biological invasions*, *13*(12), 2861-2868.

Pöckl, M. (2009). Success of the invasive Ponto-Caspian amphipod Dikerogammarus villosus by life history traits and reproductive capacity.*Biological invasions*, *11*(9), 2021-2041.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/.

S. Pawley, M. Dobson & M. Fletcher (2011). *Guide to British Freshwater Macroinvertebrates for Biotic Assessment*. Ambleside: The Freshwater Biological Association.

Sax, D. F., Stachowicz, J. J., & Gaines, S. D. (2005). *Species invasions: insights into ecology, evolution and biogeography*. Sunderland, Sinauer Associates Incorporated.

Schreiber, E. S. G., Quinn, G. P., & Lake, P. S. (2003). Distribution of an alien aquatic snail in relation to flow variability, human activities and water quality.*Freshwater Biology*, *48*(6), 951-961.

Simberloff, D., & Von Holle, B. (1999). Positive interactions of nonindigenous species: invasional meltdown?. *Biological invasions*, *1*(1), 21-32.

Smith, M. D., & Knapp, A. K. (1999). Exotic plant species in a C4-dominated grassland: invasibility, disturbance, and community structure. *Oecologia*, *120*(4), 605-612.

Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., & Midgley, P. M. (2013). Climate change 2013: The physical science basis.*Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5)(Cambridge Univ Press, New York)*.

Van der Velde, G., Rajagopal, S., Kelleher, B., Musko, I. B., & de Vaate, A. B. (2000). Ecological impact of crustacean invaders: general considerations and examples from the Rhine River. *Crustacean Issues*, *12*, 3-34.

Van Riel, M. C., & Van der Velde, G. (2006). To conquer and persist: colonization and population development of the Ponto-Caspian amphipods Dikerogammarus villosus and Chelicorophium curvispinum on bare stone substrate in the main channel of the River Rhine. *Archiv für Hydrobiologie*,*166*(1), 23-39.

Vance-Chalcraft, H. D., Soluk, D. A., & Ozburn, N. (2004). Is prey predation risk influenced more by increasing predator density or predator species richness in stream enclosures?. *Oecologia*, *139*(1), 117-122.

Wallace, J. B., & Webster, J. R. (1996). The role of macroinvertebrates in stream ecosystem function. *Annual review of entomology*, *41*(1), 115-139.

Wonham, M. J., Walton, W. C., Ruiz, G. M., Frese, A. M., & Galil, B. S. (2001). Going to the source: role of the invasion pathway in determining potential invaders. *Marine Ecology Progress Series*, *215*(1-12).

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. New York: Springer Science + Business Media.

7.1 R Package References

Andrew Gelman and Yu-Sung Su (2014). arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.7-07. http://CRAN.R-project.org/package=arm

Baptiste Auguie (2012). gridExtra: functions in Grid graphics. R package version 0.9.1. http://CRAN.R-project.org/package=gridExtra

Bates D, Maechler M, Bolker B and Walker S (2014). _lme4: Linear mixed-effects models using Eigen and S4_. R package version 1.1-7, <URL: http://CRAN.R-project.org/package=lme4>.

H. Wickham. ggplot2: elegant graphics for data analysis. Springer New York, 2009.

Henry H. Stevens and Helene Wagner (2013). vegan: Community Ecology Package. R package version 2.0-10. http://CRAN.R-project.org/package=vegan

John Fox and Sanford Weisberg (2011). An {R} Companion to Applied Regression, Second Edition. Thousand Oaks CA: Sage. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion

Navarro, D. J. (2014) Learning statistics with R: A tutorial for psychology students and other beginners. (Version 0.4) University of Adelaide. Adelaide, Australia

8.0 Appendices

8.1 Appendix 1. Pilot Study

Before commencing the main survey, presented in this dissertation, a pilot survey was undertaken to refine the survey design. Six 1 minute long surber samples were taken from *Site A* (Cropredy Manor) in different riffle micro habitats. These were described as follows:

(A) Riffle habitat, Gravel substrate, low algae abundance and macrophyte growth.

(B) Riffle habitat, Small macrophyte reed bed, gravel substrate.

(C) Shallow riffle habitat, slower flow, lots of algae and biofilm, medium sized cobble substrate.

(D) Sandy margin habitat, slow flow, large cobbles below sand.

(E) Fast riffle habitat, large cobble substrate, lots of green algae and macrophyte growth.

(F) End of riffle, approximately 50cm deep, large cobbles and sand, little algal growth.

The same method was repeated at *Site G* (Somerton Bridge) at corresponding habitats. All of the invertebrates in the samples were identified and counted in the laboratory at the University of Birmingham. Following the processing of data, an informed decision was made on the habitat type and sampling time was reduced to 30 seconds. The abundances of *G. pulex* at *Site A* and *D. haemobaphes* at *Site G*, for each habitat, are shown in *Figure 25*.

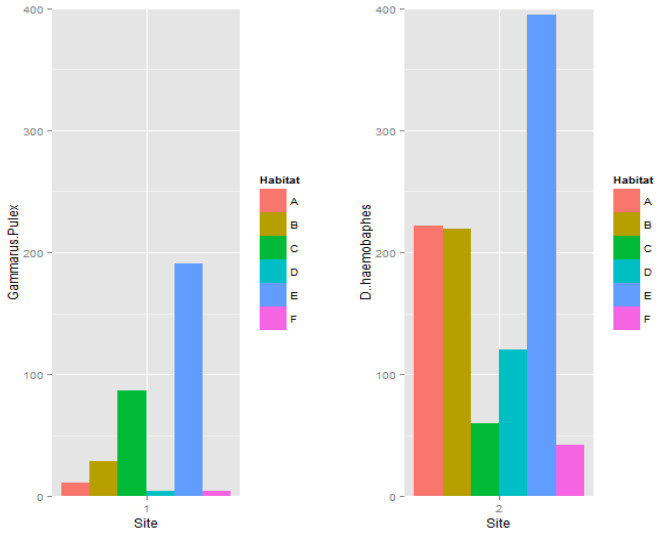

Figure 25: A summary of the abundances of the dominant Gammaridae species in different habitats. Left hand bar plot = Site A, right hand barplot = Site G.

8.2 Appendix 2. Site A Photos

Figure 26: Left – looking downstream from the footbridge at Site A. Right – looking upstream from the footbridge.

8.3 Appendix 3. Site B Photos

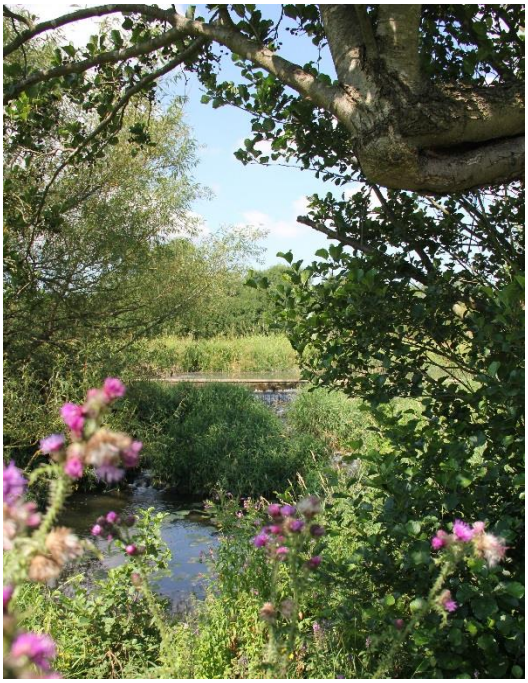

Figure 27: Weir at the upstream boundary of Site B

Figure 28: Examples of artificially created riffles from Site B

8.4 Appendix 4. Site C Photos

Figure 29: Examples of sample locations from Site C.

8.5 Appendix 5. Site D Photos

Figure 30: Sample locations, Site D

8.6 Appendix 6. Site E Photos

Figure 31: Photo of Site E

8.7 Appendix 7. Site F Photos

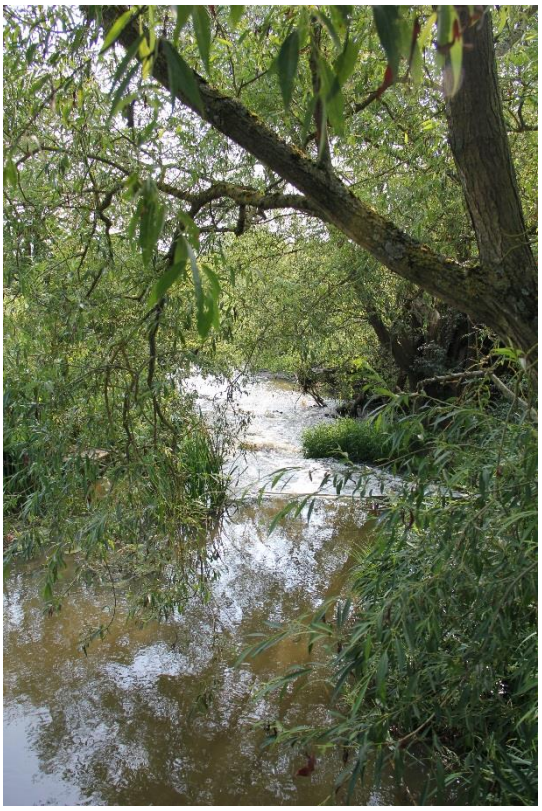

Figure 32: Photo of Site E

8.8 Appendix 8. Site G Photos

Figure 33: Upstream of Somerton Bridge, Site G. Red circle indicates the location of the experimental study.

Figure 34: Downstream of Somerton Bridge (Site G), left – broken weir, right natural riffle

8.9 Appendix 9. Logistic Regression, family level comparison R Script

```
##############################
#LOGISTIC REGRESSION ANALYSIS
##############################


rm(list=ls())   # Clear Memory

# set working directory

getwd()  # check directory location


Survey <- read.csv(file.choose())

str(Survey)
Survey$number <- as.factor(Survey$number )
```

```
Survey["All.Elmidae"] <- NA # That creates the new column named "MY_NEW_COLUMN" filled with "NA"
Survey$All.Elmidae <- Survey$Elmidae..Adult. + Survey$Elmidae.Elmis..larva. + Survey$Elmidae.Limnius..larva.
str(Survey, list.len=1000)

Survey["Diker.Haem.Y.N"] <- NA # That creates the new column named "MY_NEW_COLUMN" filled with "NA"
Survey$Diker.Haem.Y.N <- Survey$D..haemobaphes
str(Survey, list.len=1000)

Survey$Diker.Haem.Y.N[Survey$Diker.Haem.Y.N>0] <-1
######### sort out data ######

Binary.Survey <- Survey

Binary.Survey$number <- NULL
Binary.Survey$Number.of.Families <- NULL
Binary.Survey$TOTAL <- NULL
Binary.Survey$BMWP <- NULL
Binary.Survey$Site <- NULL
Binary.Survey$BMWP <- NULL
Binary.Survey$ASPT <- NULL
Binary.Survey$Dikero.Gam.ratio <- NULL


Binary.Survey[,c(1:44)] [Binary.Survey[,c(1:44)]>0] <-1
Binary.Survey["Dikero.All"] <- NA
Binary.Survey$Dikero.All <- Survey$D..haemobaphes
Binary.Survey$D..haemobaphes <- NULL
Binary.Survey["Dikero.Gam.ratio"] <- NA
Binary.Survey$Dikero.Gam.ratio <- Survey$Dikero.Gam.ratio


#########################################################################
#Acroloxidae
#########################################################################

boxplot(Dikero.All~ Acroloxidae, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Acroloxidae")
plot(Acroloxidae~Dikero.All, pch = 19, data = Binary.Survey)

Dikero.glm1 <- glm(Acroloxidae~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model
```

```
summary(Dikero.glm1)   # p = 9.025 not significant - don't plot
##############################################################################
#Asellidae
##############################################################################
boxplot(Dikero.All~ Asellidae, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Asellidae")
plot(Asellidae~Dikero.All, pch = 19, data = Binary.Survey)

Dikero.glm2 <- glm(Asellidae~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model

summary(Dikero.glm2) # p = 0.5 not significant

#####################################################################
# Baetidae
#####################################################################


boxplot(Dikero.All~ Baetidae, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Dikero")
plot(Baetidae~Dikero.All, pch = 19, data = Binary.Survey)

Dikero.glm3 <- glm(Baetidae~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model

plot(Dikero.glm3, which = 4)   # cooks distances - 33 is influential remove

Binary.Survey3 <- Binary.Survey[(-33),]    # remove influential point - run model again
Dikero.glm3 <- glm(Baetidae~Dikero.All, family=binomial, data=Binary.Survey3)
plot(Dikero.glm3, which = 4)   # that's all cool now!

summary(Dikero.glm3)      # p = 0.004793


xs<-seq(0,500,l=1000)
Baetidae.predict <- predict(Dikero.glm3, type="response", se=T, newdata=data.frame(Dikero.All=xs))
# Produce base plot
plot(Baetidae~Dikero.All, data=Binary.Survey3, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Baetidae.predict$fit~xs, type="l", col="gray")
```

```
lines(Baetidae.predict$fit+Baetidae.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Baetidae.predict$fit-Baetidae.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence", 2, line=3)
axis(2,las=1)
mtext("D.H abundance",1, line=3)
mtext("Baetidae",3, line=3)
axis(1)
box(bty="l")
text(500,0.9, expression(paste(R^2 == 0.4)), pos = 2)
text(500,0.8, "Effect Size = 0.5415076 ", pos = 2)
text(500,0.7, "P = 0.004793 ", pos = 2)

exp(cbind(OR = coef(Dikero.glm3),confint(Dikero.glm3))) # calculate OR and CIs
1-(Dikero.glm3$dev / Dikero.glm3$null)   # R2 = 0.4070425

0.9801287/1.81   # Effect size = 0.5415076

#############################################################################
#Chironomidae
#############################################################################
boxplot(Chironomidae~ Diker.Haem.Y.N, data = Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Chironomidae")
plot(Chironomidae~Diker.Haem.Y.N, pch = 19, data = Survey)

Dikero.glm4 <- glm(Diker.Haem.Y.N~Chironomidae, family=binomial, data=Survey)    # Run Model

summary(Dikero.glm4) # p = 0.4 not significant

#############################################################################################
# Diptera larva - excluding chironomidae and simulidae
#############################################################################################
boxplot(Diptera.lavra..~ Diker.Haem.Y.N, data = Survey,
        col = "red", xlab = "Presence / Absence D.H",
        ylab = "Diptera.lavra..")
plot(Diptera.lavra..~Diker.Haem.Y.N, pch = 19, data = Survey)

Dikero.glm4 <- glm(Diker.Haem.Y.N~Diptera.lavra.., family=binomial, data=Survey)    # Run Model
```

```
plot(Dikero.glm4, which = 4)   # cooks distance points are alright

summary(Dikero.glm4)      # p = 0.00347



xs<-seq(0,40,l=1000)
Dikero.predict <- predict(Dikero.glm4, type="response", se=T, newdata=data.frame(Diptera.lavra..=xs))
# Produce base plot
plot(Diker.Haem.Y.N~Diptera.lavra.., data=Survey, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Dikero.predict$fit~xs, type="l", col="gray")
lines(Dikero.predict$fit+Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Dikero.predict$fit-Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence of D. haemobaphes", 2, line=3)
axis(2,las=1)
mtext("Diptera larvae (excluding Chironomidae and Simulidae)",1, line=3)
axis(1)
box(bty="l")      # graph looks quite good!
text(40,0.9, expression(paste(R^2 == 0.25)), pos = 2)
text(40,0.8, "Effect Size = 4919101 ", pos = 2)
text(40,0.7, "P = 0.00347 ", pos = 2)
mtext("Diptera Larvae",3, line=3)

exp(cbind(OR = coef(Dikero.glm4),confint(Dikero.glm4)))
1-(Dikero.glm4$dev / Dikero.glm4$null)

0.8903572/1.81 # effectsize = 0.49

#########################################################
# ECNOMIDAE
#########################################################



boxplot(Dikero.All~ Ecnomidae, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Ecnomidae")
plot(Ecnomidae~Dikero.All, pch = 19, data = Binary.Survey)

Dikero.glm5 <- glm(Ecnomidae~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model
```

```
plot(Dikero.glm5, which = 4)   # cooks distance points - 32 is very high

Binary.Survey2 <- Binary.Survey[(-32),]    # remove influential point - run model again
Dikero.glm5 <- glm(Ecnomidae~Dikero.All, family=binomial, data=Binary.Survey2)
plot(Dikero.glm5, which = 4)   # that's all cool now!



summary(Dikero.glm)   # p = 0.0329



xs<-seq(0,500,l=1000)
Ecnomidae.predict <- predict(Dikero.glm5, type="response", se=T, newdata=data.frame(Dikero.All=xs))
# Produce base plot
plot(Ecnomidae~Dikero.All, data=Binary.Survey2, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Ecnomidae.predict$fit~xs, type="l", col="gray")
lines(Ecnomidae.predict$fit+Ecnomidae.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Ecnomidae.predict$fit-Ecnomidae.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence", 2, line=3)
axis(2,las=1)
mtext("D.H abundance",1, line=3)
mtext("Ecnomidae",3, line=3)
axis(1)
box(bty="l")
text(400,0.9, expression(paste(R^2 == 0.77)), pos = 2)
text(400,0.8, "Effect Size = 0.4526869 ", pos = 2)
text(400,0.7, "P = 0.0329 ", pos = 2)

exp(cbind(OR = coef(Dikero.glm5),confint(Dikero.glm5)))
1-(Dikero.glm5$dev / Dikero.glm5$null)   # R2 = 0.7651455

0.8193633/1.81 # effect size = 0.4526869

##############################################################################################
# All elmidae
##############################################################################################
```

```
boxplot(All.Elmidae~ Diker.Haem.Y.N, data = Survey,
        col = "red", xlab = "Presence / Absence D.H",
        ylab = "All.Elmidae")
plot(All.Elmidae~Diker.Haem.Y.N, pch = 19, data = Survey)

Dikero.glm6 <- glm(Diker.Haem.Y.N~All.Elmidae, family=binomial, data=Survey)    # Run Model



plot(Dikero.glm6, which = 4)   # cooks distance points are alright

summary(Dikero.glm6)    # p = 0.04730



xs<-seq(0,300,l=1000)
Dikero.predict <- predict(Dikero.glm6, type="response", se=T, newdata=data.frame(All.Elmidae=xs))
# Produce base plot
plot(Diker.Haem.Y.N~All.Elmidae, data=Survey, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Dikero.predict$fit~xs, type="l", col="gray")
lines(Dikero.predict$fit+Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Dikero.predict$fit-Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence of D. haemobaphes", 2, line=3)
axis(2,las=1)
mtext("Elmidae abundance",1, line=3)
axis(1)
box(bty="l")
text(120,0.5, expression(paste(R^2 == 0.097)), pos = 2)
text(120,0.4, "Effect Size = 0.5466848 ", pos = 2)
text(120,0.3, "P = 0.04730 ", pos = 2)
mtext("Elmidae ",3, line=3)

exp(cbind(OR = coef(Dikero.glm6),confint(Dikero.glm6)))
1-(Dikero.glm6$dev / Dikero.glm6$null)  #R2 = 0.09667966

0.9894995/1.81 # Effect size = 0.5466848

#############################################################################
# Epheremerellidae
#############################################################################
```

```
boxplot(Dikero.All~ Ephemerellidae, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Dikero")
plot(Ephemerellidae~Dikero.All, pch = 19, data = Binary.Survey)

Dikero.glm7 <- glm(Ephemerellidae~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model



plot(Dikero.glm7, which = 4)   # cooks distance 0.8 should be ok...



summary(Dikero.glm7)   # p = 0.02024



xs<-seq(0,500,l=1000)
Ephemerellidae.predict <- predict(Dikero.glm7, type="response", se=T, newdata=data.frame(Dikero.All=xs))
# Produce base plot
plot(Ephemerellidae~Dikero.All, data=Binary.Survey, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Ephemerellidae.predict$fit~xs, type="l", col="gray")
lines(Ephemerellidae.predict$fit+Ephemerellidae.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Ephemerellidae.predict$fit-Ephemerellidae.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence", 2, line=3)
axis(2,las=1)
mtext("D.H abundance",1, line=3)
mtext("Ephemerellidae",3, line=3)
axis(1)
box(bty="l")
text(500,0.9, expression(paste(R^2 == 0.29)), pos = 2)
text(500,0.8, "Effect Size = 0.5404264 ", pos = 2)
text(500,0.7, "P = p = 0.02024 ", pos = 2)

exp(cbind(OR = coef(Dikero.glm7),confint(Dikero.glm7))) #
1-(Dikero.glm7$dev / Dikero.glm7$null)    #R2 = 0.2914287

0.9781717/1.81 # effect size = 0.5404264

##########################################################################################
##########
```

```
# Ephemeridae
##############################################################################################
##


boxplot(Dikero.All~ Ephemeridae, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Dikero")
plot(Ephemeridae~Dikero.All, pch = 19, data = Binary.Survey)


Dikero.glm8 <- glm(Ephemeridae~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model


plot(Dikero.glm8, which = 4)   # cooks distance good


summary(Dikero.glm8)  # p = 0.0838


xs<-seq(0,500,l=1000)
Ephemeridae.predict <- predict(Dikero.glm8, type="response", se=T, newdata=data.frame(Dikero.All=xs))
# Produce base plot
plot(Ephemeridae~Dikero.All, data=Binary.Survey, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Ephemeridae.predict$fit~xs, type="l", col="gray")
lines(Ephemeridae.predict$fit+Ephemeridae.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Ephemeridae.predict$fit-Ephemeridae.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence", 2, line=3)
axis(2,las=1)
mtext("D.H abundance",1, line=3)
mtext("Ephemeridae",3, line=3)
axis(1)
box(bty="l")
text(500,0.6, expression(paste(R^2 == 0.11)), pos = 2)
text(500,0.5, "Effect Size = 0.5586638 ", pos = 2)
text(500,0.4, "P = 0.0838 ", pos = 2)




exp(cbind(OR = coef(Dikero.glm8),confint(Dikero.glm8))) #   Dikero.All  1.0111814 1.0017696 1.028593
1-(Dikero.glm8$dev / Dikero.glm8$null) # r2 0.1120439


1.0111814/ 1.81 # effect size 0.5586638
```

```
##############################################################################################
#Gammarus
##############################################################################################
boxplot(Gammarus.Pulex~ Diker.Haem.Y.N, data = Survey,
        col = "red", xlab = "Presence / Absence D.H",
        ylab = "Gammarus")
plot(Gammarus.Pulex~Diker.Haem.Y.N, pch = 19, data = Survey)

Dikero.glm9 <- glm(Diker.Haem.Y.N~Gammarus.Pulex, family=binomial, data=Survey)    # Run Model

plot(Dikero.glm9, which = 4)   # cooks distance points are alright

summary(Dikero.glm9)  # p = 0.019796



xs<-seq(0,240,l=1000)
Dikero.predict <- predict(Dikero.glm9, type="response", se=T, newdata=data.frame(Gammarus.Pulex=xs))
# Produce base plot
plot(Diker.Haem.Y.N~Gammarus.Pulex, data=Survey, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Dikero.predict$fit~xs, type="l", col="gray")
lines(Dikero.predict$fit+Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Dikero.predict$fit-Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence of D. haemobaphes", 2, line=3)
axis(2,las=1)
mtext("Gammarus.Pulex abundance",1, line=3)
axis(1)
box(bty="l")      #
mtext("Gammarus",3, line=3)
text(100,0.4, expression(paste(R^2 == 0.15)), pos = 2)
text(100,0.3, "Effect Size = 0.5415248 ", pos = 2)
text(100,0.2, "P = 0.0198 ", pos = 2)

exp(cbind(OR = coef(Dikero.glm9),confint(Dikero.glm9)))
1-(Dikero.glm9$dev / Dikero.glm9$null)  # r2 = 0.1507212

0.9801598/1.81   # effect size = 0.5415248
```

```
##############################################################################
#Glossiphoniidae
##############################################################################
boxplot(Dikero.All~ Glossiphoniidae, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "dikero")
plot(Glossiphoniidae~Dikero.All, pch = 19, data = Binary.Survey)

Dikero.glm10 <- glm(Glossiphoniidae~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model

summary(Dikero.glm10) # p = 0.3 not significant



##############################################################################
# Hydrobiidae
##############################################################################

boxplot(Hydrobiidae~ Diker.Haem.Y.N, data = Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Hydrobiidae")
plot(Hydrobiidae~Diker.Haem.Y.N, pch = 19, data = Survey)

Dikero.glm11 <- glm(Diker.Haem.Y.N~Hydrobiidae, family=binomial, data=Survey)    # Run Model

summary(Dikero.glm11) # p = 0.2507 not significant



##############################################################################
# Hydropsychidae
##############################################################################


boxplot(Hydropsychidae~ Diker.Haem.Y.N, data = Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Hydropsychidae")
plot(Hydropsychidae~Diker.Haem.Y.N, pch = 19, data = Survey)

Dikero.glm12 <- glm(Diker.Haem.Y.N~Hydropsychidae, family=binomial, data=Survey)    # Run Model

summary(Dikero.glm12)    # p = 0.53
```

```
################################################################################
# Hydroptilidae

boxplot(Hydroptilidae~ Diker.Haem.Y.N, data = Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Hydroptilidae")
plot(Hydroptilidae~Diker.Haem.Y.N, pch = 19, data = Survey)

Dikero.glm13 <- glm(Diker.Haem.Y.N~Hydroptilidae, family=binomial, data=Survey)    # Run Model

summary(Dikero.glm13)    # p = 0.036292



plot(Dikero.glm13, which = 4)   # cooks distance points are alright



xs<-seq(0,70,l=1000)
Dikero.predict <- predict(Dikero.glm13, type="response", se=T, newdata=data.frame(Hydroptilidae=xs))
# Produce base plot
plot(Diker.Haem.Y.N~Hydroptilidae, data=Survey, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Dikero.predict$fit~xs, type="l", col="gray")
lines(Dikero.predict$fit+Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Dikero.predict$fit-Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence of D. haemobaphes", 2, line=3)
axis(2,las=1)
mtext("Hydroptilidae abundance",1, line=3)
axis(1)
box(bty="l")      #
mtext("Hydroptilidae",3, line=3)
text(70,0.9, expression(paste(R^2 == 0.12)), pos = 2)
text(70,0.8, "Effect Size = 0.5218994 ", pos = 2)
text(70,0.7, "P =  0.036292 ", pos = 2)

exp(cbind(OR = coef(Dikero.glm13),confint(Dikero.glm13)))
1-(Dikero.glm13$dev / Dikero.glm13$null)  # r2 =  0.1187734

0.944638/1.81   # effect size 0.5218994

################################################################################
```

```
# Limnephilidae
#############################################################################
boxplot(Dikero.All~ Limnephilidae, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Dikero")
plot(Limnephilidae~Dikero.All, pch = 19, data = Binary.Survey)

Dikero.glm14 <- glm(Limnephilidae~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model

summary(Dikero.glm14) # p = 0.125

#############################################################################################
#Oligochaeta
#############################################################################################



boxplot(Dikero.All~ Oligochaeta, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Dikero")
plot(Oligochaeta~Dikero.All, pch = 19, data = Binary.Survey)

Dikero.glm15 <- glm(Oligochaeta~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model

summary(Dikero.glm15)   # p = 0.390

#############################################################################
# Polycentropodidae

boxplot(Dikero.All~ Polycentropodidae, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Dikero")
plot(Polycentropodidae~Dikero.All, pch = 19, data = Binary.Survey)

Dikero.glm16 <- glm(Polycentropodidae~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model

plot(Dikero.glm16, which = 4)   # cooks distance good  0.7 - accepted



summary(Dikero.glm16)  # p = 0.0665
```

```
xs<-seq(0,500,l=1000)
Polycentropodidae.predict <- predict(Dikero.glm16, type="response", se=T, newdata=data.frame(Dikero.All=xs))
# Produce base plot
plot(Polycentropodidae~Dikero.All, data=Binary.Survey, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Polycentropodidae.predict$fit~xs, type="l", col="gray")
lines(Polycentropodidae.predict$fit+Ephemeridae.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Polycentropodidae.predict$fit-Ephemeridae.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence", 2, line=3)
axis(2,las=1)
mtext("D.h abundance",1, line=3)
mtext("Polycentropodidae",3, line=3)
axis(1)
box(bty="l")
text(500,0.9, expression(paste(R^2 == 0.13)), pos = 2)
text(500,0.8, "Effect Size = 0.5456393 ", pos = 2)
text(500,0.7, "P = 0.0665 ", pos = 2)


exp(cbind(OR = coef(Dikero.glm16),confint(Dikero.glm16))) #   Dikero.All  0.9876072 0.9699198 0.9974865
1-(Dikero.glm16$dev / Dikero.glm16$null)

0.9876072/1.81 # effect size 0.5456393
##############################################################################################
# Psychomidae


boxplot(Dikero.All~ Psychomyiidae, data = Binary.Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Dikero")
plot(Psychomyiidae~Dikero.All, pch = 19, data = Binary.Survey)

Dikero.glm17 <- glm(Psychomyiidae~Dikero.All, family=binomial, data=Binary.Survey)    # Run Model


plot(Dikero.glm17, which = 4)   # cooks distance good  0.6 - leave for now...


summary(Dikero.glm17)  # p = 0.774
```

```
##############################################################################################
# Rhyacophilidae
##############################################################################################

boxplot(Rhyacophilidae~ Diker.Haem.Y.N, data = Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Rhyacophilidae")
plot(Rhyacophilidae~Diker.Haem.Y.N, pch = 19, data = Survey)

Dikero.glm18 <- glm(Diker.Haem.Y.N~Rhyacophilidae, family=binomial, data=Survey)    # Run Model


plot(Dikero.glm18, which = 4)   # cooks distance good


summary(Dikero.glm18)  # p= 0.062


xs<-seq(0,17,l=1000)
Dikero.predict <- predict(Dikero.glm18, type="response", se=T, newdata=data.frame(Rhyacophilidae=xs))
# Produce base plot
plot(Diker.Haem.Y.N~Rhyacophilidae, data=Survey, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Dikero.predict$fit~xs, type="l", col="gray")
lines(Dikero.predict$fit+Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Dikero.predict$fit-Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence of D. haemobaphes", 2, line=3)
axis(2,las=1)
mtext("Rhyacophilidae abundance",1, line=3)
axis(1)
box(bty="l")      #
mtext("Rhyacophilidae",3, line=3)
text(16,0.9, expression(paste(R^2 == 0.11)), pos = 2)
text(16,0.8, "Effect Size = 0.4172986 ", pos = 2)
text(16,0.7, "P = 0.062 ", pos = 2)

exp(cbind(OR = coef(Dikero.glm18),confint(Dikero.glm18)))
1-(Dikero.glm18$dev / Dikero.glm18$null)  # r2 =  0.1118611

0.7553104/1.81   # effect size 0.4172986
```

```
##############################################################################################
# Simulidae
##############################################################################################

boxplot(Simulidae~ Diker.Haem.Y.N, data = Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Simulidae")
plot(Simulidae~Diker.Haem.Y.N, pch = 19, data = Survey)

Dikero.glm19 <- glm(Diker.Haem.Y.N~Simulidae, family=binomial, data=Survey)    # Run Model



plot(Dikero.glm19, which = 4)   # cooks distance good



summary(Dikero.glm19) # p = 0.96

##############################################################################################
# Sphaeriidae
##############################################################################################
boxplot(Sphaeriidae~ Diker.Haem.Y.N, data = Survey,
        col = "red", xlab = "Presence / Absence ",
        ylab = "Sphaeriidae")
plot(Sphaeriidae~Diker.Haem.Y.N, pch = 19, data = Survey)

Dikero.glm20 <- glm(Diker.Haem.Y.N~Sphaeriidae, family=binomial, data=Survey)    # Run Model



plot(Dikero.glm20, which = 4)   # cooks distance good



summary(Dikero.glm20)  #p = 0.36
##############################################################################################

##Corophiidae

## Analysis of the interaction between corphiidae and Dh is not possible due to the sample size and number of zeros in
the data - let's plot a
## barplot just to show what could be going on...
## boxplot is better - more information
```

```
library(ggplot2)

#p1<-ggplot(Survey, aes(x=factor(Site), y=D..haemobaphes)) + stat_summary(fun.y="mean", geom="bar", fill = "red")

#p2<-ggplot(Survey, aes(x=factor(Site), y=Corophidae)) + stat_summary(fun.y="mean", geom="bar", colour= "blue", fill = "blue")

library(gridExtra)
#pushViewport(viewport(layout = grid.layout(2, 1)))
#print(p1, vp = viewport(layout.pos.row = 1, layout.pos.col = 1))
#print(p2, vp = viewport(layout.pos.row = 2, layout.pos.col = 1))

box1<- ggplot(Survey, aes(x=Site, y=D..haemobaphes)) + geom_boxplot(fill = "red") + xlab ("Site") +
  ylab("D.haemobaphes Abundance") + ggtitle ("Comaprison of D.haemobaphes and Corophiidae abundances")
box2<- ggplot(Survey, aes(x=Site, y=Corophidae), fill="blue") + geom_boxplot(fill = "blue") + xlab ("Site") +
  ylab("Corophiidae Abundance")

pushViewport(viewport(layout = grid.layout(2, 1)))
print(box1, vp = viewport(layout.pos.row = 1, layout.pos.col = 1))
print(box2, vp = viewport(layout.pos.row = 2, layout.pos.col = 1))



##############################################################################################
### Comparison of Gammaridae communities
#p3<- ggplot(Survey, aes(x=factor(Site), y=Gammarus.Pulex)) + stat_summary(fun.y="mean", geom="bar", fill = "springgreen3")
#p4<- ggplot(Survey, aes(x=factor(Site), y=Dikero.Gam.ratio)) + stat_summary(fun.y="mean", geom="bar", fill = "turquoise4")



#pushViewport(viewport(layout = grid.layout(3, 1)))
#print(p1, vp = viewport(layout.pos.row = 1, layout.pos.col = 1))
#print(p3, vp = viewport(layout.pos.row = 2, layout.pos.col = 1))
#print(p4, vp = viewport(layout.pos.row = 3, layout.pos.col = 1))
box5<- ggplot(Survey, aes(x=Site, y=D..haemobaphes)) + geom_boxplot(fill = "red") + xlab ("Site") +
  ylab("D.haemobaphes Abundance") + ggtitle ("Comaprison of Gammaridae communities")
box3<- ggplot(Survey, aes(x=Site, y=Gammarus.Pulex)) + geom_boxplot(fill = "springgreen3") + xlab ("Site") +
  ylab("G. pulex abundance")
box4<- ggplot(Survey, aes(x=Site, y=Dikero.Gam.ratio), fill="blue") + geom_boxplot(fill = "turquoise4") + xlab ("Site") +
  ylab("D.haemobaphes / G.pulex ")
```

```
box6<- ggplot(Survey, aes(x=Site, y=Percent.DH), fill="blue") + geom_boxplot(fill = "turquoise4") + xlab ("Site") +
  ylab("Proportion of D.haemobaphes (%)")

pushViewport(viewport(layout = grid.layout(3, 1)))
print(box5, vp = viewport(layout.pos.row = 1, layout.pos.col = 1))
print(box3, vp = viewport(layout.pos.row = 2, layout.pos.col = 1))
print(box6, vp = viewport(layout.pos.row = 3, layout.pos.col = 1))

############################
```

8.10 Appendix 10. Sorensen Similarity index Analysis R script

```
########################################
# SORENSEN ANALYSIS
###################################


rm(list=ls())   # Clear Memory

# set working directory

getwd()  # check directory location


Survey <- read.csv(file.choose())

str(Survey)
Survey$number <- as.factor(Survey$number )

Survey["All.Elmidae"] <- NA # That creates the new column named "MY_NEW_COLUMN" filled with "NA"
Survey$All.Elmidae <- Survey$Elmidae..Adult. + Survey$Elmidae.Elmis..larva. + Survey$Elmidae.Limnius..larva.
str(Survey, list.len=1000)

###### Similarity Tests ###################

Survey$number <- NULL
Survey$Number.of.Families <- NULL
Survey$TOTAL <- NULL
```

```
Survey$BMWP <- NULL
Survey$Site <- NULL
Survey$BMWP <- NULL
Survey$ASPT <- NULL
Survey$Dikero.Gam.ratio <- NULL
Survey$All.Elmidae <- NULL
Survey$Diker.Haem.Y.N <- NULL
head(Survey)

require("vegan")        # trying outa new package - I'm a bit lost.

#test1 <-vegdist(Pilot2, method="euclidean", binary=FALSE, diag=FALSE, upper=FALSE,
#na.rm = FALSE)                                    #    it did something but I don't have a clue what????

#plot(test1)
#summary(test1)

test2 <-vegdist(Survey, method="jaccard", binary=FALSE, diag=FALSE, upper=FALSE,
            na.rm = FALSE)
summary(test2)
plot(test2)
hist(test2)
# jaccard test measures similarity based on presence absence thereore the sites have similar bugs based
# on this analysis - must use an index that considers abundance...



test3<- vegdist(Survey, method="bray", binary=TRUE, diag=FALSE, upper=FALSE,
            na.rm = FALSE)
plot(test3)
summary(test3)
boxplot(Survey$D..haemobaphes ~ test3 )
hist(test3)



test4 <- vegdist(Pilot2, binary=TRUE)              # these two methods produce a sorensen similarity index...
mean(test4)
plot( Pilot2,11)

boxplot(test4)
str(test4)
```

```
test5 <- betadisper(d = test3, group = Survey$D..haemobaphes)
plot(test5)
boxplot(test5)
aov(test5)
aovtest<- anova(test5)

str(test5)

####### THE FOLLOWING CODE LOOKS AT THE SIMILARITY INDICES AT EACH SITE #########
## Bray-Curtis distances between samples
dis <- vegdist(Survey, method="bray", binary=TRUE, diag=FALSE, upper=FALSE,
         na.rm = FALSE)
## First 16 sites grazed, remaining 8 sites ungrazed
groups <- factor(c(rep(1,6), rep(2,10),rep(3,4),rep(4,6),rep(5,4),rep(6,2),rep(7,8)), labels = c("A","B","C","D","E","F","G"))
## Calculate multivariate dispersions
mod1 <- betadisper(dis, groups)
mod1
## Perform test
anova(mod1)
## Permutation test for F
permutest(mod1, pairwise = TRUE)
## Tukey's Honest Significant Differences
(mod1.HSD <- TukeyHSD(mod1))
plot(mod1.HSD)
## Plot the groups and distances to centroids on the
## first two PCoA axes
plot(mod1)
## can also specify which axes to plot, ordering respected
plot(mod1, axes = c(3,1), main = "Site Comparison of Sorensen Distances")
text(0.3,0.05, "P = 0.01463 ", pos = 2)
text(0.1,-0.08, "A ", pos = 2)
text(0.03,-0.07, "B ", pos = 2)
text(-0.11,-0.05, "C ", pos = 2)
text(-0.0,-0.11, "D ", pos = 2)
text(-0.02,0.08, "E ", pos = 2)
text(0.03,0.23, "F ", pos = 2)
text(0.03,0.19, "G ", pos = 2)

## Draw a oxplot of the distances to centroid for each group
boxplot(mod1)
## `scores` and `eigenvals` also work
```

```
scrs <- scores(mod1)
str(scrs)
head(scores(mod1, 1:4, display = "sites"))
# group centroids/medians
scores(mod1, 1:4, display = "centroids")
# eigenvalues from the underlying principal coordinates analysis


### DO THE SAME THING BUT COMPARE SITES WITH AND WITHOUT DH

newtable <-  Survey[ order(-Survey[,14], Survey[,1]), ]
newtable$row.names <- NULL
str(newtable)

## Bray-Curtis distances between samples
dis <- vegdist(newtable, method="bray", binary=TRUE, diag=FALSE, upper=FALSE,
         na.rm = FALSE)

## First 16 sites grazed, remaining 8 sites ungrazed
groups <- factor(c(rep(1,30), rep(2,10)), labels = c("Present","Absent"))

## Calculate multivariate dispersions
mod2 <- betadisper(dis, groups)
mod2
## Perform test
anova(mod2) # 0.00323
## Permutation test for F
permutest(mod2, pairwise = TRUE)
## Plot the groups and distances to centroids on the
## first two PCoA axes
plot(mod2)
## can also specify which axes to plot, ordering respected
plot(mod2, axes = c(3,1), main = "Presence/Absence of D. haemobaphes - Soresen distances")
text(0.3,0.0, "P = 0.00323 ", pos = 2)
text(0.1,-0.02, "Pres. ", pos = 2)
text(0.12,0.14, "Abs. ", pos = 2)

## Draw a boxplot of the distances to centroid for each group
boxplot(mod2)
## `scores` and `eigenvals` also work
scrs <- scores(mod2)
```

```
str(scrs)
head(scores(mod2, 1:4, display = "sites"))
# group centroids/medians
scores(mod2, 1:4, display = "centroids")



### The above worked well -


#######  ATTEMPTING TO DO THE SAME WITH DH AS INTEGER ######  Not POSSIBLE
## Bray-Curtis distances between samples
dis <- vegdist(Survey, method="bray", binary=TRUE, diag=FALSE, upper=FALSE,
            na.rm = FALSE)
## First 16 sites grazed, remaining 8 sites ungrazed
groups <- as.integer(Survey$D..haemobaphes)
## Calculate multivariate dispersions
mod3 <- betadisper(dis, groups)
mod3
## Perform test
anova(mod3)       # anova proves highly significant but as DH is a integer plotting isn't possible
## Permutation test for F
permutest(mod3, pairwise = TRUE)
## Tukey's Honest Significant Differences
#(mod3.HSD <- TukeyHSD(mod3))
#plot(mod3.HSD)
## Plot the groups and distances to centroids on the
## first two PCoA axes
#plot(mod3)
## can also specify which axes to plot, ordering respected
#(mod3, axes = c(10,100))
## Draw a boxplot of the distances to centroid for each group
#boxplot(mod3)
## `scores` and `eigenvals` also work
#scrs <- scores(mod3)
#str(scrs)
#head(scores(mod3, 1:4, display = "sites"))
# group centroids/medians
#scores(mod3, 1:4, display = "centroids")



##  EXTRACT SORENSEN DATA AND ANALYSE AS A PLOT (lOGISTIC REGRESSION)
```

```
head(mod3)
str(mod3)



df <- data.frame(distances = mod3$distances, group = mod3$group)
#  mod.aov <- aov(distances ~ group, data = df)                    # creates data frame with sorenson data
#summary(mod.aov)              # anova significant

#op <- par(mfrow = c(2, 2))     #
#plot(mod.aov)            # but plots terrible can't accept
#par(op)

#str(df)

df["Dikero"] <- NA # That creates the new column named "MY_NEW_COLUMN" filled with "NA"
df$Dikero <- as.numeric(as.character(df$group))
as.numeric(as.character(df$Dikero))                          # converts dikero data to numeric
#plot (distances ~ Dikero, data = df)
#scatterplot (log10(distances+0.01) ~ log10(Dikero +0.01), data = df)            #scaterplot - looks crap

#mod.aov2 <- aov(log10(distances+0.01) ~ log10(Dikero +0.01), data = df)  # log both axis
#summary(mod.aov2)
#op <- par(mfrow = c(2, 2))     #
#plot(mod.aov2)                    # again plots are terrible
#par(op)

# try a logistic regression with this data - presence absence of DH. This is required because the sorensen data is
proportional.

df["Diker.Haem.Y.N"] <- NA # That creates the new column named "MY_NEW_COLUMN" filled with "NA"
df$Diker.Haem.Y.N <- df$Dikero
str(df, list.len=1000)

df$Diker.Haem.Y.N[df$Diker.Haem.Y.N>0] <-1  # changes the column to binary



#explore

boxplot(distances~ Diker.Haem.Y.N, data = df,
        col = "red", xlab = "Presence / Absence D.H",
```

```
      ylab = "sorensen differences")
plot(distances~Diker.Haem.Y.N, pch = 19, data = df)

Dikero.glm2 <- glm(Diker.Haem.Y.N~distances, family=binomial, data=df)    # Run Model



plot(Dikero.glm2, which = 4)   # cooks distance points are alright

summary(Dikero.glm2)   # p = 0.03387

xs<-seq(0,0.18,l=1000)
Dikero.predict <- predict(Dikero.glm2, type="response", se=T, newdata=data.frame(distances=xs))
# Produce base plot
plot(Diker.Haem.Y.N~distances, data=df, xlab="", ylab="", axes=F, pch=16)
# Plot fitted model and 95% CI bands
points(Dikero.predict$fit~xs, type="l", col="gray")
lines(Dikero.predict$fit+Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
lines(Dikero.predict$fit-Dikero.predict$se.fit ~ xs, col="gray", type="l", lty=2)
# Axes titles
mtext(" presence/absence of D. haemobaphes", 2, line=3)
axis(2,las=1)
mtext("sorenson distances",1, line=3)
axis(1)
box(bty="l")      #
text(0.07,0.6, expression(paste(R^2 == 0.1238)), pos = 2)
text(0.07,0.5, "P = 0.03387 ", pos = 2)
mtext("Change in Sorensen Distances with presence/absence of D. haemobaphes",3, line=3)

exp(cbind(OR = coef(Dikero.glm2),confint(Dikero.glm2)))
1-(Dikero.glm2$dev / Dikero.glm2$null)
8.324370e-08/1.81
```

8.11 Appendix 11. GLMM R script

```
############### Working GLMM  ######################

rm(list=ls(all=TRUE))

library(lme4)

invlogit <- function(x){1/(1+exp(-x))}

Dune=read.table("data/SurveyGLM2.txt", header=TRUE,sep=" ")
```

```
Dune0= Dune  # we will modify Dune starting from the original, Dune0
colnames(Dune)
glPredict <- function(fm1, newdat, conf = 95) {
  # Predicts occurrence probability with confidence limits from an glmer object at
  #  the points provided as rows of newdat
  # fm1 = glmer  object
  # newdat = data frame with values for predictors for which prediciton must be made
  # confidence value (in %)
  # for related code see package ez
  # Value:
  #  y, lo, hi = prediction with confidence limits on link scale
  #  p, plow, phigh =   occurrence probability with confidence limits
  frac = 1 - (100-conf)/200
  mm = model.matrix(terms(fm1),newdat)
  y = mm %*% fixef(fm1)  # prediction on link scale
  Var <- Matrix::diag(mm %*% tcrossprod(vcov(fm1),mm)) # variance on link scale
  lo = y-qnorm(frac)*sqrt(Var)
  hi = y+qnorm(frac)*sqrt(Var)
  newdat$y = y
  newdat <- data.frame(newdat, ylo = lo, yhi = hi,
                       p = invlogit(y), plow = invlogit(lo), phigh = invlogit(hi))
  newdat
}

##############################################################################

# Table 4 Quantitative environmental variable; Factor trait  - THIS WORKED WOOOOOOOOOP!
#####################################################################
Dune = Dune0
Trophic.NicheLab = c("Scraper","Predator","Gatherer/Collector", "Shredder","Filter/Collector")
#Dune$Trophic.Niche= cut(Dune$Trophic.Niche, breaks = c(1,2,3,4,5), labels= Trophic.NicheLab)
```

```
print(fm3<-glmer(y~D..haemobaphes+Trophic.Niche+(1+D..haemobaphes|sp)+(1|site)
                 , family=binomial, Dune),corr=FALSE)


newdat <- expand.grid( D..haemobaphes=seq(0,500,length.out=1000),Trophic.Niche=Trophic.NicheLab, y =0)
newdat <- glPredict(fm3, newdat)


par(mfrow=c(2,3))
for ( j in Trophic.NicheLab){
  data.f<- subset( newdat , Trophic.Niche %in% j)
  x<- data.f$D..haemobaphes
  plot(0,0,ylim=c(0,1),xlim=range(x),ylab="Pr(sp presence)" ,xlab="D..haemobaphes" ,
       yaxs="i" , main="",type="n")
  mtext(paste("",j ),    font= 2, col= "black" )
  polygon(c(x, rev(x)), c(data.f$phigh, rev(data.f$plow)),col = 'gray', border = FALSE)
  points(x, data.f$p, type='l',lwd=2.5)
}


summary(fm3)
capture.output(summary(fm3),file="nicheGLMM.txt")
plot(fm3)


library('arm')



binnedplot(fitted(fm3),resid(fm3), main = "Binned residual trophic Niche")    # binned plot seems to be the only way to validate binary models - ask Jon...



#############################################################################################################

# plot all mean trends in one graph - not too helpful - pred and scraper over lay each other. gatherer and filterer overlay each other

plot(0,0,ylim=c(0,1),xlim=range(x),ylab="Pr(sp presence)" ,xlab="D..haemobaphes" ,
```

```
    yaxs="i" , main="",type="n")



for ( j in Trophic.NicheLab){
  data.f<- subset( newdat , Trophic.Niche %in% j)
  x<- data.f$D..haemobaphes
  lines(0,0,ylim=c(0,1),xlim=range(x),ylab="Pr(sp presence)" ,xlab="D..haemobaphes" ,
       yaxs="i" , main="",type="n")


  points(x, data.f$p, type='l',lwd=2.5, col = "2")
}



##########################################################################################################
#######



##############################################################################
## EFFECT SIZE

1-var(residuals(fm3))/(var(model.response(model.frame(fm3)))) # very negative... is this a true effect size?


##############################################################################
 ### Calculate pseudo R2

r2.corr.mer <- function(m) {
  lmfit <-  lm(model.response(model.frame(m)) ~ fitted(m))
  summary(lmfit)$r.squared
}

r2.corr.mer(fm3)    # 0.62
```

8.12 Appendix 12. Experimental ANOVA Validation Plots

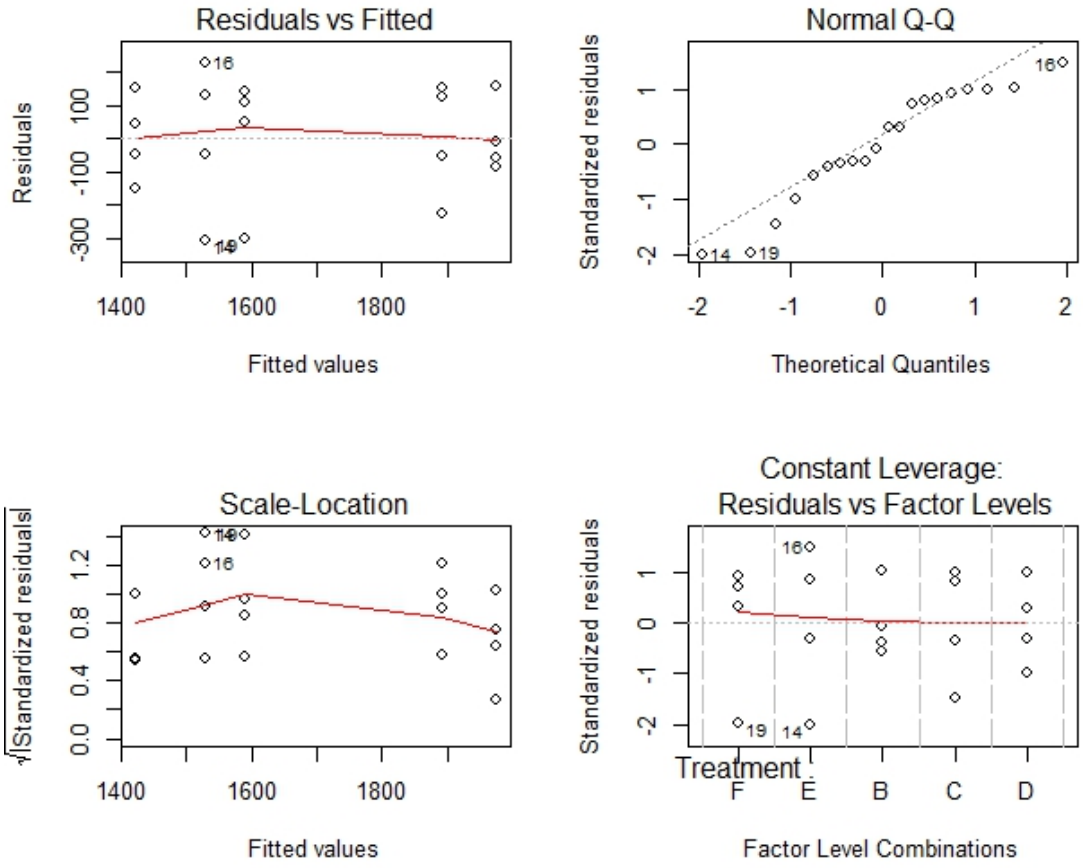

Figure 35: ANOVA validation plots for experimental enclosure study.

8.13 Appendix 13. Experimental Analysis R script

```
################################################

## Experiment Analysis Code

################################################


rm(list=ls())   # Clear Memory


# set working directory


getwd()  # check directory location
```

```
Experiment <- read.csv(file.choose())


str(Experiment)

Experiment["Start.number.of.DikeroF"] <- NA # That creates the new column named "MY_NEW_COLUMN" filled with "NA"

Experiment$Start.number.of.DikeroF <- Experiment$Start.number.of.Dikero.

Experiment$Start.number.of.DikeroF <- as.factor( Experiment$Start.number.of.DikeroF)


boxplot(Mass.consumed ~ Start.number.of.DikeroF, data = Experiment, ylab = "leaf litter consumption (mg)", xlab = "Start number of D. haemobaphes")



scatterplot (Mass.consumed ~ Start.number.of.Dikero., data = Experiment, ylab = "leaf litter consumption (mg)", xlab = "leaf litter consumption (mg)")


aov1<- aov(Mass.consumed ~ Start.number.of.DikeroF, data = Experiment)   # run anova


op <- par(mfrow = c(2, 2))     # plots look good

plot(aov1)

par(op)


summary(aov1)  # significant - 0.0017

#capture.output(summary(aov1),file="ExperimentFAOV.txt")



lm1<- lm(Mass.consumed ~ Start.number.of.DikeroF, data = Experiment)

summary(lm1)                              # run the linear model for the graph


plot(Mass.consumed ~ Start.number.of.DikeroF, data = Experiment, ylab = "Leaf litter decomposition (mg)", xlab = "Number of D. Haemobaphes out of 30",

    col = "brown1", main = "Change in decomposition with increasing proportion of D. Haemobaphes", cex.main=1.0) # plot

text(5,2000, "P = 0.0017 ", pos = 2)
```

```
text(5,1900, "Eta-squared = 0.6633259 ", pos = 2)

abline(lm1)    # final image                                                    # add the line

########################################

# attempt at final plot with control treatment included...



Experiment2 <- read.csv(file.choose())


str(Experiment2)


Experiment2$Treatment3 <- factor(Experiment2$Treatment2, as.character(Experiment2$Treatment2))


Experiment2$Control <- as.factor( Experiment2$Control)


str(Experiment2)

str(Experiment3)


Experiment3 <- Experiment2


Experiment3 <- Experiment3[(-21),]

Experiment3 <- Experiment3[(-21),]

Experiment3 <- Experiment3[(-21),]

Experiment3 <- Experiment3[(-21),]


lm2<- lm(Mass.consumed ~ Treatment3, data = Experiment3)

summary(lm2)


library(ggplot2)


ggplot(Experiment2, aes(x=Treatment3, y=Mass.consumed, fill=Control, )) + geom_boxplot() + xlab
("Dikerogammarus:Gammarus Ratio") +

 ylab("Leaf mass consumed (mg)") + ggtitle ("Leaf litter Decomposition under changing Gammaridae community")   #
plot without lm
```

```
#########################################################

## Effect size



library(lsr)



etaSquared(lm1)                 # calculates eta squared value 0.663

etaSquared(aov1, anova = TRUE)


confint(aov1)
```

8.14 Appendix 14. Full Survey Data Set

Macro Invertebrates	BMWP value	Site A1	Site A2	Site A3	Site A4	Site A5	Site A6	Site B1	Site B2	Site B3	Site B4	Site B5	Site B6	Site B7	Site B8	Site B9	Site B10	Site C1	Site C2	Site C3
Acroloxidae	6							13	8	15	13	19	9	1		15	10	37	28	3
Aphelocheiridae	10																			
Asellidae	3																1			
Astacidae	8																			
Baetidae	4	100	118	58	90	28	55	92	69	16	83	44	68	75	20	26	19	52	87	42
Bithyniidae	3	1																		
Brachycentridae	10		1	1																
Caenidae	7																			
Calopterygidae	8																			
Chironomidae	2	7		38	140	16	7	432	224	136	624	168	99	117	195	432	162	252	246	246
Chrysomelidae	0																			
Corophidae	6																			
Crangonyctidae	6																		4	
D. haemobaphes	0									4	2	1		1	2		1	3	3	17
Diptera lavra *	5	33	2	17	35	11	33	38	13	25	25	2	24	25	6	21	21	10	4	12
Economidae	8	2	35	56	38	22	10	28	51	7	6	11	11	28	2	10	10	43	60	14
Elmidae (Adult)	2	15	19	25	28	14	39	118	97	46	24	10	99	74	7	16	23	6	10	47
Elmidae Elmis (larva)	2	11	18	28	7	24	12	68	58	24	6	12	76	44	1	18	28	4	6	3
Elmidae Limnius (larva)	2	24	27	96	36	93	52	79	48	33	29	91	104	41	8	40	123	6	1	136
Ephemerellidae	10	3	6	12	6	7	3	3	6		4		2	1		1	4	14	11	6
Ephemeridae	10			2		6	1									1	1			
Epobdellidae	3															1				
Gammarus Pulex	6	154	61	115	36	42	43	79	84	17	53	25	214	102	8	20	31	14	73	22
Glossiphoniidae	3							1	1					1			1			8
Hydrachnida	0					2			1				2	2		2	3			
Hydrobiidae	3		1	27	38	54	252	7	22	17	1	21	32	12		4	1	47	13	9
Hydrophilidae	5																	2		
Hydropsychidae	5	9	36	31	36	25	18	387	355	64	198	177	71	239	47	245	257	381	336	177
Hydroptilidae	6	2	30	66	42	25	23	9	4	3	9		12	4	3	5	2	19	11	6
Leptoceridae	10																			
Limnephilidae	7	1		9	4		1	2	1	1	9	9	5	2		3	7			2
Lymnaeidae	3																			1
Neritidae	6	1		1								1							1	
Oligochaeta	0	6		2	1	10	6		9	3		4		1			3		1	24
Physidae	3																			
Planorbidae	3																		1	
Polycentropodidae	7				1					1	1		1	2		5	4	7	43	2
Psychomyiidae	8			3	2					1		1					2	1		1
Rhyacophilidae	7	2	2	1				8	16	1	2	2	3	7	1	4	1	1	3	
Simulidae	5	19	136	26	24	24	3	160	11	1	60	3	11	13	21		1	28	462	1
Sphaeriidae	3	106	1	5	12	6	17	15	15	17	9	51	36	38	1	24	43		7	32
Unionoidea	6																			
Valvatidae	3																			1
Number of Families		16	13	19	16	15	15	16	18	18	17	17	17	20	12	18	23	17	20	21
TOTAL		496	493	619	576	409	575	1539	1093	432	1135	682	879	830	322	893	759	927	1411	812
BMWP		79	74	107	81	69	76	82	82	84	86	77	86	89	53	94	110	89	94	93
ASPT		4.9375	5.692307692	5.631578947	5.0625	4.6	5.066666667	5.125	4.555556	4.666667	5.058824	4.529412	5.058824	4.45	4.416667	5.222222	4.782609	5.235294	4.7	4.428571
* not Chimomid or Simulidae																				

Table 4: Full Survey Data set Part 1.

Site C4	Site D1	Site D2	Site D3	Site D4	Site D5	Site D6	Site E1	Site E2	Site E3	Site E4	Site F1	Site F2	Site G1	Site G2	Site G3	Site G4	Site G5	Site G6	Site G7	Site G8
17	1	12				7	4	15	1	6			16	17	1	6		4	95	87
														1			1		2	2
48	14	1		1								2								
																				1
32	61	89	21	11	15	73	2	1	1				1		1		2	1	2	
4	1					4														
														1	2					
				2																
	1																			
420	402	46	396	438	492	504	33	29	15	16	25	64	108	100	234	156	76	30	15	9
															1					
											45	25	5	8	217	17	22	11	82	82
44							6				1							1		
4	3	1	17	6	5	14	3	8	2	6	98	485	475	180	217	136	41	25	166	215
7	2	7		1	7	13		13	7	7		1	3	13	3	27	3	2	6	11
20	7	11	1	6		7	9	1	1	1		2								
37	17	61	47	63	90	92	11	11	10	8	1		1	1	7	1		2		
23	19	70	17	12	7	13	7	9	2	1	2		5	2	12	2	4	3	3	5
63	54	81	71	67	79	90	91	58	9	80		5	37	20	37	6	16	38	32	32
3	7	8	2	4	2	12	4	3								1				
1			1	1	2	2	5			1		12	33	5	14		2	7	21	4
19																				
149	19	50	38	16	4	25	3	8	2	3	28	127	31	16	62	10	8	23	79	82
18		7		1	8	1													1	
2																				
15	19	53	1	3	19	190	34	249	246	210		4	59	73	156	32	270	378	270	324
352	60	149	21	45	11	183	3	2	28	13	1	3	7	3	1	11	2		3	7
28	24	6	20	10	10	24	48	20		4		1	1	1	8	1	2		1	
	1			1	2		4	1											1	
8		6	1	1	6	6								7	1			2	1	
1																				
	2	10				2								1	3					
3		12	1		4		16	6	6	12	3	14	2		4			18	24	18
1																				
7						1														
8	8	5	5	6	8	1	7		2	1				1						
		1					4	6		13	1			11	1	1	3	3	3	15
		7	1		1	2														
14	5	53	5				23	30	384	2	4	6		17	30	17		3		5
6	39	86	15	5	32	23	1	2		2	1	3	5	9	19		8	2	13	13
				1	1														1	
1			4		1															
28	20	22	18	20	20	21	19	17	13	16	11	15	14	19	20	13	14	16	20	16
1355	766	832	685	701	806	1289	318	472	716	386	210	754	789	487	1031	424	460	553	821	912
119	102	106	85	107	99	106	101	83	53	76	43	64	58	107	94	64	70	73	102	79
4.25	5.1	4.818182	4.722222	5.35	4.95	5.047619	5.315789	4.882353	4.076923	4.75	3.909091	4.266667	4.142857	5.631579	4.7	4.923077	5	4.5625	5.1	4.9375

Table 5: Full survey data set part 2

8.14 Appendix 14. GLMM Summary

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']

Family: binomial (logit)

Formula: y ~ D..haemobaphes + Trophic.Niche + (1 + D..haemobaphes | sp) + (1 | site)

Data: Dune

AIC	BIC	logLik	deviance	df.resid
1321.3	1375.5	-650.6	1301.3	1670

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.8013	-0.3486	-0.1548	0.3143	6.0384

Random effects:

Groups	Name	Variance	Std.Dev.	Corr
sp	(Intercept)	5.646e+00	2.376182	
	D..haemobaphes	4.512e-05	0.006717	-0.26
site	(Intercept)	3.224e-01	0.567792	

Number of obs: 1680, groups: sp, 42; site, 40

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	1.032532	0.909201	1.136	0.25610	
D..haemobaphes	-0.002638	0.001503	-1.755	0.07921	.
Trophic.NicheGatherer/Collector	0.024194	1.073571	0.022	0.98202	
Trophic.NichePredator	-3.481148	1.307137	-2.663	0.00774	**
Trophic.NicheScraper	-2.106319	1.196957	-1.760	0.07845	.
Trophic.NicheShredder	-3.471517	1.655965	-2.096	0.03605	*

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:

	(Intr)	D..hmb	T.NG/C	Trp.NP	Trphc.NchSc
D..haembphs	-0.112				
Trphc.NcG/C	-0.820	-0.013			
Trphc.NchPr	-0.699	-0.006	0.582		
Trphc.NchSc	-0.743	0.000	0.624	0.534	
Trphc.NchSh	-0.528	-0.001	0.446	0.370	0.403